

Econ 354

Greg Dow

March 10, 2021

Notes on Robert Ellickson, "Order Without Law"

Introduction (first set of lecture notes)

Robert Ellickson is an emeritus professor at the Yale Law School. I will start off with a few general comments about the book and how it fits into the rest of the course.

The book has two main parts. The first six chapters are mostly a case study of Shasta County, which is located in northern California. You can think of this as a bit like the anthropological case studies from the Johnson and Earle book, except that it involves a tribe of Americans.

The rest of the book is a general discussion of social norms: what are they, where do they come from, what determines their content, and so on. RE also includes some shorter case studies later in the book.

His main point is simple: informal social norms are often more important than formal law as a way of ensuring cooperation and resolving conflicts.

This should not be a surprise at this point in the course.

- (a) Johnson and Earle talked a lot about non-state societies. In a sense these societies have no law. The only thing holding them together is social norms.
- (b) Ostrom was concerned with societies where states exist, so law enforcement was at least possible. But social norms, reciprocity, and informal punishments were often more important than the legal system.

The value of Ellickson's book is that he provides a lot of detail about the kinds of social norms people actually use, and he has some things to say about how economic and game theory models can be linked to a theory of social norms. He also approaches issues from a legal point of view, which is different from what we have seen so far.

From an economic standpoint, this book gives us an opportunity to talk about bargaining and to apply what we have already learned about repeated games. We will also consider the concept of externalities.

As with the Ostrom book, I will begin by developing a theoretical background that should help you understand what Ellickson is doing. After that, I will start talking about the case study of Shasta County (in my next set of notes).

The Coase Theorem. This idea motivates a lot of the research RE describes in the book. Unfortunately although he mentions it from time to time, he does not explain it in much detail, so I will need to do that.

The Coase Theorem comes from a famous article by Ronald Coase called "The Problem of Social Cost", published in 1960. The theorem is central to the study of externalities, an idea I will discuss later.

By the way, if you like economics but are not good at math, the example of Coase should give you some hope. Coase received a Nobel Prize in economics essentially for writing two journal articles. One was published in 1937 and the other (mentioned above) is from 1960. Neither article contains any math. However, both were hugely influential and are still cited by many economists today.

Here is what the Coase Theorem says (see p. 2 of Ellickson):

"When *transaction costs* are zero a change in the *rule of liability* will have no effect on the *allocation of resources*."

I put the three key concepts in italics. In order to understand the theorem, you need to know what each of these concepts means.

I'll start with an example that resembles the situation in Shasta County described by RE. We have a farmer (person 1) and a rancher (person 2). They live next to each other. The rancher's cattle could get loose and stampede through the farmer's garden, which would cause damage (amount in dollars = d) for the farmer. This damage can be prevented by building a fence to keep the cattle from escaping (cost of a fence in dollars = c).

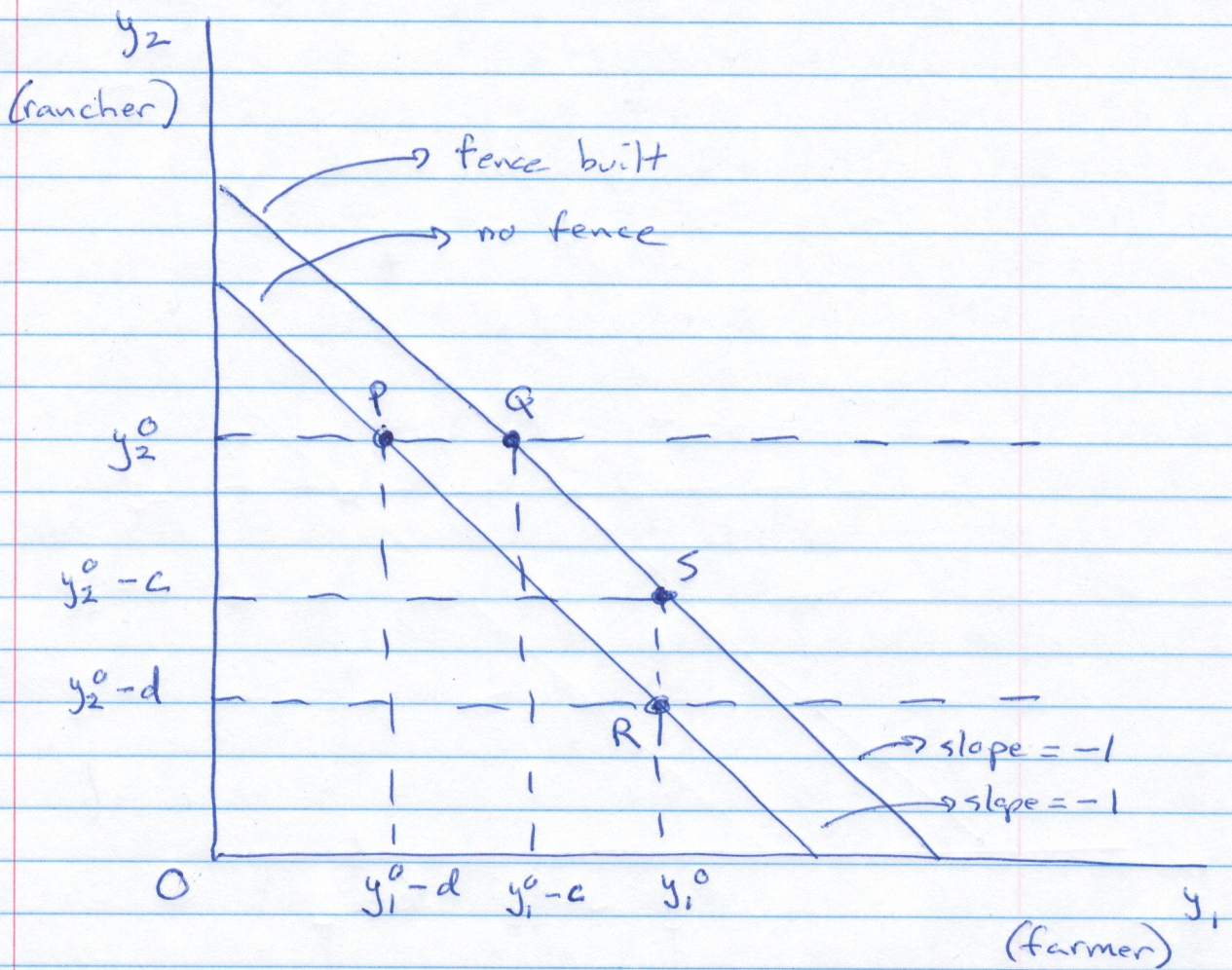
We will assume $c < d$ so it is worthwhile to build a fence. But we don't yet know who is responsible for building it. Each person would like the other person to pay the cost c .

There are two possible legal rules:

- (A) The rancher is not liable. This means that if the damage occurs, the farmer must suffer the loss d without any compensation.
- (B) The rancher is liable. This means that the rancher must pay d in compensation to the farmer if the damage occurs.

Now look at Figure 1. Payoffs to the farmer are on the horizontal axis and payoffs to the rancher are on the vertical axis. I will assume that if there is no fence and no damage, the farmer has income y_1^0 and the rancher has income y_2^0 . I also assume that either person is capable of building a fence (no special skills are required).

Let's consider what happens under each of the two legal rules described above.



