

1: Markets

“De gustibus non est disputandum”

August 23, 2019

1.1: Introduction: What Is Justice?

At the start of Plato’s most important dialog, *The Republic*, Socrates asks Cephalus, and old man, for a definition of justice, the subject of the dialogue.¹ Cephalus says justice consists of telling the truth and respecting other people’s property. Reasonable enough, but Socrates sees a problem:

“What you say is very fine indeed, Cephalus,” I said. “But as to this very thing, justice, shall we so simply assert that it is the truth and giving back what a man has taken from another, or is to do these very things sometimes just and sometimes unjust?”

Take this case as an example of what I mean: everyone would surely say that if a man takes weapons from a friend when the latter is of sound mind, and the friend demands them back when he is mad, one shouldn’t give back such things, and the man who gave them back would not be just, and moreover, one should not be willing to tell someone in this state the whole truth.”

“What you say is right,” he said.

“Then this isn’t the definition of justice, speaking the truth and giving back what one takes.”

From ancient Greece to the present day, an important function of government is providing incentives for “speaking the truth and giving back what one takes.” That is what courts do when they enforce contracts— promises to do something in exchange for something else— and punish theft— taking other people’s property. We all feel that it is good to be able to make contracts and do what you want with the things you own. It’s also very common, however, to think that the economy would work better if the government intervened to require contracts to be a certain way and to restrict what people can do with their property. In particular, we think that maybe the economy would work better if the government intervened to control the prices people charge when selling their property or the kinds of property they are allowed to sell. Or, one might think that the government should own businesses, and manufacturing in particular, rather than have them be owned by individual people as private companies or as corporations, and run the businesses for the public good rather than for profit. This is the idea of *socialism*, either in its communist form or in its milder democratic forms. And indeed, although in the United States (unlike in countries such as the United Kingdom) the government ran few businesses, at one time government regulation of prices was common. From the 1930’s to the 1980’s, however, there was a considerable diminution in price regulation, and after experiencing deregulation of industries such as air travel, trucking, and telephone sales, the public is less interested in government regulation of prices. The virtues of *laissez faire* pricing are part of the conventional wisdom. At the same time, other forms of regulation have vastly increased— regulation not of prices, but of what kinds of products businesses can sell, how they sell them, and how they make them.

¹ Plato , *The Republic*, translated by Allan Bloom (1968) at 331c.

In this chapter we will start by thinking about what it means for an economy to work well, and what it means for a regulation to help the economy instead of hurting it. Economists do not go much beyond Cephalus's common-sense notion of what justice means. We take the satisfaction of human wants to be society's objective, with contracts and property rights as crucial means to that end. After the *Republic* passage I just quoted, Cephalus admits things are complicated and goes off to perform religious sacrifices. He hands off the argument to his sons, younger men, and in the remaining 98% of *The Republic* is devoted to exploring what justice really means, looking at it from all angles but in failing to come to a definite conclusion. Economists leaves the discussion of the most difficult questions to other people. To be sure, much of what government does, and much of what justice is revolves around the question of the meaning of life. The discipline of economics just tries to figure out how to make a society wealthy, meaning how to organize production and consumption to satisfy people's desires, taking those desires as given rather than saying what desires are good and what are bad. The Latin phrase "De gustibus non est disputandum" sums up the attitude: "About tastes there is no arguing." In this chapter you will see how to pin down the meaning of satisfaction as a goal and you will see that private markets do surprisingly well at achieving it even with a government that does very little to help it.

1.2 Surplus Maximization

How should we decide whether a regulation is good, or bad? The first step is to choose some valuation rule—some way to compare two policies, and, if possible, to assign numerical values to how good each policy is.

Suppose, for example, that we are trying to decide whether a rule requiring the arsenic level in drinking water to be less than 23 parts per billion is a good rule or not. A strident environmentalist might say that the more stringent the rule, the better—that a level of 23 parts per billion is better than 30, but 4 parts per billion would be even better. Someone else might say that cost should be considered too, and that reducing the level to 4 parts per billion would cost more than the entire budget of the city government, requiring taxes to double. Still another person might say that "the government is best which governs least," so there should be no rule at all—the policy should be **laissez faire**, French for "let them do it".

The standard valuation rule used by economists is **surplus maximization**. The idea is simple. Add up how much each person who likes the regulation would pay to have it, and subtract out how much each person who dislikes the regulation would need to be paid to accept it. If the resulting number is positive, adopt the regulation.²

²There are variations on the rule—we could ask both people how much they would pay to get their preferred rule, or ask both how much they would accept to not have it, but we won't go into those subtleties.

A concrete example is the best way to understand surplus maximization. Suppose Anderson and Brown want a stricter arsenic regulation and would pay up to \$30 and \$70 to get it, whereas Corman and Daniels don't want it, and would require payments of at least \$20 and \$10 to balance out their dissatisfaction with the new regulation. Since supporters would pay \$100 and opponents would accept \$30, adopting the regulation maximizes surplus.

The arsenic regulation maximizes surplus even if no payments actually take place. Corman and Daniels end up worse off, but their loss is less than the gain of Anderson and Brown. If payments do take place, on the other hand everybody can be made better off after the new rule is adopted. If we adopt the regulation and make Anderson pay \$25 to Corman and Brown pay \$25 to Daniels, all four of them are happy that the deal went through, compared to the initial situation. We call this a **Pareto improvement**, after the economist Wilfried Pareto who came up with the criterion that a policy should be adopted if it makes some people better off and no one worse off.

FIGURE 1.1
WILFRIED PARETO
(1848-1923)



whether the world is better off if you're dead or if you're alive. Doctors stick to the task of healing the sick, and try not to think about whether it's worth healing them, which is one reason why they are not the best people to ask about ways to control health care costs. Similarly, don't ask economists questions like whether legalizing marijuana is a good idea; they will have opinions, but your own ethical beliefs are just as good a guide.

Surplus maximization is more useful as a criterion than Pareto improvement, though. We can not only say that the arsenic regulation is a Pareto improvement; we can put a number on the size of the improvement. The supporters would pay \$100 and the opponents would pay \$30, so the surplus rises by \$70. In economics, we call this the

Pareto optimality seems like an obvious idea— too obvious for Pareto to deserve getting his picture as Figure 1.1. It isn't as simple as it looks, though. Whose well-being should count when we define "some people" and "no one"?³ What if Person X gets happier if Person Y's freedom is restricted?⁴ Those are more good questions for the philosophers. Economists evade them by saying that although we have personal opinions as to the answers, for policy purposes we'll just take the answers you give us and use them to do objective reasoning. After all, if you ask your doctor whether you should take blood pressure medication, you don't expect him to supply an answer to the question of

³See Figure 1.2 and Lawrence A. Hansen "Animal Research: Groupthink in Both Camps," *The Chronicle of Higher Education*, (November 7, 2010).

⁴See the Story of Prude and Lewd in: Amartya Sen , "The Impossibility of a Paretian Liberal," *The Journal of Political Economy*, 78 (1): 152-157 (January-February 1970).

“value” created by the regulation. We have actually made precise how good the regulation is, and done it in an objective way. That’s surprising—how can a political issue, on which these four people disagree, be made objective? The trick is that it is objective only from the point of view of a neutral observer. Our basic data is the subjective values each of the four people—Anderson, Brown, Corman, and Daniels— put on the policy.⁵ But once we have those values, we have an objective method of putting them all together. It is like determining the value of a corporation. Once we have the sales figures, prices, and costs, determining the value of a company is an objective process (though not easy, since we’re trying to forecast future profits, and forecasters will disagree). Future sales, though, depend on how much consumers like the company’s product, which depends on their subjective values. Placing a dollar value on the regulation’s improvement combines subjective data objectively in the same way.

FIGURE 1.2
PARETO OPTIMALITY



(a) Anderson



(b) Brown



(c) Corman



(d) Daniels

In practice, most government policies are not Pareto improvements. Almost always, at least one person loses from a change in policy, and the policies do not include side payments to losers. Such payments would be highly impractical to make. Imagine trying to tax people whose preferred presidential candidate won so as to pay compensation to the voters who lost. If we insist that every new policy make nobody worse off, we’d rule out virtually all new policies. Surplus maximization is a more practical criterion. Whichever arsenic regulation is chosen, someone is going to be unhappy, so it seems sensible to try to evaluate how strongly each side feels. Moreover, since governments make policies on a vast number of subjects, anyone who loses from use of the surplus maximization criterion to change one policy is likely to win with its use for other policies.

Surplus analysis is also useful because even if maximizing total surplus is not the

⁵“Four people” introduces another problem, illustrated by Figure 1.2. Think too of the 1981 Gary Larson *Far Side* cartoon of three men and a dog in a boat with the caption “Fair is fair, Larry... We’re out of food, we drew straws, you lost.”

goal, it is useful to figure out who is helped and who is hurt by a policy, and by how much. This will help us predict the political battlelines, as we will see in later chapters.

1.3: The Surplus from a Single Transaction

Let us now apply the idea of surplus maximization to a market transaction instead of government policy. Buyer Brown approaches seller Smith and asks if Smith will sell a bottle of whisky for \$10. Smith agrees, and the whisky changes hands. Is the transfer of the bottle from Smith to Brown a good thing?

Surplus maximization says it is. Since Brown offered a price of \$10, we know his willingness to pay was at least that high, and probably higher. Suppose it is \$15. Since Smith accepted the price of \$10, we know his value for the bottle was no more than that. Suppose it is \$8. Using the figures of \$15 and \$8, the net benefit from Smith giving the bottle to Brown is \$7, an increase in total surplus, an increase in value. If we didn't know the number \$8 and \$15, the gain might be more than \$7, or less, but since both parties agreed to the trade we know that Smith's value must be below Brown's and the gain in value is positive. Total surplus has risen.

We are calling this valuation rule "surplus maximization," but it is also called **the Kaldor-Hicks criterion** after the economists who first defined it, or **potential Pareto improvement**, or **wealth maximization**, or **economic efficiency**. These names all say things about the rule. Kaldor-Hicks tells the names of the two English economists who came up with it in 1939.⁶

FIGURE 1.3
THE ECONOMISTS BEHIND KALDOR-HICKS EFFICIENCY



Nicholas Kaldor (1908-1986)



John Hicks (1904-1989)

"Potential Pareto improvement" is an idea we've already discussed. If surplus has increased, that means that potentially we could have a Pareto improvement by undertaking the new policy or transaction and then compensating any losers. Of course, in the case of a voluntary transaction, there are no losers, so the Pareto improvement is not just potential.

⁶Nicholas Kaldor, "Welfare Propositions in Economics and Interpersonal Comparisons of Utility," *The Economic Journal*, 49 (195): 549-552 (1939). John Hicks, "The Foundations of Welfare Economics," *The Economic Journal*, 49 (196): 696-712 (1939).

How about “wealth maximization”? If Brown values the bottle of whisky at \$15 and Smith at \$8, then moving the bottle from Smith to Brown at a price of \$10 has a benefit of \$5 for Brown and \$2 for Smith. The effect of the trade on their satisfaction is the same as if Brown and Smith had never met, but by a miracle \$5 suddenly appeared in Brown’s pocket and \$2 in Smith’s. Thus, the name “wealth maximization”: a surplus-maximizing trade increases the dollar amount at which people value their possessions. That is the right way to think of society’s wealth, as a concept. It is crucial to realize that this is *not* the same as the market value of everybody’s possessions, which is the usual way to measure wealth since we observe market prices but not the most people are willing to pay. Wealth in the sense of total surplus measures how much people value their possessions, not how much they had to pay for them or how much they could sell them for. My surplus for example, would rise if I sold a \$200,000 house which I was barely willing to buy at that price and used the money to buy a different \$200,000 house for which I would have been willing to pay \$300,000. My well-being has risen after I make the change, but my dollar wealth as measured at market prices has not.⁷

The term “economic efficiency” has a similar motivation. In everyday language, something is efficient if it achieves its goal with minimum effort or expense. In economics, efficiency refers to whether we could increase total surplus without increasing the amount of inputs. If we can, then the original situation is inefficient, and the size of the inefficiency is the amount of potential surplus lost. If the problem is **productive inefficiency**, the goods could be produced more cheaply. If the problem is **allocative inefficiency**, total satisfaction— that is, total surplus— could be increased by rearranging who has which goods without increasing the total amount. Moving the bottle from Smith to Brown is necessary for allocative efficiency because it increases satisfaction, even though there are no more bottles of whisky than before the transfer. It is as if the economy had found a technology that increased the amount of output by \$7.

We’ve started with an easy case for our value criterion. It may seem obvious that it is good for Smith to sell the bottle of whisky to Brown. But notice that surplus maximization does not “assume the answer,” as would a criterion such as “voluntary trades are good.” Surplus maximization applies generally. It is a multipurpose tool. We can compare any two outcomes by looking at whether surplus has increased, so

⁷The concept of “gross domestic product ” does use prices rather than values to measure well-being. GDP measures how much people pay for goods traded in markets. It tries to get at the idea of how much value is produced, but it’s impossible to measure in practice how much Brown *would* pay for the whisky, and easier to measure how much he *does* pay. Also, GDP measures “value-added” to avoid double counting inputs, so if Smith paid \$7 for corn to make the whisky, GDP would increase by \$3 (= \$10 - \$7) as a result of the trade, and the other \$7 would be counted in GDP by looking at the trade the corn farmer made with Smith.

we don't need one rule for whether voluntary transactions are good, another one for whether government regulations are good and another one for whether innovation is good. Also, surplus maximization allows a numerical value to be put on how good the trade is, rather than just saying it is good or bad.

Since surplus maximization doesn't assume the answer, though, it may not always come up with the answer you want. What if Brown is only ten years old when he buys the whisky? Or what if Brown steals the bottle of whisky from Smith instead of buying it? Surplus maximization says that this, too, is a good thing. The effect on total surplus is exactly the same as the sale at the price of \$10. The sale benefited Brown by \$5 and Smith by \$2, a total of \$7. The theft benefits Brown by \$15 and hurt Smith by \$8, which also makes for a total benefit of \$7. Surplus maximization doesn't care about how much each person gets as a result of the event. All that matters is that the bottle has moved from someone who values it less to someone who values it more.

Most people would say that Brown's theft is a bad thing, not a good one, even though it raises surplus. That's why we have laws against theft. But surplus maximization is morally neutral. It leaves out morality and just looks at personal satisfaction.

You might feel like rejecting surplus maximization since it can reach a perverse result such as theft being a good thing, good, but moral neutrality is one reason economists use the concept. The best moral rule is not controversial in the case of theft, but for many government policies it's unclear how morality would apply or whose moral belief should count.⁸ But what is more important is that using a general principle such as surplus maximization instead of a specific one such as "stealing is sinful" can often (though not always) help us pin down what is bad about bad actions more precisely.

Surplus maximization, in fact, usually agrees with everyday morality when it comes to prescribing what the government should do. Surplus maximization actually does imply that stealing should be illegal and that people should have moral scruples against stealing. In the particular case of Brown and Smith, surplus maximization says that the theft is good. It is good, however, only because we started with a story in which Brown was willing to pay as much as \$15 and Smith was willing to accept as little as \$8. That information is crucial to whether moving the bottle from Smith to Brown was good or bad.

What would happen if we didn't know the particular values Brown and Smith put on the whisky? Suppose the government announced that it was going to simply give the bottle to whoever valued it most, without actually requiring money to change hands. Not knowing the numbers, the government would have to ask each person

⁸Of course, if you ask me, I'd could tell you exactly whose moral beliefs should count—mine, since I hold the correct moral beliefs. But people would rather hear what economists have to say about surplus than what they have to say about sin and virtue, so that's what we talk about.

his value. What would they say? Brown could say he should get the bottle because he wants it so much he would pay \$150 for it if he had to. Smith could say that he, too, really wants that bottle and his value is \$500, even more, so the government shouldn't take it away from him and give it to Brown. They would lie—or at least, they would lie if they didn't have moral scruples. And even if we put aside the possibility of lying, until someone has to actually pay for something he usually doesn't know the absolute most he would pay for it because he hasn't thought much about the question. He hasn't thought about it, because that absolute most number rarely matters. It's enough for me to know that my value for a cup of coffee is greater than the current Starbucks price; I don't have to calculate the most I would pay if I had to.