Payoffs to Farmer and Rancher

Figure 1

- (A) The rancher is not liable. If the fence is not built, the farmer suffers the damage d and the rancher pays nothing. The resulting payoffs are $(y_1^0 - d, y_2^0)$ as shown by point P in Figure 1. On the other hand, if the farmer builds a fence, the farmer has to pay the cost c but no longer suffers any damage. The resulting payoffs are now $(y_1^0 - c, y_2^0)$ as shown by point Q. Because we are assuming $c < d$, the farmer is better off building the fence and point Q is to the right of point P. They are on the same horizontal line because the rancher's payoff is the same in each case. Under the legal rule (A), the farmer's personal incentive to minimize cost ensures that the fence will be built, and the outcome will be point Q.
- (B) The rancher is liable. If there is no fence, the farmer suffers the damage d , but the rancher must compensate the farmer for the loss. This means that it is the rancher, not the farmer, who bears the cost d . The resulting payoffs are now $(y_1^0, y_2^0 - d)$, which puts us at point R. However, the rancher has the option of building a fence. This prevents the damage, so the rancher only has to pay the cost of the fence c . The resulting payoffs are $(y_1^0, y_2^0 - c)$, which is indicated by the point S. Notice that because $c < d$, point S is above point R and the rancher is better off at S. The two points are along the same vertical line because the farmer's payoff is the same in each case. We conclude that under the legal rule (B), the rancher's incentive to minimize cost ensures that the fence is built, so the outcome is S.

At this point we can start to interpret the Coase Theorem. The "rule of liability" is either (A) or (B). Now define "the allocation of resources" to mean whether the fence is built or not. Ignore transaction costs for the moment. What we have shown is that regardless of whether the rule of liability is (A) or (B), someone will build the fence. Hence, the rule of liability does not affect the allocation of resources.

We can go further with Figure 1. At point P, the total payoff of the two people is $y_1^0 + y_2^0 - d$. This is also the total payoff of the two people at point R. So when no fence is built, the total payoff is the same. The rule of liability only determines who incurs the cost of the damage d .

Now consider the set of all points where the sum of the payoffs y_1 and y_2 is equal to the same total $y_1^0 + y_2^0 - d$. This is true when $y_1 + y_2 = y_1^0 + y_2^0 - d$. Rearranging this gives $y_2 = y_1^0 + y_2^0 - d - y_1$. This is the equation for a straight line in Figure 1, which has the slope -1 . The reason is that whenever we give one more payoff unit to person 1, we must give one less payoff unit to person 2. Notice that points P and R are both located on this line because they give the required total payoff.

We can do the same thing with points Q and S. In each case, the total payoff is $y_1^0 + y_2^0 - c$. The only difference between Q and S involves the question of who pays for the fence. Construct the line in Figure 1 with $y_1 + y_2 = y_1^0 + y_2^0 - c$. Again this leads to an equation for a straight line with the slope -1 , but this time going through points Q and S. The line through Q and S yields a higher total payoff than the line through P and R because $c < d$. The two lines are parallel because they both have the slope -1 .

Now we can provide a further interpretation of the Coase Theorem: because total income is greater if the fence is built than if it is not, the farmer and the rancher will at least agree that someone should build the fence. This means that they will agree on the allocation of resources, in Coase's definition of that concept. To put it another way, they agree that it is better to be along the line through points Q and S than the line through points P and R.

The only disagreement is about who should pay for the fence (should they be at point Q or point S?). Clearly the farmer prefers point S (it is better to have the rancher pay), but the rancher prefers point Q (it is better to have the farmer pay). This issue is settled by the rule of liability, which in effect determines how income will be distributed between the farmer and the rancher. As we have seen above, under rule (A) the farmer will pay for the fence and under rule (B) the rancher will pay for the fence.

We can generalize this idea. Think about some level of total income Y , and think about the ways of distributing this income that add up to Y . The result is a line defined by the equation $y_1 + y_2 = Y$. It makes economic sense for the two people to do whatever would be necessary to maximize Y . After that, they can argue about how to divide up Y (what the individual payoffs (y_1, y_2) will be). In general, it is better to argue over a larger pie than a smaller pie.

Next let's do a slightly more complex example. Suppose that only the farmer knows how to build a fence. The rancher lacks such skills and cannot build a fence. You might think that now the liability rule would have some effect on whether or not a fence is built, but it turns out that this is not true.

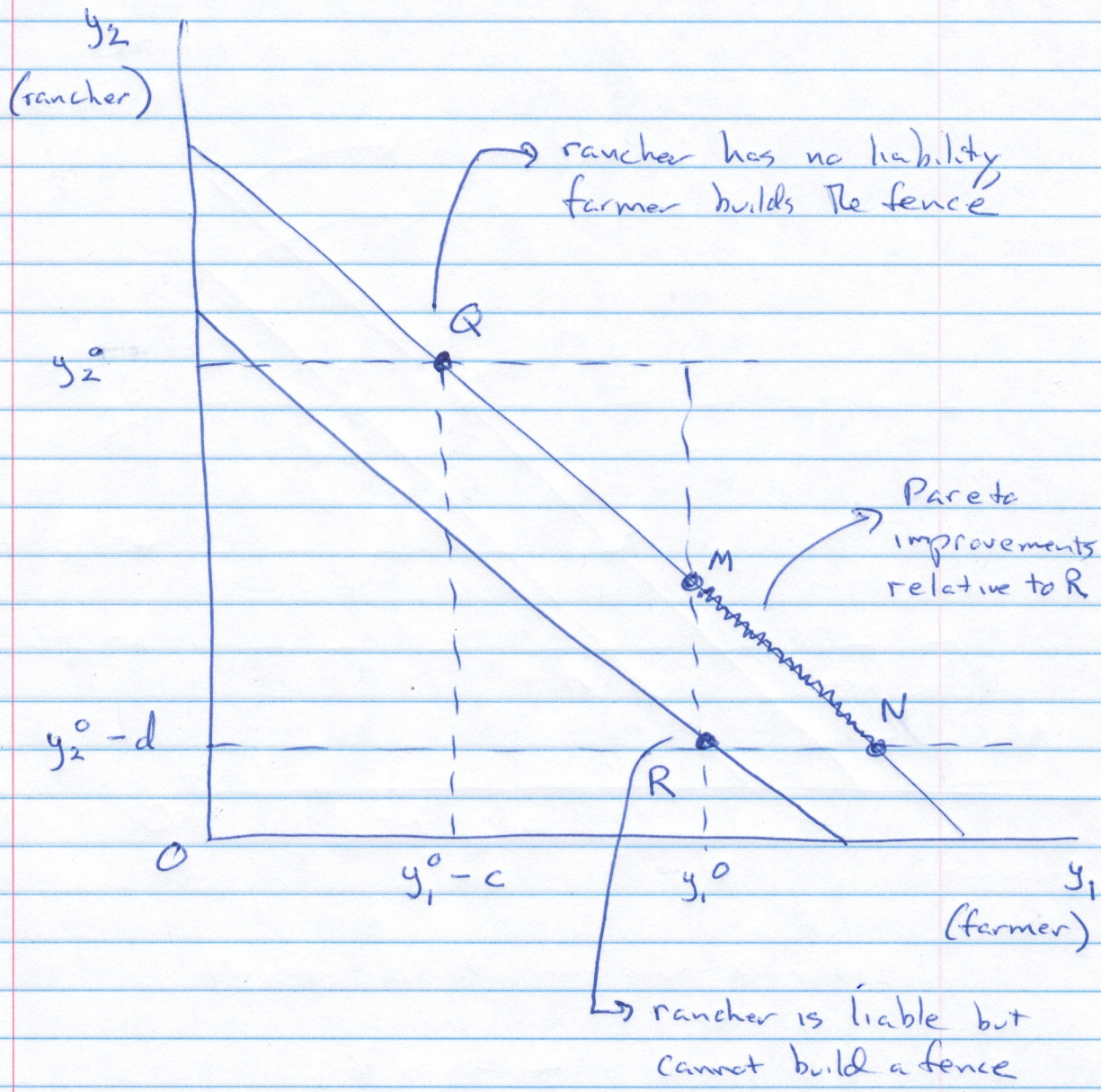
Under rule (A) where the rancher is not liable, we get the same outcome as before. The farmer wants to build the fence because c is less than d . Because the farmer knows how to build a fence, we go to point Q in Figure 1.