Without knowing everyone's values, though, how can we use surplus maximization as a policy tool? Sometimes the economist just does his best to try to estimate values by looking at people and thinking how much that kind of person values something, or by asking them and trusting to their honesty. The wonderful thing about market transactions, however, is that market prices force people to reveal their true values. Brown puts his money where his mouth is when he offers \$10 for the bottle. He's not going to exaggerate and say he's willing to pay \$150 if he actually has to pay the money. Smith reveals something about his value too, when he accepts \$10. If he wants to credibly claim his value is \$500, he has to give up on the \$10 sale. When a voluntary transaction takes place, we do not learn the exact values, but we learn a little bit more, and we can deduce that there is a gain in social surplus.

Now let's go back to theft. After a theft, unlike after a voluntary transaction, we do not know that surplus has increased. We learn something—that Brown was willing to go to the trouble of stealing the bottle, and Smith was not willing to go to the trouble of guarding it effectively—but that is not enough to guarantee surplus maximization. Maybe Brown stole the bottle of whisky because he values it at \$3 (unlike in our earlier example, where it was \$15) and the effort of stealing is only a cost of \$1 for him. If Smith's value is \$8, the theft has reduced social surplus.

Notice, too, that using \$3 for Brown and \$8 for Smith, the social surplus is not just -\$5, the difference between Jones's value and Brown's, but -\$6. That's because not only does the theft reduce allocative efficiency, it also creates productive inefficiency: Brown bears that \$1 cost of stealing, which isn't producing any new goods for anyone.

The cost rises even higher if Smith reacts by investing time and money to protect his goods against Brown's thievery. Smith's defensive tactics of buying locks and security guards may be morally more justifiable than Brown's offensive expenditure on burglar tools and night-time outings, but both offense and defense are a drain on society's wealth. Thus, theft, while it may increase surplus in particular cases, has a bad effect overall. Moreover, if the government forbids theft but allows selling, then if Brown really does value the whisky at more than \$8 he can buy the bottle from Smith anyway.

In this way we've derived a reason why theft is bad, a reason based on surplus maximization, rather than having to accept the evil of theft as one on a list of many separate moral rules. The Eighth Commandment says "Thou shalt not steal," but even if you don't believe that God ordained the Ten Commandments, if you accept surplus maximization you come to the same conclusion. We can not only derive "Thou shalt not steal," but "Thou shalt not kill," "Thou shalt not commit adultery," and many other moral rules, which become mere corollaries or surplus maximization. In this way surplus maximization like the Golden Rule: "Love thy neighbor as thyself" (Matthew 19:19). The Golden Rule doesn't say the decisionmaker should use each person's own valuation to allocate goods in the economy, but like surplus maximization it is a general guide to behavior and treats people in an unbiased way.

Think too about what surplus maximization is not. It is not materialism, or economic goals, or maximizing gross domestic product, or making a country as rich as possible, even though it's measured in dollars. Crucially, it's based on how much people value things, not on the things themselves. Suppose Brown owns an old walnut tree like the one in Figure 1.4 and refuses to sell it to Smith for \$8,000 to cut down as timber.⁹ That outcome maximizes surplus, since clearly Brown values the walnut at more than \$8,000 and Smith at less. If Brown sold the tree, however, GDP would rise by \$8,000, even though in a more precise sense "wealth" has fallen.

The most pernicious confusion policymakers have about economic goals is the idea of "creating jobs" and "reducing unemployment". I say "confusion" because although people talk about increased employment as the goal, that isn't really what they're after. If reducing unemployment is really the goal, a quick way to achieve it would be to make unemployment illegal. A law like this was actually passed once in Mississippi. It said:

All freedmen, free negroes and mulattoes in this State, over the age of eighteen years, found on the second Monday in January, 1866, or thereafter, with no lawful employment or business, . . . shall be deemed vagrants, and on conviction thereof shall be fined in a sum not exceeding, in the case of a freedman, free negro or mulatto, fifty dollars, . . . If any freedman, free negro or mulatto shall fail . . . to pay . . . it shall be, and is hereby, made the duty of the sheriff of the proper county to hire out said freedman, free negro or mulatto, to any person who will, for the shortest period of service, pay said fine and forfeiture and all costs . . .¹⁰

Probably the authors of this law, passed while freedmen still were not able to vote, would have said this was for the freedmen's benefit, to prevent them from becoming vagrants. It also helped white employers, though, because it meant that black workers

⁹For a discussion of walnut tree prices, see <http://www.woodweb.com/cgi-bin/forums/forestry.pl?read=512463>. Another tree one might use as an example is the Boole Tree, a giant sequoia, the largest sequoia in a grove and the only large one remaining after logging between 1892 and 1918, saved by the foreman of the operation.

¹⁰Mississippi Vagrant Law, §2 http://wps.ablongman.com/long_longman_lahdemo.1/0,8259,1546454-,00.html. "§" means "Section".

could not be choosy when it came to picking a job. When unemployment is banned, it is illegal to quit without having a new job already in hand. It's bad being unemployed when you want a job, but not so bad as being forced to take the first miserable job you're offered. What people mean when they say they want a job, or want to create jobs, is that they want good job opportunities.

So it isn't that we just want to reduce unemployment. Nor is it that we want to increase *employment*. If we required all women to work and forced students into the workplace by closing colleges, that would raise employment. Nor do we want to guarantee that anybody searching for a job gets one. We could require them to take their best opportunity within six months— fast-food and lawnmowing included— or face penalties. Or, if it is high-paying jobs that are the goal, we could have the government hire people at \$100,000/year to pick dandelions out of lawns. No— what people really mean when they say they want there to be more jobs is that they want there to be more ways that people can productively use their time. In other words, they want to maximize surplus. In the ideal world, people standing idle are not maximizing surplus. In the actual world, we want them to stand idle till they find the right job. There is tremendous movement into and out of employment both in good times and bad. Job movement is ordinarily good, not bad. Someone is worth less in their current job than in some other job, so he switches. An economy cannot have people moving to better jobs unless they leave worse jobs. Ideally, everyone would find a better job instantly, but in reality workers have to search for a job and employers have to search for workers. In a recession people stay unemployed for longer than usual because of a shock to the system. Overbuilding of housing followed by a banking crisis, for example, would mean that a larger number of people (and capital) are in a job this is less productive than another position somewhere in the economy. Finding that new position isn't easy, so unemployment rises. Subsidizing housing jobs would reduce unemployment, but at the cost of building still more empty houses. Jobs voluntarily created by employers, on the other hand, will produce goods that consumers want. Employers, too, need to figure out what those goods are that have higher value than the overbuilt housing. This takes time, and meanwhile both wages and profits suffer, but adjustment is necessary if the economy is to produce goods people want.

1.4: The Surplus in an Entire Market

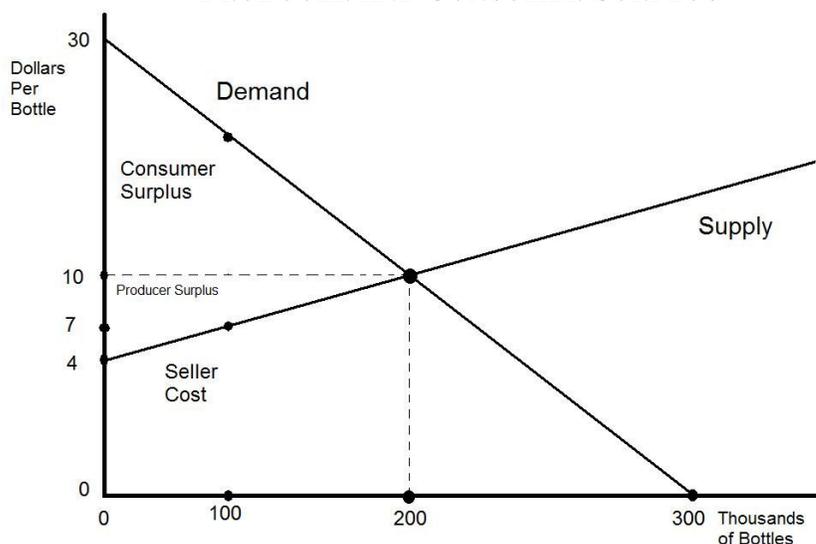
The first step to understanding why the free market maximizes surplus was to understand why the single transaction between Brown and Smith maximizes surplus. We slid over the question of why they traded at a price of \$10 rather than some other price between the \$8 that is the least Smith would accept and the \$15 that is the most Brown would pay. Now we will return to that.