Under rule (B) where the rancher is liable, things become more interesting. The rancher will have to pay for the damage d if the fence is not built. Thus, the rancher would like to have a fence but can't build one. As a result the rancher will be stuck at point R in Figure 1 and can't get to point S.

What is the solution? In the real world, an obvious possibility is for the rancher to have a conversation with the farmer and ask the farmer to build a fence. The farmer would have to pay c in order to do this, and will be reluctant to agree. However, the rancher can offer the farmer something in return.

Let's say the rancher offers to give the farmer a payment t as compensation for building a fence. I will sometimes call this a transfer. Clearly we must have $t \geq c$ or the farmer will not agree to build the fence. Also we must have $d \geq t$ or the deal is not worthwhile to the rancher (it would be cheaper just to pay compensation to the farmer for the damage d).

Now look at Figure 2, which has the same axes as Figure 1. Assuming that the rancher is liable, we start from point R. We know that if the fence is built, total income will be the



Payoffs When Rancher Cannot Build Fence

Figure 2

same as at point Q: $y_1 + y_2 = y_1^0 + y_2^0 - c$. The line for this total income passes above the point R, where the total income is lower ($y_1 + y_2 = y_1^0 + y_2^0 - d$).

If the farmer and the rancher can reach a point along the segment MN, they can achieve a Pareto improvement relative to point R. At point M or N, one person is better off than at R and the other person is no worse off. Anywhere between M and N, both people have a higher payoff than at point R.

When the farmer and the rancher discuss the payment t the rancher will give to the farmer in exchange for building a fence, they are discussing payoffs of the form

$$y_1 = y_1^0 - c + t \quad \text{and} \quad y_2 = y_2^0 - t$$

Together these always add up to the total payoff for the line through point Q, which is the same as the total payoff at all points along the segment MN (the transfer t cancels out and does not affect the sum $y_1 + y_2$).

To get a Pareto improvement starting from point R, we need

$$\begin{aligned} y_1 = y_1^0 - c + t &\geq y_1^0 && \text{so the farmer is no worse off than at R, and} \\ y_2 = y_2^0 - t &\geq y_2^0 - d && \text{so the rancher is no worse off than at R.} \end{aligned}$$

The first inequality reduces to $t \geq c$ and the second reduces to $d \geq t$, which is consistent with what we said earlier. For the deal to make sense, the transfer cannot be less than c (otherwise the farmer will refuse) or more than d (otherwise the rancher will refuse).

You should be able to convince yourself that when the transfer t is at the lower bound c , we go to point M in Figure 2. In this case the farmer is indifferent toward building the fence and all of the gains from having a fence go to the rancher.

When the transfer t is at the upper bound d , we go to point N in Figure 2. In this case the rancher is indifferent and all the gains from having a fence go to the farmer.

Any value of the transfer t in the intermediate range where $c < t < d$ corresponds to some point along the line segment between M and N. Such points make both people better off than they were at R. However, there is still a disagreement over income distribution: the farmer wants a big t (closer to N) and the rancher wants a small t (closer to M).

The outcome depends on the bargaining skills of the two people. If the farmer is a good bargainer while the rancher is not, we will get a value of t near d and we will end up close to point N on the graph. If the opposite is true, we will get a value of t near c and end up close to point M on the graph.

Now go back to the Coase Theorem. If there is no obstacle to bargaining between the two people, they should be able to agree that the farmer will build a fence. They should also be able to agree on a specific transfer t the rancher will pay to the farmer, which is equivalent to choosing a specific point along the line segment MN.

Again we get a result consistent with the Coase Theorem: no matter what the liability rule says, the fence will be built. Under rule (A) where the rancher is not liable, the farmer builds it because this minimizes cost for the farmer. Under rule (B) where the rancher is liable, the farmer builds it because bargaining with the rancher leads to a compensation payment t that exceeds the cost of building a fence.

Another way to think about this is that when bargaining is feasible, the two parties will always agree to maximize total income, and this always involves building a fence. Any conflict over the distribution of this total income is resolved separately by bargaining.

Or to put it yet another way: if there is no fence then a Pareto improvement is feasible, and the bargaining mechanism provides a way to achieve a Pareto improvement. But notice that unless we know something about the bargaining skills of the two people, we cannot say for sure which Pareto improvement will occur (anything between M and N would count as a Pareto improvement).

Here are a few observations about this reasoning.

- (a) It should be clear that the same argument would apply if the rancher can build a fence but the farmer cannot. Suppose this is true, and we have rule (A) where the rancher is not liable. In that case the farmer would need to offer compensation to the rancher. However, bargaining would still get us to the line that passes through points Q and S in Figure 1, and the fence would still be built. You should draw a graph in order to make sure you understand how this works.
- (b) Although the rule of liability has no effect on whether the fence is built, it does affect the distribution of income. In the example I gave above, where the farmer knows how to build the fence but the rancher does not, the rancher clearly prefers not to be liable. This puts the rancher at point Q, which gives the rancher a higher payoff than any point along MN. The farmer prefers for the rancher to be liable, because from the farmer's point of view, anything along MN is better than Q.
- (c) For this reason, we have to be careful about the term "allocation of resources" in the Coase Theorem. If we define this to mean "whether the fence gets built", then Coase is correct. But most economists would define the allocation of resources in a way that includes the distribution of income. With that definition, Coase is not correct, because the rule of liability clearly does affect income distribution.
- (d) Because there is conflict over income distribution, this translates into conflict over the rule of liability. For example, if the state legislature in California is deciding which rule of liability should be used, the farmer will lobby for the rancher to be

liable, and the rancher will oppose this. Together they may waste a lot of time and money fighting over which liability rule should be adopted.

- (e) What if $d < c$? In this case the cost of a fence exceeds the damage done by the cattle. As a result, total income is higher when there is no fence. Accordingly, no fence will be built (it is impossible to get a Pareto improvement by building one). This will be true regardless of the rule of liability.
- (f) What if there are other ways to avoid the damage d besides building a fence? For example, the rancher could watch the cattle to make sure they don't cross into the farmer's property, or hire employees to keep the cattle where they belong, or have fewer cattle so there is less danger that some animals will wander away. In all of these situations, there is some cost to the rancher of preventing the damage, and it is like the case where the rancher can build a fence. If the rancher is liable and the costs of preventive measures are less than d , the rancher will take the preventive measures in order to avoid paying d . If the rancher is not liable, then the farmer can offer a payment to the rancher and use bargaining to avoid the damage.
- (g) You can think about the liability rule as assigning property rights. In one case the rancher has a right to allow the cattle to wander into the farmer's land, while in the other case the farmer has a right to insist that the cattle be kept out. In a sense this determines whether the farmer has weak or strong property rights over her land.
- (h) Another way to think about the liability rule is that it establishes the starting point (or status quo point) from which bargaining begins. If the rancher has no liability then the starting point for a bargaining process will be point P in Figure 1. If the rancher does have liability then the starting point for bargaining will be point R.