The next step is to look at how the market price is chosen. Let us look at the entire market for whisky. Let the market consist of 300,000 potential buyers, each

of whom would buy at most 1 bottle, and 5,000 potential sellers, each of whom might sell 100 bottles. Buyers vary in their willingness to pay. Some buyers would pay at most \$.01 per bottle, but others would pay as much as \$30. (We say that \$30 is the **reservation price**, the most any buyer would pay.) Sellers vary in their minimum acceptable price. Some would accept as little as \$4, but others would require as much as \$19. Figure 1.4's supply and demand diagrams show this more precisely using the supply and demand equations $P = 30 - 0.1Q^d$ and $P = 4 + 0.03Q^s$ (with quantities of bottles measured in thousands). The demand curve shows that there are 100,000 consumers willing to pay at least \$20 and 200,000 willing to pay at least \$10.

The supply curve in Figure 1.4 shows that there are 1,000 sellers (with a total of 100,000 bottles) willing to take as little as \$7 and 2,000 willing to take as little as \$10. At the price of \$10, the quantity supplied equals the quantity demanded. This price of \$10/bottle, the price generated by market forces, is the **equilibrium price**.

FIGURE 1.4
PRODUCER AND CONSUMER SURPLUS



Why do market forces generate a price of \$10 and not some other price? One might answer “because the supply and demand curves cross at \$10” or “because supply equals demand at \$10,” but those answers are superficial. The real answer must explain why the equilibrium price is where quantity supplied equals quantity demanded.

Think about what would happen if the price were higher—say, \$20 per bottle. Sellers would be delighted to sell everything they had, but only 100,000 consumers would be willing to buy. As a result, some sellers would end up unable to sell. Those frustrated sellers would shave the price to \$19.99, causing buyers to switch to them. That would leave other sellers customerless, and those sellers would shave the price to \$19.98. The price of \$20 per bottle is thus unstable. And so is any other price above \$10.00.

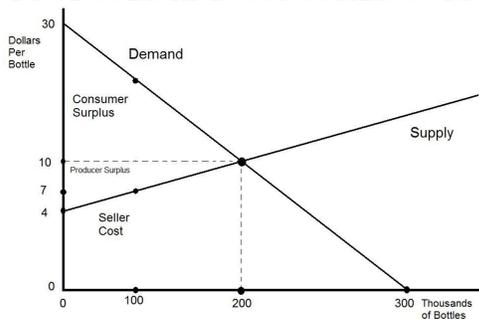
The same reasoning shows that any price below \$10.00 is unstable. At any price below \$10.00, buyers are more eager to buy than sellers are to sell, and some buyers would be unable to find a seller. Suppose the price were \$7.00/bottle. Consumers who couldn't find a willing seller because of the excess demand would offer \$7.01 to outcompete the other buyers. The outcompeted buyers would offer \$7.02 in response, and the process would continue until the price was bid up to \$10.00.

At the price of \$10.00, each buyer willing to buy at that price would find a willing seller, and there would be no incentive to for any buyer to offer a higher price or for any seller to offer a lower price.

This reasoning shows that the free market equilibrium price is stable. That it is stable, however, says nothing about whether it is good. Buyers would always like to pay less and sellers would like to get more, so there are always people ready to complain about the equilibrium price. Philosophers have put much thought into deciding when those complaints might be justified. What is the “just price” for a bottle of cheap whisky, the price that is fair? Economists shy away from the words “just” or “fair,” but they have their own criterion for whether a price is good: does it end up helping to maximize surplus? And as it happens, the equilibrium price results in surplus being maximized.

To show that the equilibrium price is efficient, we need to think about the benefits of the transactions to buyers and sellers. First, calculate the benefits for the equilibrium price of \$10.00 and quantity of 200,000 bottles, bought and sold by the buyers and sellers to the left of the intersection of supply and demand in Figure 1.4. Those are the buyers with the highest gross benefits from buying bottles (from \$30 down to \$10) and the sellers with the lowest costs of parting with them (from \$4 to \$10). (The **gross benefit** is the value to the buyer of suddenly possessing a free bottle. The **net benefit** is the gross value minus the price he has to pay.) You will see immediately that to maximize surplus we need whisky to be bought by these high-valuing buyers and sold by these low-valuing sellers, as happens in the free market. Government allocation of the whisky might not reach the same result. Government allocation would reach a different result if it gave the bottles to the consumers with the lowest income, or the most morally deserving consumers, or the consumers best-connected politically, rather than to the consumers willing to pay the most.

FIGURE 1.4 (REPEATED FROM ABOVE)
PRODUCER AND CONSUMER SURPLUS



The gross benefit to the sellers who are selling those 200,000 bottles is their sales revenue, which is (200,000 bottles) (\$10/bottle) = \$2,000,000. This is not their net benefit, because the sellers place positive value on those bottles, even if their values are not as high as the buyers'. The seller values range from \$4/bottle to \$10/bottle, as shown by the height of the supply curve. Combined, the seller values are the area labelled "Seller Cost" in Figure 1.4, since for typical sellers the value they place on what they are selling is their acquisition or production cost (though the cost might also be an opportu-

nity cost—that they cannot drink the whisky themselves).

We can numerically calculate the size of the seller cost. Geometrically, it is the area of the rectangle \$4/bottle high and 200,000 bottles wide (which is \$800,000) plus the area of the triangle with a height of (\$10/bottle - \$4/bottle) and a width of 200,000 bottles, which is $(1/2) (\$6/\text{bottle})(200,000 \text{ bottles}) = \$600,000$. That sums to a seller cost of \$1,400,000.

Since the sellers' net benefit is their gross benefit (the revenue) of \$2,000,000 minus their lost value (the production cost) of \$1,400,000, their net benefit is \$600,000. This is the area labelled "producer surplus" in Figure 1.4. Producer surplus is the standard name for net seller benefit.¹¹ More formally:

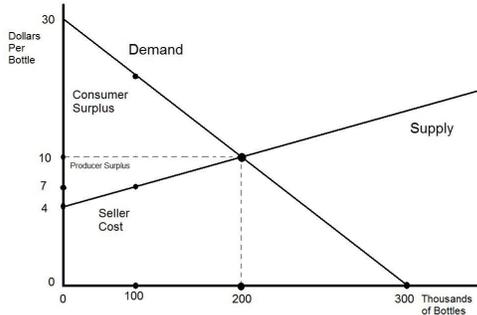
Producer surplus is the sum across sellers of how much they receive in actual prices minus the minimum prices they would accept.

Producer surplus is not the same as profit. Profit as used in ordinary language includes the business's return to its equity capital. Those costs will often show up in the height of the supply curve, since a business will not be willing to keep supplying a given quantity if it is not earning a competitive return to its capital. Instead, the business will shrink or go out of business. On the other hand, in the short run the capital of a business is sunk, like certain its costs (e.g. an office lease that can't be terminated early), so it may be earning positive producer surplus but not be making an accounting profit or a long-run economic profit. Although producer surplus and

¹¹Sometimes the concept of producer surplus is taught a different way, calculating the area of the producer surplus directly as a triangle rather than as the revenue rectangle minus the cost trapezoid. I'm using the slower way here so the reader will understand the concept better.

profit are similar, they are not identical concepts.

FIGURE 1.4 (REPEATED FROM ABOVE)
PRODUCER AND CONSUMER SURPLUS



\$10/bottle) and width 200,000 bottles, which is $(1/2) (\$20/\text{bottle}) (200,000 \text{ bottles}) = \$2,000,000$. Adding up the two areas (which have the same size here— pure coincidence) yields the gross benefit of the consumers, which is $\$4,000,000$.

The net value for the buyers is less than the gross value, because they have to pay the sellers $\$10/\text{bottle}$. This is a payment of $\$2,000,000$ for all 200,000 bottles, so the net value is $\$4,000,000 - \$2,000,000 = \$2,000,000$. In Figure 1.4, this is the area labelled “consumer surplus”.

Consumer surplus is the sum across buyers of the maximum prices they would pay minus the actual prices they do pay.