Now we know about "the rule of liability" and "the allocation of resources". The third key concept in the Coase Theorem is "transaction cost". Let's spend some time on that.

A *transaction cost* is any cost associated with making a transaction or exchange. Such costs fall into three general categories: search costs, bargaining costs, and enforcement costs.

A search cost could include the cost of finding out who is selling a good or service, who is willing to buy it, what prices people are asking, what the quality of the good or service is, and so on. These costs are important in many economic models and in the real world, but we will not be concerned with them here.

A bargaining cost could include the time it takes to negotiate an agreement, to find out what the other person is willing to pay or accept, or the costs of hiring lawyers to write contracts. I'll say more about these costs below.

An enforcement cost involves monitoring whether the other person is doing the things they agreed to do, and imposing penalties on them if they violate the agreement. For

example, someone might not deliver a good or service of the same quality as what was promised, or might fail to make a payment they promised to make. It may be costly to find out what the true quality was (maybe you only discover this later when you use the product), or to watch the activities of another person. It is also costly to hire lawyers in order to go to court and sue someone for violating a contract.

To see why transaction costs are important in the Coase Theorem, go back to the case where the rancher is liable but does not know how to build a fence (Figure 2). Earlier I said that this problem could be resolved through bargaining, which gets us from point R (the status quo) to some Pareto improvement along the segment MN.

This ignored any potential bargaining costs. In effect, I assumed that any time a Pareto improvement was feasible, it could be obtained costlessly. But if there are large costs of time or money in the bargaining process, it might not be possible for the rancher and the farmer to agree on a transfer t . If they cannot do this, they may get stuck at point R with no fence (without compensation, the farmer won't build one).

In this situation the Coase Theorem is still true (if transaction costs were zero, the rule of liability would have no effect). However, transaction costs are positive rather than zero, so the rule of liability can affect the allocation of resources. In particular rule (A) where the rancher has no liability leads to a fence (the farmer builds one to avoid the damage), but rule (B) where the rancher is liable leads to no fence (there is no bargaining process through which the rancher can persuade the farmer to build one). This implies that total income is higher under rule (A) than under rule (B).

Next I want to define externalities and explain their relationship to the Coase Theorem.

An *externality* (either positive or negative) is a direct effect of one person's actions on another person's welfare, without any financial compensation for this effect.

For example: suppose each morning I walk by my neighbor's garden and I enjoy looking at her flowers. This is a *positive externality*. My neighbor's actions (tending a garden) make me better off, and I am not paying her anything in exchange for this service.

Suppose my other neighbor likes to play his electric bass at 4 AM and this wakes me up. This is a *negative externality*. My neighbor's musical tendencies make me worse off, and he is not paying me anything in exchange for interrupting my sleep.

The damage done by the cattle in our earlier example was a negative externality. The rancher's actions (raising cattle) decrease the welfare of the farmer (because the animals can escape and destroy the garden). Under rule (A) where the rancher is not liable, the farmer receives no compensation. Under rule (B), the rancher does pay compensation and therefore has incentives to avoid the damage.

The Coase Theorem has caused a lot of controversy because some people use it to argue that there is no need for government regulation to deal with negative externalities. For

example, suppose a factory emits pollution into the air or water and the pollution reduces the welfare of people who live close to the factory. One might argue that according to the Coase Theorem, we shouldn't care about the rule of liability, because if it makes sense to cut back on pollution (reducing the pollution costs less than the damage it does), then the pollution will be eliminated regardless of the liability rule. Therefore it doesn't make any difference whether the factory owner is liable for the pollution damage or not.

There are several problems with this argument. First, even if transaction costs are zero, you might have moral reasons to care about rules of liability. Keep in mind that the rule of liability in effect determines property rights. If you say that the factory owner has no liability, you are saying that the factory has a right to pollute, and if the neighbors don't like it, that's too bad. If you say that the factory owner is liable, then you are saying that the neighbors have a right to clear air. Many people would make the latter argument.

Second, you might care about income distribution. Even if transaction costs are zero, we know that the rule of liability can affect the way in which income is distributed. It may be true that there is no effect on the level of pollution, but you might still be concerned about how income is distributed between the factory owner and the neighbors.

Third, in many cases transaction costs are not zero (not even close). For example, there may be hundreds or thousands (or millions) of people who have to breathe the air that is polluted by the factory, or have to drink the water contaminated by the factory. In cases like this, bargaining and enforcement costs are likely to be extremely high, and it would be unreasonable to think that people negatively affected by the pollution will somehow band together to negotiate with the factory. For this reason, we normally rely on some type of government regulation to solve large-scale environmental problems.

Fourth, everything I have just said assumes that people have good information about the damage being done by the pollution and the costs of cleaning it up. In the real world, it is frequently true that people don't have such information (for example, they may not know the health risks caused by the pollution). Again, we normally expect governments to deal with such problems because it is too difficult for every individual to gather the relevant information (they may not have the time or the skills to do it). Anyway, it doesn't make much sense to have individuals independently gathering the same information multiple times, rather than having the government do it once on behalf of everyone (remember that information is frequently a public good).

Fifth, there are enforcement costs. Suppose the neighbors do reach some agreement with the factory owner about cutting back on pollution. Who is going to monitor the factory to make sure they respect the agreement? Who is going to impose penalties on the factory owner if the agreement is violated? In principle the neighbors might hire some experts to monitor the factory, and hire lawyers to impose penalties through a class action lawsuit, but it is usually a lot cheaper and easier for government regulators to do these things.

Coase was well aware of the issues arising when transaction costs are positive. The first half of his 1960 article explores what happens when there are no transaction costs, while

the second half explores what happens when transaction costs are positive. In a number of ways, the second half is more important and interesting. But people sometimes focus on the first part even in cases where transaction costs are clearly not zero. This can lead to misleading interpretations of what Coase was saying.

You might consider how the Coase Theorem relates to common pool resources. The big problem in the Ostrom book is that individual people could take too much of the resource because they ignore the costs they are imposing on other people. Another problem is that individuals could free ride on investments in maintaining a resource because they ignore the benefits their own investments provide for other people.

The first problem can be interpreted as people ignoring a negative externality, and thus doing too much of something from a social point of view. The second problem can be interpreted as people ignoring a positive externality, and therefore doing too little from a social point of view.

If transaction costs were zero, we could try to use the Coase Theorem to argue that there is no real problem with CPRs: people can maximize total income or total welfare through some kind of bargaining process. But of course, if you read the Ostrom book it becomes obvious that transaction costs are not zero, so people have to create complex institutions in order to manage CPRs.

Another point I want to make is that bargaining costs could be high even if there are only two people involved (we don't need hundreds of people for this to be an issue). Suppose I am thinking about selling my condo to you. There is some minimum price that I would be willing to accept, but you don't know what it is. There is also a maximum price that you are willing to pay, but I don't know what it is.