Why is this called “consumer surplus” rather than “buyer surplus”? Historical accident. I wish we used “buyer surplus” instead, because sometimes the buyers are not individual people, but companies. Even in that situation, though, we refer to the surplus of the buyer as “consumer surplus”.

Adding together the producer surplus and the consumer surplus gives the total surplus created by the existence of this market. When the quantity is 200,000 bottles and the price is $\$10/\text{bottle}$, the total surplus is thus $\$600,000 + \$2,000,000 = \$2,600,000$.

Total surplus is the sum across everyone in society of the net benefits received from the market for a good.

Usually total surplus is just the sum of consumer and producer surplus, but it could also include benefits to taxpayers from taxes buyers and sellers pay, or costs

Now think about the buyers of whisky. Their gross benefit from the 200,000 bottles is the sum of the values for each of the buyers. Some buyers have a value of $\$30$, some $\$29$, some $\$28$, and so forth down to the last buyer who actually makes a purchase, whose value is only $\$10$. (There exist other “buyers” who are inactive, but they will not be getting any benefit, so we can ignore them.) The sum of the values is the area under the demand curve up to 200,000 bottles. This equals the area of the rectangle $\$10/\text{bottle}$ high and 200,000 bottles wide, which is $\$2,000,000$; plus the area of the triangle above it with height $(\$30/\text{bottle} -$

and benefits to external bystanders such as neighbors of the liquor store who don't participate directly (the "externalities" that we will come to in the next chapter).

Notice that consumer and producer surplus don't equal each other. Indeed, often producer surplus is zero. That happens if producers are all so similar that they all have the same cost of production. Usually, though, there is surplus on both sides, but it is unequal. Is this unfair? Hard to say— it depends on what you mean by 'fair'. But if we want prices to signal the marginal benefit of the most reluctant active buyer and the most reluctant active seller, we can't obscure that information by deciding that one side or the other deserves more of the surplus. George Orwell said of the economist F. A. Hayek, "He does not see, or will not admit, that a return to 'free' competition means for the great mass of people a tyranny probably worse, because more irresponsible, than that of the State. The trouble with competitions is that somebody wins them."¹² Orwell, a socialist, didn't like having the free market deciding prices without some wise hand to guide it, and he trusted the government to provide that wisdom. Orwell was a good writer, so he knows how to turn a phrase. But, of course, in the competitions of the marketplace, it is not a bad thing that someone wins. That winner is whichever company provides the best product at the lowest price. The company that wins, gets a profit. But consumers win too, no matter which company is the winner. The "great mass of people" always wins from free competition; it is the companies which wish the competition was replaced by some kind of cartel that let them all charge high prices and earn high profits. And of course the tyranny of the company winning the competition is merely the ability to attract voluntary purchases from consumers. It can't compare with the tyranny of the State, which consists of the ability to take people's money without giving anything in return, and to put them in prison or execute them if they refuse to pay up. Moreover, a company selling products is responsible to consumers in the sense of losing their business if the company does a bad job. The State may or may not be responsible to the citizens, depending on whether the regime is democratic and on the power of the unelected judiciary and bureaucracy, but the penalty the State pays for poor performance is less immediate than that of the business that fails to provide an adequate product. We will return to this in detail later when we discuss government failure and government design.

At any rate, having calculated the total surplus created by the free market, we must now see whether government regulation could do better, as Orwell thinks is the case. Here we come to a remarkable fact: it is the quantity of 200,000 bottles that determines the total surplus, not the price of \$10/bottle. If we change the price, but keep the quantity the same, total surplus won't change.

To see this, suppose that the quantity traded remains 200,000, being bought by the

¹²Review by George Orwell: "The Road to Serfdom by F.A. Hayek / The Mirror of the Past by K. Zilliacus," *The Observer*, 9 April 1944.

same consumers and sold by the same sellers as in the free market, but the price rises to \$20/bottle. This must be backed up by government force, as a two-part regulation. The first part is that those 200,000 consumers must be forced to buy whisky on pain of prison, since many of them would rather not buy at such a high price. The second part is that sellers must be forbidden to reduce their price, on pain of prison, since there won't be enough customers to satisfy all the sellers at that high price and sellers will be tempted to offer discounts.

Calculate the surpluses again with that two-part regulation. The gross consumer benefit and the seller cost have not changed from their free market levels. Since the same people are buying and selling the same 200,000 bottles, the buyers still value the bottles at \$4,000,000 and the sellers still value them at \$1,400,000. All that has changed is that the buyers now pay a much higher price—a total amount of (\$20/bottle) (200,000 bottles) = \$4,000,000—and the sellers get higher revenue. Thus, now the consumer surplus is $(\$4,000,000 - \$4,000,000) = \$0$, and the producer surplus is $(\$4,000,000 - \$1,400,000) = \$2,600,000$. The total surplus is unchanged from its free market level of \$2,600,000; all that has happened is that now the sellers get all of it and the buyers get none of it.

Whatever price is chosen under this two-part regulation, the total surplus will stay the same. When the quantity is fixed, the price is just a transfer from buyer to seller. The total surplus is

$$\begin{aligned} & (\text{Consumer Surplus}) && + && (\text{Producer Surplus}) \\ & (\text{Gross Buyer Benefit} - \text{Price} * \text{Quantity}) && + && (\text{Price} * \text{Quantity} - \text{Seller Cost}) \\ & (\text{Gross Buyer Benefit}) && - && (\text{Seller Cost}) \end{aligned}$$

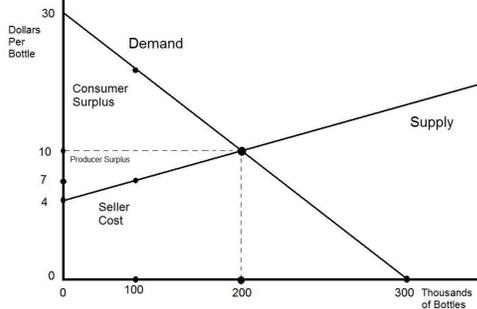
The $\text{Price} * \text{Quantity}$ terms cancel each other, so the total surplus is $(\text{Gross Buyer Benefit} - \text{Seller Cost})$, which does not depend on the price.

The total surplus does depend on the quantity, however, which we were keeping fixed at 200,000 bottles. To see why, consider increasing the quantity. We will need the most reluctant seller to become active, one whose value for a bottle is \$10.00 and who has already sold some but not all of his 100 bottles, since sellers with lower values are already selling. We will need a new buyer to become active too, one whose value is less than \$10.00, since buyers with higher values are already buying. Even if the new seller's value is \$10.00 and the new buyer's value is \$9.99, this new exchange reduces total surplus by \$.01 rather than increasing it. Increasing sales further would reduce total surplus by even more.

How about reducing the quantity below 200,000? If a buyer with a value of \$10.01 stops buying from the most reluctant seller, the seller whose value is \$10.00, then total

surplus drops by \$.01. Reducing the quantity below the free market equilibrium level loses some of the gains from trade between active sellers and active buyers.

FIGURE 1.4 (REPEATED FROM ABOVE)
CONSUMER AND PRODUCER SURPLUS



Thus, we see that the workings of the free market maximize surplus. First, the free market arrives at the equilibrium price, without any government intervention necessary. Then, the equilibrium price elicits an equilibrium quantity which maximizes the sum of producer and consumer surplus.

The fact that this is a two-step process is why the price actually does matter in the end. Earlier, we saw that if the government required the quantity to be 200,000, with the same buyers and sellers as in the free market, then the government could require any level it wanted to for the price, without altering total surplus. The two-part regulation maximized surplus just as well as the free market could, even if the regulation could not do any better. But notice what we took for granted: that the government had detailed information at its disposal and could enforce its regulation costlessly.

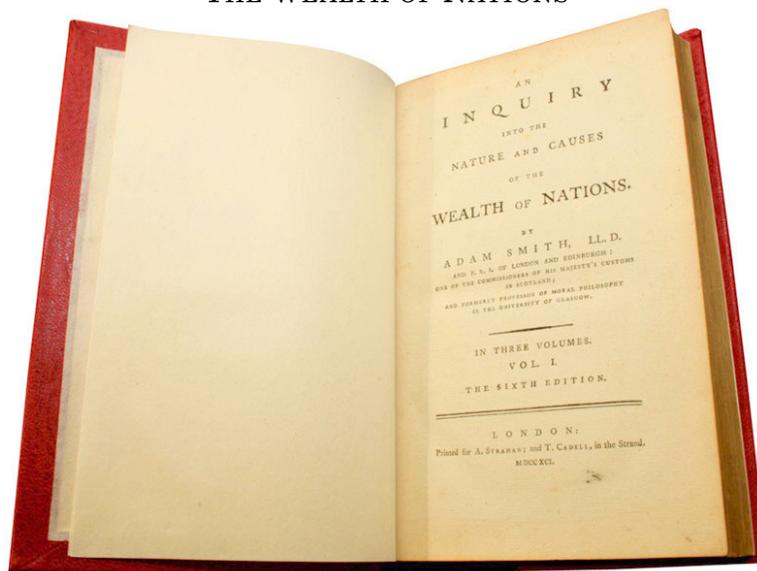
How would the government know that the optimal quantity was 200,000? Supply and demand curves are not written down in books that the government can consult. Economists measure them with intricate statistical procedures using market-generated data. Simply asking people how much they would pay won't work, in the same way it didn't work for Brown and Smith. People don't think hard enough about it if they don't really have to pay, and they don't tell the truth. What would happen if the government said it was going to set the price of a bottle of whisky to one penny per bottle, and asked buyers to step forward if they were one of the 200,000 buyers with values above \$10? Moral scruples about lying aside, all 300,000 consumers would step forward, since all of them would like the chance to buy whisky at that price. If the government then asked which 2,000 sellers had costs below \$10/bottle, no seller would step forward, since none of them want to be forced to sell at a price of a penny per bottle.

The free market, unlike the government, needs very little information. The competitive process of bidding up or discounting moves the price to \$10.00 without any need for people to tell the truth. In an actual market, a buyer has no reason to claim that he would only pay \$5.00 if his true value is \$13.00 and the result of his claim is that he loses the chance to buy at \$10.00. A seller has no reason to claim that his minimum acceptable price is \$11.00 if it is really \$9.00 and his lie will lose him the sale. Economizing on information is a huge advantage of the marketplace, even more

important than the fact that it does not need to pay police to enforce its prices and quantities.¹³

The logic is similar to why the single transaction between Brown and Smith maximized surplus. Brown and Smith are just two of the thousands of participants in the market, and the members of each pair are trading because both benefit. What is different in the case of the market is that anonymous market forces determine that the price will be \$10, whereas in the single transaction example I said that Brown offered Smith a price of \$10 without explaining where that price came from. But in both the single transaction and the market, the essential idea is that if both buyer and seller voluntarily agree to a transaction, it benefits both of them, and it increases total surplus.

FIGURE 1.5
THE WEALTH OF NATIONS



In ordinary economic transactions the free market maximizes surplus without the need for government intervention. This is a modern interpretation of Adam Smith's idea of the Invisible Hand in his 1776 book, *The Wealth of Nations*.¹⁴ In competition

¹³It is for the insight that information aggregation is a key accomplishment of the free market that Frederick Hayek received the Nobel Prize in economics. Every economics student should view the video, "Fear the Boom and Bust: The Original Keynes vs. Hayek Rap Battle," by Russell Roberts, *Cafe Hayek: Where Orders Emerge* blog (January 25, 2010).

¹⁴Smith, Adam, *An Inquiry into the Nature and Causes of the Wealth of Nations* (1st ed. 1776). . A good way to appreciate the beauty of the idea is to read Leonard E. Read, "I, Pencil: My Family Tree as told to Leonard E. Read," *The Freeman* (December 1958).

with each other, producers bid down the price to where the price equals the benefit to the lowest-valuing customer in the market. If a little more were produced, its marginal cost would be greater than the marginal benefit to consumers. If a little less were produced, we would lose a little bit of consumer and producer surplus.

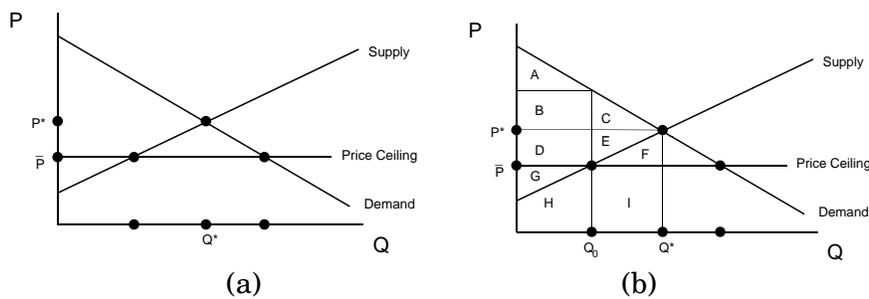
The Invisible Hand is why economists criticize many regulations. Markets generally achieve good results on their own, without the need for anyone to intervene to change the market price or quantity. The idea of the Invisible Hand is exceedingly important, and much of an introductory economics class is devoted to trying to explain it. But economists also recognize that situations do exist in which some premises underlying the reasoning fails, so government regulation could help increase total surplus. These situations of “market failure” will be discussed in the next chapter. First, though, we will look at what happens with misguided government regulation of prices.

1.5: Measuring the Surplus under Rent Control

I have taken some care in explaining how surplus is measured, because it is so fundamental. Now let’s think about how to tackle a given policy change. Rent control is a classic example. Suppose that due to increasing demand the rental rate for a one-bedroom apartment has risen to \$500/month, so the city council passes an ordinance saying that rents can be no more than \$400/month (a **price ceiling**).

In Figure 1.6a, I drew the curves and decided what the equilibria were. In Figure 1.6b, I cut the areas up into triangles and rectangles. Once the areas were cut up and labelled, I could make up lists of rectangles and triangles. We’ll now go through that step by step.

FIGURE 1.6
RENT CONTROL



1. *Draw the supply and demand curves and find the free market equilibrium price and quantity.*

Landlords are supplying apartment units, and potential tenants are demanding them. Figure 1.6a shows how as the price falls the quantity of apartments demanded increases because more people prefer substitute from owning their residences and more want to live in the city instead of the suburbs. Thus, the demand curve slopes down.

Should the supply curve be vertical, or should it slope up? It would be vertical if the quantity supplied of apartments did not depend on the price—**perfectly inelastic**, perfectly insensitive to price. That could be a good approximation, but I've drawn the supply curve slanted, to represent the long-run situation in which the number of apartment does rise with rents because suppliers construct new apartment buildings, rent out condominium units, convert office buildings into apartments, and so forth.

The laissez faire equilibrium is where quantity supplied equals quantity demanded, the intersection of the two curves. I've labelled that price and quantity P^* and Q^* .

2. *Find the market equilibrium price and quantity after the policy or other change.*

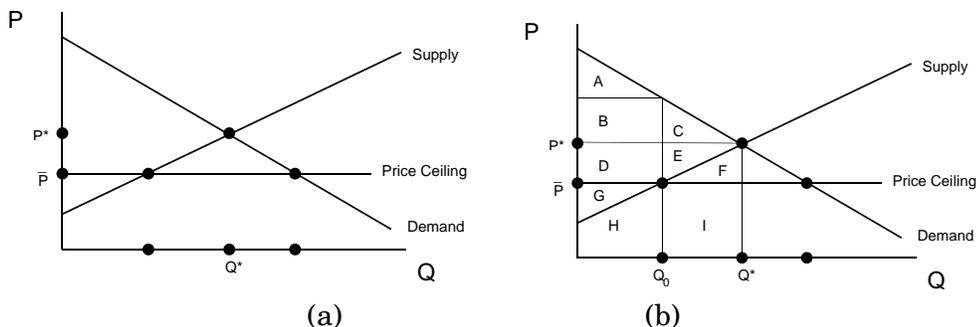
The regulation imposes a price ceiling, which I've labelled as \bar{P} . That price ceiling is less than the free market price, so now the quantity landlords are willing to supply is less than the quantity tenants demand, the situation called **excess demand**. The rent control regulation doesn't say that landlords must rent out their units, only that they can't charge more than \bar{P} , so the quantity rented out will be the landlords' quantity supplied (Q_1 in Figure 1.6a), rather than the tenants' quantity demanded.