The result is generally some negotiation process where I ask for a price well above the minimum I am willing to accept, and you respond with an offer that is well below the maximum you are willing to pay. Each of us has incentives to delay in order to get the other person to make their offer more generous. The problem of incomplete information could make it impossible for us to come to an agreement, even in cases where we would both potentially benefit (you are genuinely willing to pay more than the minimum I am genuinely willing to accept). Therefore we may not achieve a Pareto improvement, even if one is available in principle.

Similar things happen in many other bargaining situations. Another example involves collective bargaining between an employer and a union, where the employer knows the true profitability of the company and the union knows what wage its members are willing to accept. Neither side knows precisely what the other is willing to pay or accept. This can lead to strikes or lockouts that may go on for months or even years, where both sides are losing money in relation to what they could have had from an agreement, but each is trying to signal that they are tough bargainers and won't back down.

The point is that you should not assume bargaining is easy or cheap just because only a few people are involved. Bargaining costs (and therefore transaction costs) are almost always high when there are many people. Therefore having a small number of people is typically a necessary condition to have low transaction costs. But this is not a sufficient condition, because transaction costs also depend on how information is distributed.

The main lesson I want you to take away from this discussion is that the Coase Theorem is important, at least theoretically and sometimes also in the real world. But you should not get too excited about it because the assumption of zero transaction costs often fails to hold in real situations. In fact institutions are often created precisely because transaction costs are substantial. Whether transaction costs are significant in a particular situation is an empirical question, and it can't be determined just through theory or speculation.

For most large externality problems there are many people involved, individuals may not have reliable information, transaction costs are high, and property rights do matter. Such problems are typically handled through government regulation and this does influence the allocation of resources. In fact, changing the allocation of resources (such as the amount of pollution emitted) is usually the point.

Econ 354

Greg Dow

March 14, 2021

Notes on Robert Ellickson, "Order Without Law"

Chapters 1-6 (second set of lecture notes)

These notes are about the case study of Shasta County in chapters 1-6. Along the way, I will discuss how this case relates to my first set of lecture notes on the Coase Theorem.

Shasta County is divided into areas called open range and closed range.

- (A) In *open range*, generally the ranchers are not legally liable for the damage done by their cattle, assuming the land owned by the victim is not fenced. Therefore, the cost of any damage is borne by the victim. In this case, farmers and others must decide whether it is worthwhile to try to keep the cattle out.
- (B) In *closed range*, generally the ranchers are liable for the damage done by their cattle. In fact they are strictly liable, which means that legally it doesn't matter whether the rancher was negligent or not. Therefore, the cost of any damage is borne by the owner of the animals. In this case, ranchers must decide whether it is worthwhile to try to keep the cattle in.

Ellickson knew about the Coase Theorem, and he first became interested in Shasta County because he wanted to find out whether the rule of liability had any effect on the allocation of resources, such as decisions about building fences.

Here is the bottom line: legal rules do seem to be mostly irrelevant to resource allocation decisions (as Coase predicts in the case of zero transaction costs). However, RE found that this was not for the same reasons as Coase gave. For example, he found no explicit bargaining between neighbors over transfers (as the reasoning behind the Coase Theorem would lead you to expect) and no situations where the legal rules determined the starting point for bargaining (again, as Coase's reasoning might lead you to expect).

In fact, RE found that people typically don't know what the law says. Instead they tend to cooperate and resolve disputes through social norms. The social norms in Shasta County say that people should not hire lawyers when a conflict occurs and that people should not generally give cash compensation or transfers to their neighbors.

Even so, RE believes that the social norms do lead to economically efficient behavior. Specifically, they tend to maximize the total income or welfare of the people involved. We'll see later in the book why he thinks this.

Note: Ostrom was mainly concerned with whether we observe institutions that succeed or fail. Ellickson is different because he is not concerned about failures. Instead, he studies two possible ways of succeeding (Coase Theorem versus social norms), and he wants to know why social norms provide a better description of the real processes through which resources are allocated.

Here are some big picture implications.

1. Hobbes is wrong (recall Ostrom's discussion of Hobbes). We don't necessarily have chaos in the absence of government. Of course, we already know this from the case studies of pre-state societies in Johnson and Earle. Rather, social norms can provide a lot of order in the absence of a state. Even when a state exists (such as the state of California, where Shasta County is located), people may ignore it and settle disputes in other ways, instead of using police, courts, and lawyers.
2. There are many examples of order without law, where unplanned or spontaneous behavior leads to stable patterns for social systems as a whole. Here are a few:
 - (a) Languages evolve over time, and people can communicate with each other, even if no one passes any laws regulating the languages people use.
 - (b) Cities tend to evolve over long periods of time in ways that are not determined by the law, but rather by economic incentives and social processes. For example, no one in the 1800s made any decision about what the city of Vancouver should look like in the 21st century. But gradually residential areas developed, business areas developed, a downtown core evolved, and so on.
 - (c) Markets can often allocate resources quite well without any central planning. If you are familiar with competitive markets (supply and demand), you know that it is often possible to predict how prices and quantities will change in response to a change in some exogenous variable, even though each individual person takes the market price as given and ignores their own influence over it.

None of this is news to economists (especially not the last point). However, economists probably pay too little attention to social norms, and we don't have a lot of theory about them. So we may be able to learn some things from RE's book.

I don't want to exaggerate the point here. Government is not useless or evil. Often there are collective problems that can only be solved through law, taxation, regulation, etc., or solved most efficiently in these ways. But we don't want to overlook the potential value of social norms, especially in small-scale communities where people tend to stay around for long periods of time.

Chapter 1. This chapter gives a short description of Shasta County.

Before I get into this, I want to make a general point. One thing I hope you have learned in this course is that the local features of a place, such as its climate, geography, and other characteristics of the natural environment, can be important if you want to understand the economic problems people face, and the institutions they develop to solve their problems. This is certainly true for Shasta County.

Shasta is in northern California. If you want to drive there from Vancouver, you can get on Oak Street, go to the US border, and get on interstate 5. You will drive south through Washington and Oregon and eventually you will reach California. You know you are in Shasta County when you reach Redding, which is the largest city in the county. It takes about two days of driving. I don't know the area well, but I have stopped a few times to have lunch and get gas. It seems like a nice place.

The region has a rainy season from December to May. It is dry in the summer and very hot. The land is generally bad for farming. Common activities include cattle ranching and forestry (at the higher elevations). There are three main types of land.

The *grasslands* are between 500-1500 feet above sea level (sorry for not using the metric system). As the name suggests, grass grows well there and it is possible to raise cattle all year as long as you use irrigation during the dry months (there is plenty of natural pasture in the spring so irrigation is less important then). However, some places cannot be easily irrigated, so they are less valuable for ranching.

The *foothills* are between 1500-3500 feet above sea level. Instead of grass there are trees and open scrub. This area is less good for cattle ranching because the land would have to be cleared and maintained.

The *mountains* are above 3500 feet. There are forests and it is cold in the winter. Also it is hard to clear the land for ranching. Most of the land is owned by the government or by private timber companies, but it can be leased by cattle ranchers for summer grazing.