3. *Cut up the possible surplus area into as many rectangles and triangles as you think will be relevant, labelling them with letters. It's better to make too many than too few.*

I've done that chopping in Figure 1.6a, which I've drawn separately from Figure 1.6b so you would be less distracted by all the surplus areas when I was explaining the

earlier steps.

FIGURE 1.6 (REPEATED FROM ABOVE)
RENT CONTROL



4. Figure out for the free market which areas are in producer surplus, consumer surplus, effects on third parties (for example, tax revenue), and total surplus.

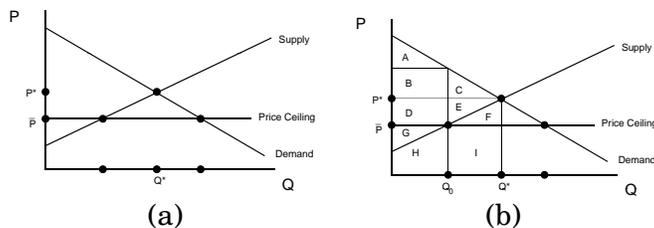
In the free market, the quantity is Q^* , so the surpluses will only cover areas from the quantity 0 to the quantity Q^* . Consumers' total benefit (the tenants') is the area under the demand curve from quantity 0 to quantity Q^* , but they have to pay P^* , so their surplus is the area between the demand curve and the P^* flat line, area $A+B+C$. Producers' total benefit (the landlords') is their revenue, $P^* \cdot Q^*$, but their surplus is the difference between that revenue and their costs, which is the area $D+E+G$ between the supply curve and the P^* flat line. There are no effects on third parties—no tax revenue, for example. The total surplus is therefore $A+B+C+D+E+G$.

$$CS \text{ (laissez faire)} = A+B+C$$

$$PS \text{ (laissez faire)} = D+E+G$$

$$TS \text{ (laissez faire)} = A+B+C+D+E+G$$

FIGURE 1.6 (REPEATED FROM ABOVE)
RENT CONTROL

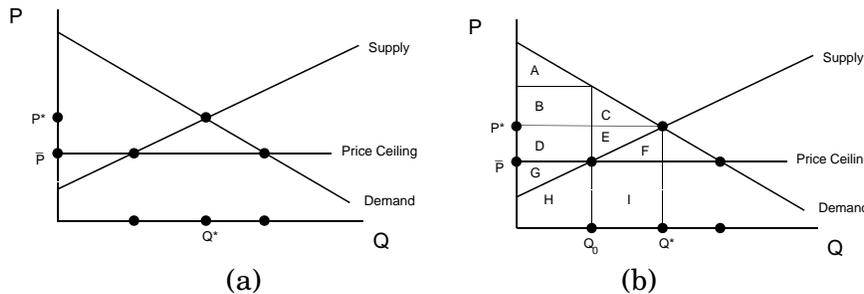


Notice that areas F, H, and I end up not being part of anybody's surplus. I didn't have to label them, but I did it anyway to illustrate how it doesn't do any harm to label areas you don't end up using. Also, I could have labelled area $A+B+C$ just as one area, going straight to consumer surplus, but we'll see that not all that area will be surplus under the regulation, so chopping it into pieces will make comparison easier.

5. Figure out for the regulated market which areas are in producer surplus, consumer surplus, effects on third parties (for example, tax revenue), and total surplus.

With rent control, the quantity is Q_1 , so the surpluses will only cover areas from quantity 0 to Q_1 . Consumers' total benefit is the area under the demand curve from quantities 0 to Q_1 , but they have to pay \bar{P} , so their surplus is the area A+B+D between the demand curve and the \bar{P} flat line. Producers' total benefit (the landlords') is their revenue, $\bar{P} \cdot Q_1$, but their surplus is the difference between that revenue and their costs, the area G between the supply curve and the \bar{P} flat line. There are no effects on third parties. The total surplus adds up to A+B+D+G.

FIGURE 1.6 (REPEATED FROM ABOVE)
RENT CONTROL



$$\begin{aligned} \text{CS (rent control)} &= A+B+D \\ \text{PS (rent control)} &= G \\ \text{TS (rent control)} &= A+B+D+G \end{aligned}$$

6. See which has the bigger total surplus, the free market or the regulated market.

The free market has the biggest total surplus, because it includes all the areas in the regulated surplus plus C+E. This difference C+E is called the **deadweight loss** or **triangle loss** or **allocative inefficiency**. Consumer surplus has risen by amount D-C because of rent control, but producer surplus has fallen by D+E. The losers lose more than the winners gain, so rent control does not maximize surplus.

1.6 Quality, Rationing, and Rent-Seeking

Our surplus calculations have been based on hidden assumptions that simplify the analysis but leave out important effects of what happens when regulation leaves a market with excess supply or excess demand.

Our first assumption was that product quality was the same before and after price controls. In fact, if sellers cannot lower their prices below a government ceiling, so the market has excess supply, they will start raising quality to compete for the scarce customers. This was the case when airline fares were regulated in the 1960's; airlines

competed on the quality of their service and their meals, which dropped once the market was deregulated and they could compete on price instead. If there is a price floor, such as with rent control, there is excess demand, and so sellers can spend less on quality and still find buyers.

Before rent control, landlords had to maintain the quality of their apartments to attract tenants willing to pay the equilibrium rent of \$500 in our example. Rent control reduces the rent to \$400, so even if the landlord stops maintaining the apartment he will still find tenants willing to rent it. Landlords will paint the walls less often, make repairs more slowly, and in general try to pass along all the costs they can to the tenants. Regulators know about this effect, so rent control is accompanied by regulations to try to force landlords to provide maintenance, but since it is hard to specify and enforce everything a landlord does to make his apartments attractive, apartment quality in rent controlled areas falls over time. If quality falls enough, the rent of \$400 actually becomes the equilibrium price and there is no excess demand, but there is still a surplus loss because of the inefficiently low level of quality.

BOX 1.1

RENT CONTROL IN NEW YORK

New York instituted rent control after World War II because of “the post-war rental housing emergency”. By 2010, only 40,000 apartments were still covered— just ones built before 1947 and still occupied by the same family as in 1971. Whoever does live in one is paying a below-market rent and won’t want to move, even if someone else would get more value from it. Something over one million apartments are “rent stabilized”, a form of rent control in which rents can only rise at a certain rate each year.

Changes in quality actually reduce the deadweight loss from price controls, but we neglected a second feature of markets that increases it: rent seeking. **Rent seeking** refers to people engaging in activities that transfer surplus from other people to themselves. The transfer itself does not change total surplus; it is just redistributed. Rent-seeking activities often are costly, however, and those costs can cause real surplus loss in the same way that production costs do. Production costs create surplus, but rent-seeking costs just take it away from somebody else.

The phrase “rent seeking” is standard, but it was chosen badly by Anne Krueger in 1974 when she picked a phrase to describe Gordon Tullock’s 1967 idea of costs that just transfer surplus.¹⁵ In the early 1800’s David Ricardo used the word “rents” to refer to producer surplus of any kind, taking the word from the payment to land. There is nothing inefficient about seeking producer surplus— that’s what motivates sellers. But “rent seek-

plus of any kind, taking the word from the payment to land. There is nothing inefficient about seeking producer surplus— that’s what motivates sellers. But “rent seek-

¹⁵See Gordon Tullock (1967) “The Welfare Cost of Tariffs, Monopolies, and Theft,” *Western Economic Review* and Anne O. Krueger (1974) “The Political Economy of the Rent-Seeking Society,” *American Economic Review*.

ing” is now used by economists to refer to costs that take surplus rather than making it.

In the context of rent control, the extra search time spent trying to find a rent-controlled apartment is a form of rent-seeking cost. Suppose in a city that is phasing out rent control Mr. Smith has a choice between immediately taking an uncontrolled new apartment at a rent of \$500 or waiting and searching further for a rent-controlled old apartment at a rent of \$400. Smith would be willing to spend hundreds of dollars of his time— up to the equivalent of \$100/month— to find one of the old apartments that was available.

Of course, “rent seeking” applies to any kind of product, not just apartments. One example is fresh fruit in the Soviet Union in the 1970’s. The socialist government set the price of fruit low, so there was excess demand. The result was long lines of consumers showing up early to stores wanting to buy fruit when it was available. The time spent standing in line was a rent-seeking cost, that did not increase the total amount of fruit.

A third assumption that was behind the surplus analysis in the previous sections was that the people who were able to rent the scarce apartments were the ones who valued them the most. In a free market with a market price of \$500, we know that anybody who values the good at more than \$500— that is, the people who make up the top part of the demand curve— will be able to buy it. At the rent-controlled price of \$400, however, it might be that someone willing to pay \$600 cannot find an apartment but someone willing to pay just \$450 can find one. Whether this happens or not depends on the form of rationing.

Rationing means the allocation of goods by some means other than price when there is excess demand or supply (if it’s excess supply, the rationing determine which sellers are the ones lucky enough to be able to sell). Under socialism and during wartime in capitalist economies, there is often excess demand because the government imposes price ceilings. Sometimes the government distributes ration tickets to consumers that they have to present to a seller whenever they buy something. That is one form of rationing, but the term refers to whatever mechanism determines who gets to buy and sell. Ration tickets, political influence, willingness to stand in line, and willingness to wait for delivery are all rationing mechanisms.

Inefficient rationing occurs when not only is the quantity too low because of some flaw in the market or in policy, but the buyers who are able to find a seller are not the buyers who value the product the most. If the rent is set at \$400, below the market price of \$500, then consumer Brown, with valuation \$425, and consumer Smith, with valuation \$800, are both willing to pay that \$400 rent. The landlord doesn’t care, because he gets \$400 either way. Thus, he might well rent the apartment to Brown instead of Smith. Ration tickets also result in inefficient rationing, because the people who value the good the most usually aren’t given enough ration tickets to be able to

buy it.

Our analysis above assumed that whichever consumer valued the apartment the most would get it, so the consumers left apartmentless are the consumers with the lowest values. That form of priority is **efficient rationing**, because it gives the apartments to the people who value the good the most. Any kind of excess demand or supply and rationing creates inefficiency— to have excess demand because at the regulated price more units are demanded than are supplied— but at least under efficient rationing the deadweight loss is smaller than it might be. Oddly enough, the rent seeking discussed earlier can actually make rationing efficient. Consumers with higher valuations are willing to spend more on rent seeking, and this will tend to make them obtain more of the scarce good than low-valuation consumers.

FIGURE 1.8
THREE WAYS RATIONING MIGHT WORK OUT

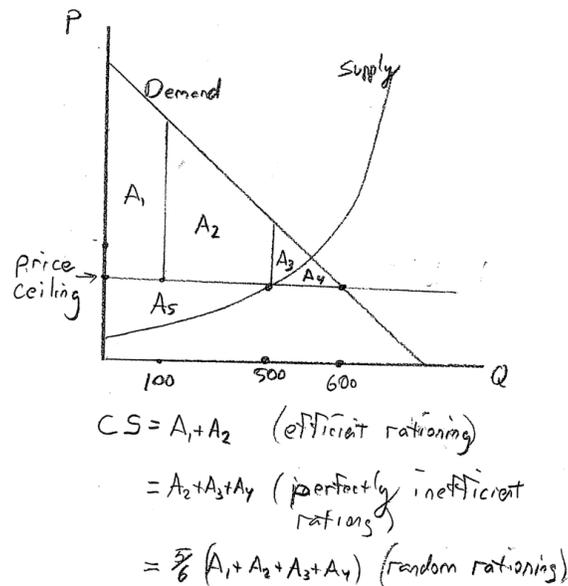


Figure 1.8 shows how three different kinds of rationing would affect surplus. At the price ceiling, the quantity supplied is 500, but the quantity demanded is 600. The producer surplus is going to be area A_5 regardless of how consumers are rationed, because it is precisely the suppliers lowest on the supply curve who are willing to supply. With efficient rationing, the top 500 consumers on the demand curve are the ones who get to rent apartments, so consumer surplus is $A_1 + A_2$. We can imagine **perfectly inefficient rationing**, which would give the apartments to the 500 consumers who *least* value them but are still willing to pay the price ceiling for them. Those are the consumers between $Q = 100$ and $Q = 600$ on the demand curve; the highest-valuing consumers are rationed out and don't get apartments. The consumer surplus from per-

fectly inefficient rationing would thus be $A_2 + A_3 + A_4$. The third type of rationing in Figure 1.8 is **random rationing**, in which each type of consumer has an equal chance of getting an apartment. Since there are 500 apartments and 600 consumers who want them, that means each type of consumer has a $5/6$ chance. It's easier to think of the equation for the resulting size of consumer surplus than to draw areas. That equation will be that consumer surplus is $5/6(A_1 + A_2 + A_3 + A_4)$ —that is, it's $5/6$ of what the consumer surplus would be if *all* the 600 consumers could find apartments. Each type of consumer will have a $5/6$ chance of getting his valuation height minus the price ceiling and a $1/6$ chance of getting zero consumer surplus. Random rationing is quite realistic. It generate more consumer surplus than perfectly inefficient rationing but more than efficient rationing.

1.7 Conclusion: Markets Work

Before we start talking about regulations in particular industries, we need to establish a theoretical framework for talking about regulation at all. Basic microeconomics is the framework we'll use, with particular focus on supply and demand and surplus analysis.

This chapter has talked about how free markets work and what it means for a market to work well. The policy objective economists focus on is efficiency: the maximizing of total surplus. Total surplus, in turn, is the sum of the benefits the various players in the economy receive from the market for a particular good. Maximizing surplus is a reasonable goal, but a limited one. It ignores considerations of morality, fairness, and equality. That sounds like a serious drawback, but it is actually an advantage. It allows us to focus on the dollar effects of policies first, and doesn't prevent the application of other criteria afterwards.

Before looking at regulation, one must understand what happens in unregulated markets, so we reviewed the theory of supply and demand. The free market equilibrium price leads to the efficient quantity being traded, so that surplus is maximized by the free market. A price ceiling such as rent control or a government-imposed cost such as a tax reduces the quantity traded to below the efficient level and thus reduces total surplus.

If willingness to buy and prices do not decide who produces which things and who consumes which things, then something else decides it. When rent control is imposed, for example, the people who get to rent apartments are those who are best at searching or those whom the landlord thinks will be the best tenants. If the landlord is indifferent otherwise, he might break his tie by something arbitrary such as only renting to handsome people, since he cannot decide based on who is willing to pay the most. If the government were to set the price of cars based on what it considered fair, there would be shortages or surpluses and someone— in government or in the regulated

business— would use his personal likes and dislikes to decide who gets the scarce cars or the scarce customers. Such regulation would replace competition based on wealth and tastes with competition based on political power.

This chapter has taught that markets work well in maximizing surplus. That is a good starting point for thinking about regulation, but not a good stopping point. We will see in the next chapter that there are situations of “market failure” in which the standard model of supply and demand does not tell the whole story and where government intervention can actually increase total surplus. We will also see that besides the Invisible Hand of the market helping to maximize surplus, there is also an invisible effect of government that has been in the background even in this chapter. Governments enforce property rights, and without property rights, our supply and demand curves wouldn’t apply—there would be stealing instead of buying and selling. So don’t stop here— read on to learn when markets go wrong and what’s good about government.

REVIEW QUESTIONS

1. How do supply and demand reach the equilibrium price and quantity?
2. What does surplus maximization mean, and what are its limitations as an objective?
3. How is it that the market maximizes surplus, and why is it more difficult for a planner to do the same thing?
4. What is the effect of government-imposed prices or quantities on social surplus?
5. How does the type of rationing when there is excess supply or demand affect surplus?

READINGS

1. “Adam Smith and the Invisible Hand,” Helen Joyce, *PlusMath.org*.
2. *The Merchant of Venice*, William Shakespeare, Antonio in Act III, Scene iii. Shylock in Act IV, Scene i. See <https://www.youtube.com/watch?v=gSUuczZcUaU>.
3. “Free Market Food Banks,” Alex Tabarrok, *Marginal Revolution* weblog.
4. “I, Pencil,” Leonard Reed, *The Freeman* (1958).
5. “A.G. Schneiderman Warns against Price Gouging during Winter Storm Stella in New York State,” press release, <https://ag.ny.gov/press-release/ag-schneiderman-warns-against-price-gouging-during-winter-storm-stella-new-york-state>.