Some old pioneer families own very large ranches in the grasslands (as you might expect, the most economically valuable land was settled first). Today this land is mostly fenced in, because this is in the interests of the ranchers themselves. For example, they want to protect their cattle from predators like mountain lions and coyotes, they want to keep out humans who might steal their cattle, they want to prevent the cattle from wandering into deep snow, and they don't want the cattle to eat poisonous plants. In addition to having fences, this land is mostly irrigated.

At the time RE was doing his research, the population of the foothills was growing. The existing landowners were often willing to convert their land for residential development, and people moving into the area from large cities like San Francisco or Los Angeles often preferred to live in the foothills because this area is cooler in summer and has more trees.

Newcomers in the foothills have smaller lots (5-40 acres) than the professional ranchers. These lots are often called "ranchettes". They are not used for commercial ranching; they are primarily places for people to live, although the owners may keep a few animals as a hobby. Most of the foothill ranchers have several "ranchette" neighbors.

Among the ranchers, there are two main groups:

The *traditionalists* let cattle roam in unfenced mountain areas during the summer, and then round them up in the fall. This requires a lot of land.

The *modernists* keep their cattle fenced at all times and irrigate their pastures as needed to ensure year-round use.

Fences are expensive. In 1982, a fence cost \$2000 per mile just for the materials. In addition to this, there were labor costs, which generally doubled the expenditure (unless the people building the fence did the work themselves). These were not trivial costs for the residents of Shasta. Beyond the cost of building the fence, there would also be costs for maintenance, monitoring the condition of the fence, and making emergency repairs.

Ranchette owners tend not to build fences because they have few (if any) animals and it is an expensive thing to do. This leads to the question of what happens when cattle stray into unfenced areas, and who pays for fences when they exist? Before exploring these issues, I will make a few comments on the history of open and closed range, as well as chapter 2 on the political conflicts that have arisen over this issue.

History.

In 1850 when California became a state, the entire state was open range. Someone who suffered damage from animal trespass could only receive compensation in court if they had a "lawful fence". This means that the victim had a fence that satisfied various legal requirements but the animals got through the fence and caused damage anyway. In that case the law took the view that the victim had done their best to avoid damage, so if the damage still happened, the owner of the animals owed compensation to the victim. But in the absence of a lawful fence, the owner of the animals had no legal liability, and any damage was the victim's problem.

With increasing settlement and more agriculture, by the late 19th century California had closed the range in certain counties. This provided more legal protection for farmers and shifted the costs of damage (or fences) over to the ranchers.

In 1915 most of California went to strict liability for owners of livestock, except for six counties in northern California where population density remained low and ranching was still important. Shasta was the least rural of these six counties.

In 1945 the state gave Shasta local autonomy: the county government was free to decide what areas would be open range and what areas would be closed range.

Chapter 2. The politics of range closures.

I won't go into all the details, but this chapter gives the flavor of the political fights that occurred in Shasta over the issue of whether some areas of the county should be closed.

Here is an interesting question in thinking about this chapter. Suppose RE is correct that open versus closed range has very little effect on who pays for the damage done by cattle (see chapter 3) and who pays for fences (see chapter 4). Then why did the residents have huge political battles (see chapter 2) about "Caton's folly" (an area that was changed from open to closed) and Oak Run (an area for which an attempted closure was defeated)? The answer is not obvious. I will come back to this issue when I discuss chapter 6.

Chapter 3. What does the law say about animal trespass? What do social norms say?

First I will summarize what the law says, according to Ellickson.

Closed range. The principles here are derived from English common law. The owner of the animals is responsible for any property damage they do. In fact, the owner is strictly liable, so it doesn't matter whether the owner was being careful or not.

Open range. The owner is not responsible for damage, even if the owner is negligent (so the owner of the animals was not being careful to keep them under control). But there are three exceptions to this rule: (a) it does not apply to goats, swine, or vicious dogs; (b) the owner is strictly liable if the victim had a lawful fence but the animals got in anyway; (c) the owner is always responsible in cases of intentional trespass (you cannot deliberately allow your animals to enter someone else's property).

In cases where animals do damage and the owner is liable, the victim has three possible legal remedies: ask for compensation in court, ask for an injunction from a court (this is granted when there have been repeated violations, and essentially tells the owner of the animals to stop it), or seize the animals, which involves notifying a state official.

What if you are a victim but you don't want to go to court (maybe this is too expensive)? Does the law allow you to do anything yourself? Yes, you could do several things. First, you can use force to drive away the animals. Second, you can herd them to a place where they won't do too much damage. Third, you can seize control of the animals but then you must care for them. You are never allowed to harm the animals.

Do people know what the law says? Not much.

1. Landowners mostly know whether their own land is open or closed.
2. They generally don't know the details (such as the various exceptions to the rules). They just assume that in open range the owner "has the rights" and in closed range the victim "has the rights". This is a good approximation 99% of the time.

3. Legal specialists tend to believe negligence must be important, because this issue is frequently important in other areas of the law. However, it does not matter for the question of open versus closed range.
4. Specialists don't know where the open and closed areas are located. They could probably find out, but they don't need to.
5. Insurance company representatives don't know and don't care (we'll come back to insurance issues later).

What do the social norms say about animal trespass?

First I will list what RE calls *substantive* norms. These are:

- (a) The owner of the animals is responsible (regardless of what the law says).
- (b) Victims should overlook minor damage from isolated incidents.
- (c) When the owner of the animals does provide compensation for damage, it should be done "in kind" (for example, by doing a favor for the victim) rather than by a cash payment.

Next I will list what RE calls *remedial* norms (what victims should do if the owner of animals does not provide adequate compensation for significant damage). These are:

- (a) Don't rely on the law (only weird people hire lawyers).
- (b) Use self-help (I'll explain this below).
- (c) If the situation is sufficiently bad, complain to public officials.

For minor incidents, a victim would normally call the owner and report the problem. The owner would come and take the animals away, while apologizing to the victim. This is a common thing and it is not a big deal.

When the owner of some animals does not behave 'properly' as described above, or when there are repeated significant problems, the victim goes to the remedial norms. Usually this involves "self-help". The lowest level is to gossip to other people about the problem and hurt the reputation of the animal owner. The next level is that when animals come into your land and do damage, you can put them someplace inconvenient for the owner. A third level is death threats (to the animals, not the people), or "rustling" (stealing the animals and selling them, while pretending to know nothing about what happened).

If self-help isn't working, victims can complain to public officials. Long time ranchers don't do this, but ranchette owners do. This is not surprising because many of them are from large cities and they are in the habit of calling up city officials or police if there is a serious problem. The result of calling a public official is that the official gives a warning

to the owner of the animals. They may threaten to take the animals away or prosecute the owner (although this never happens). Officials may also threaten to close the range.

Victims almost never use the law or try to get cash compensation. This norm is valuable because everyone can avoid legal costs, as well as the destruction of trust that occurs in a long-term relationship when someone hires a lawyer.

Chapter 4. Now let's shift gears and think about the question of who builds fences and pays for them. From the standpoint of two neighbors, a fence is often a public good in the sense that both people benefit from it. But there is potentially a free rider problem because each person would like the other to pay.

RE found in Shasta that the law (open versus closed range) was irrelevant for whether a fence would be built. If you know about the Coase Theorem, this is not a surprise (with zero transaction costs, it is what Coase predicts). What is more surprising is that the law was also irrelevant for the question of who pays for the fence. This is not consistent with the logic of the Coase Theorem, where a change in the rule of liability should have some effect on the distribution of income (see my previous lecture notes about this).

What does the law say about building fences? Not much. There are two main points:

- (a) If you enclose your land and partly rely on a fence that was previously built by your neighbor, you owe some compensation to the neighbor.
- (b) If both lands are enclosed, the costs of fence maintenance must be shared between the neighbors.

According to RE, there are no lawsuits about fences and no one knows what the law says.

What do social norms say about building fences? Several things.

The most important is the "norm of proportionality". This says that costs should be shared roughly in proportion to the number of animals on each side. In practice this is not exact but there are a few basic rules. If the neighbors are both full time ranchers, they should share the costs 50/50. If one neighbor has many animals while the other has few or none, the person with the large number of animals pays the entire cost. In complex cases, there may be some compromise that splits the costs. For example, one neighbor may supply the materials for the fence while the other supplies the labor (this goes back to an example discussed in my previous notes where one neighbor knows how to build a fence while the other does not).

A related norm is that one neighbor should not pay cash for personal labor by the other. However, it is acceptable to offer cash as compensation for out of pocket expenses (like the materials needed to build a fence).

A third substantive norm is that on big projects, you should consult in advance with your neighbor instead of just going ahead and making a decision on your own.

As in chapter 3 there are also remedial norms. For example, if someone does not invest the time or money they should be providing in order to maintain an existing fence, their neighbor might start gossiping about it, or escalate to more serious forms of retaliation.

Chapter 5. What about highway collisions involving a driver and an animal?

This issue is different from animal trespassing and fence construction or maintenance in two key ways. First, the law is relevant (there are lawsuits). Second, negligence issues do matter under the law.

A quick summary: according to RE the ranchers have incorrect ideas about the law. They believe that whether the range is open or closed has a serious effect on legal liabilities for a highway collision where a driver hits an animal. In particular, they believe that in open range the driver is more likely to bear the legal liability, while in closed range the rancher is more likely to have the legal liability. This might help explain why they oppose efforts to close the range, as in chapter 2.

The problem is that according to RE, they are wrong. What the law actually says (and keep in mind that RE is a law professor, so he should know) is that liability in a highway collision is determined by negligence (who is at fault). For example, if the animal is in a fenced lane and the owner of the animal was being careless by allowing the animal to be there, then the owner is at fault. But if the driver of the car was speeding under dark and rainy conditions, then the driver is at fault.

The ranchers also believe that open versus closed range has a serious effect on what they have to pay for insurance. Again, according to RE they are wrong. There is no evidence that the location of a ranch in open or closed range has any significant effect on what the ranchers pay for insurance. The insurance companies ignore this factor when they set the level of the insurance premium. Instead, the companies normally use negligence ideas to determine the premiums (if you were at fault in a highway collision, it is likely that your insurance premium will go up).

Note: in chapter 6 RE says there may be some (very small) grains of truth in the ranchers' beliefs. In particular, he says that range closures might lead to more insurance claims and therefore higher premiums or a need for more insurance coverage. However, he still says that these effects are greatly exaggerated by the ranchers.

Chapter 6. So why do ranchers care about range closures?

And why do they continue to have false beliefs about the effect of range closures on their legal liabilities and their insurance costs?

Another puzzle: why do the modernists support traditionalists in fighting against range closures (as in chapter 2), even though they have already built fences voluntarily?

Note that although the beliefs of the ranchers are mostly false, they have consequences for the allocation of resources. When there is a closure, the traditionalists tend to move their cattle away from the affected area. This imposes costs on them: presumably what they were previously doing had lower costs, which is why they were doing it, and when they changed their practices, they were probably shifting to something more expensive. Furthermore, there is a loss of grazing fees for people who own unfenced pasture in the land that is now closed, because the traditionalists aren't leasing that land any more.

It is also costly for modernists to spend time and money fighting a closure (for example, they have to attend public meetings about it), although it doesn't directly influence their ranching operations.

An economist would probably think all of this is quite strange. It doesn't make much economic sense for people to continue holding false beliefs despite a steady stream of information contradicting those beliefs, especially when holding the beliefs is costing them time and money.

The ranchers dismiss the claim that their beliefs are false. What they tend to say is that the lawyers and judges are incompetent or corrupt, the insurance companies don't know what they're doing, and so on.

This is a standard phenomenon in psychology called "confirmatory bias". We all tend to notice information that confirms beliefs we already have, and tend to ignore information that contradicts our current beliefs. When someone points out the conflict between what we believe and what the evidence says, we tend to claim that the evidence is not 'real' or look for reasons to downplay its significance. (Exercise: try to think of cases where you have done this yourself. I know I have.)

There is nothing especially unusual about confirmatory bias. However, it does conflict with the standard economic assumption that people update their beliefs in rational ways when they get new information. (Of course, an economist with confirmatory bias might try to downplay the importance of the psychological evidence on this subject, claim that the psychologists are doing bad research, and so forth.)

Another side point: psychologists have found that people who have more education tend to have more confirmatory bias. One possible reason is that people with more education are better at inventing alternative explanations for evidence that contradicts their beliefs, so they are less likely to question their own beliefs. They may also be more confident in their reasoning abilities (whether this confidence is justified or not).

In any event, RE's theory about the behavior of the ranchers is that they aren't concerned with dollar costs. Fighting over range closures is more about emotions and symbols.

Remember the social context: at the time RE was doing his research, the population of Shasta County had been expanding rapidly. Redding had recently become a city, many non-ranchers were moving to the area, and Shasta was increasingly urban and suburban.

In this context, both the traditionalists and the modernists were concerned about a loss of social status and political power. They interpreted range closures as a "kick in the teeth" and fought them for these reasons. RE does not say that the ranchers were explicit about any of this, or even that they were consciously thinking in this way, but he believes such attitudes influenced their behavior.

One last question from an economic standpoint: why don't market forces weed out the false beliefs eventually? You might think that if the owner of a firm holds false beliefs about the world, they will have lower revenue or higher costs than the owners of firms who hold correct beliefs. Maybe in a competitive market, firms owned by people with correct beliefs will drive the others out of business in the long run.

RE makes two points about this. First, he says there really isn't that much at stake in terms of dollars. Yes, the ranchers with false beliefs have higher costs than necessary, but not by a lot.

Second, he argues that most ranchers are getting rents (think of the rent as the difference between the revenues they actually get and the minimum amount they would be willing to accept in order to stay in business). This is particularly likely to be true for ranchers who own valuable land. Their false beliefs reduce their rent slightly, but will not cause them to be driven out of the market due to competition from other ranchers.

I think there is a third factor that may be important. The supply of accurate information is a public good. If someone devotes time and money to finding out what the effects of the liability rules really are, or how insurance companies really set their premiums, that person must incur all of the cost of doing this research. However, once the information spreads around, it will be beneficial for other people (positive externalities). This is the type of situation where I expect a free rider problem, in which voluntary contributions to a public good lead to a supply that is too small from a social perspective. Because there is too little investment in generating accurate information, false beliefs tend to persist.

One final thought: why do we happen to know that the ranchers' beliefs are false? Only because RE was a professor who wanted to do the research and had the right expertise to understand the legal rules. I don't know whether any of the ranchers in Shasta have read his book, but if they have, maybe RE helped to overcome the free rider problem. On the other hand, confirmatory bias might cause them to dismiss his arguments.

Econ 354

Greg Dow

March 14, 2021

Notes on Robert Ellickson, "Order Without Law"

Chapters 7-8 (third set of lecture notes)

These two chapters begin Part II of Ellickson's book. In these chapters he provides some general theoretical background on the subject of social norms. Later on he will construct his own theory about norms (where they come from, what they accomplish, and so forth). My notes here will be relatively short and will only cover the main points.

Chapter 7 (systems of social control).

RE is interested in the question of why there is so much cooperation in human societies, more than can be explained by the existence of laws.

This causes him to explore systems of social control at a more general level. A system of social control is a set of rules about normatively appropriate behavior. Some behavior is defined as 'good' and is rewarded; some behavior is defined to be ordinary or neutral, and does not receive any reward or punishment; and some behavior is defined as 'bad' and is punished. It makes sense to treat most behavior as ordinary, because then the system of social control doesn't have to pay any attention to it.

RE distinguishes five types of social control (think of these as alternative mechanisms):

1. Self-control (ethics, morality)
2. Second party control (such as self-help in Shasta, or self-enforcing agreements in a repeated prisoner's dilemma game)
3. Social forces (decentralized networks, social norms)
4. Organizational control (such as rules for employee behavior in a firm)
5. Governmental control (laws and regulations)

RE defines government as a hierarchical organization generally regarded as having the authority to inflict detriments on people who live in a given territory, and who have not necessarily accepted this authority.

This is a fancy way of saying that governments can punish people even if those people don't accept the authority of the government.

If we are interested in social control, how do we know that a rule about behavior exists?

One thing we can do is observe whether the behavior is rewarded or punished. However, this may not be very useful in situations where people are deterred from engaging in bad behavior because they know it will be detected and punished. In that case, there is a rule but we will not see punishment very often because deterrence works.

Other things we could do are look for systematic patterns of behavior, or seek statements by the people involved about what the rules are.

Note: courts often enforce rules that would usually be self-enforced, or enforced through social norms. So if there is an occasional court case, the outcome might reflect patterns that are widespread within the community (see the whaling case study later in the book).

Next RE describes different types of rules (there are a lot of lists in this chapter):

1. Substantive rules (rules that directly control a specific kind of behavior)
2. Remedial rules (what people do when a substantive rule is violated; an example would be the graduated penalties discussed by Ostrom)
3. Procedural rules (this involves things like what counts as evidence, who has the burden of proof if there is a dispute, and so on).
4. Constitutive rules (this is a fairly abstract idea but roughly it means the rules that set up a controlling body; for example, governments are set up by constitutions).
5. Controller-selecting rules (this means rules about which controlling body has the jurisdiction to deal with a specific issue: for example, if there is bad behavior at SFU, is this a problem for the university administration or for the police?)

Chapter 8 (theories of social control)

Ellickson discusses two views about the role of law.

1. Legal centralism. People in this category believe that the law is central to the process of social control. RE is quite critical of this view, as you would expect from the title of his book. His criticisms include the following points.
 - (a) The government may not have a monopoly on the use of force.
 - (b) People often don't know the details of the law.
 - (c) Governments may not observe some violations, or may choose to overlook them.
 - (d) In the private sector, people are frequently punished for bad behavior by the loss of their reputation or by having other people end a relationship. This can create a lot of order without requiring any use of the law.
2. Law and society. People in this category believe that norms are the main system of social control. You might think RE would have some sympathy for this view. But actually he is critical of it because people in this group tend to treat norms as exogenous (they don't explain why a society has a particular set of norms, rather than some other possible set of norms). This is disappointing for RE, who wants

to construct a theory to explain why certain norms exist (he wants them to be endogenous). He also argues that the law is not completely irrelevant.

- (a) The law has some influence on norms (consider legislation about civil rights or environmental protection, which may influence how people think about what is good or bad behavior).
- (b) Changes in the law do have consequences (consider the range closures in Shasta County, which did influence people's behavior even if their beliefs were false).

In general, RE doesn't like either approach very much. He calls them a desert (legal centralism) and a swamp (law and society).

Now let's look at how social scientists have thought about the content and evolution of norms. RE distinguishes several broad schools of thought.

1. Functionalism. This is the idea that norms exist because they contribute to the survival of a society or benefit most of its members. This is similar to the idea of integration theory in Johnson and Earle, where the state existed because it solved social problems. RE is critical of this framework because it doesn't give strong answers to questions like these: What is the causal mechanism that leads to the evolution of good norms? Are all norms functional, or can there be bad norms? Will societies really fail to survive if they have non-functional norms? Also the theorists in this group tend to ignore the motivations and welfare of individuals, focusing instead on society as a whole. Specifically they tend to ignore conflict. Overall, RE thinks the arguments in this category are theoretically loose.
2. Interest groups/power. One alternative to functionalism is to say that powerful groups impose norms on society or manipulate their content. This argument was common among Marxists, who argued that the working class suffered from false consciousness due to propaganda by capitalists, that religion was the opiate of the masses, and so on. However, it has also been a common argument among other social scientists. In a way, it is a modified form of functionalism: instead of just looking at what is good for society as a whole, you look at what is good for some powerful subset of society. RE again makes several criticisms: people need to be specific about how the causal mechanisms operate; there are various widespread norms that don't seem to fit the theory (such as norms about honesty, reciprocity, and so on); and it is not always obvious when or how social control occurs.
3. Genetics and evolutionary psychology. In this view, some norms are hard wired in humans due to our biological evolution. For example, some people argue that we rely on the same ethical principles as our foraging ancestors because humans evolved in this type of society, and we continue to use 'forager ethics' even though we don't live in foraging societies today. It is not clear whether this can explain cooperation beyond what would normally be expected due to kinship (the genetic relationships among family members). Obviously humans have some capacity to learn social norms and obey them, but it is a lot less obvious whether biology can

explain their detailed content (norms often vary across countries or historical time periods). We will come back to these questions when Putterman discusses human nature in his book.

4. Rational actor theories. I will talk much more about this in my next set of notes on chapters 9-16. The short version is that in theories about rational actors, it is assumed that people are self-interested and rationally choose means to achieve a given end. The idea that people are rational actors is central to most of economics and game theory. For example, it is the basis for the study of repeated games and for the Coase Theorem. RE is willing to be a little flexible on how he defines a rational actor. He allows the possibility that people may be 'boundedly rational' (they are not infinitely fast calculating machines) and thus learning, culture, and evolution are important. He also says that they may have cognitive biases (such as ignoring new information that contradicts things they already believe).

That's all for chapters 7 and 8.

Econ 354

Greg Dow

March 21, 2021

Notes on Robert Ellickson, "Order Without Law"

Chapters 9-16 (last set of lecture notes)

These notes are mainly about Ellickson's own theory of social norms. I summarize some of the main points, but you need to read the book to get the full story. Often RE provides important examples, details, or explanations that I will skip over here.

Chapter 9.

In this chapter, RE deals with the question of how cooperation is possible.

He assumes that people are "rational actors". He defines this to mean that they are

- (a) self-interested and
- (b) rational in selecting the best means to achieve a given end.

These ideas are consistent with what we usually assume in economics or game theory.

However, as I mentioned at the end of my previous notes, RE does allow that people are "boundedly rational" so their actions are influenced by learning, culture, evolution, etc.

He also recognizes that people can have cognitive biases -- for example, they may not update their beliefs in a rational way when new information arrives. This can include what is known as "confirmatory bias", where people pay attention to information that confirms their existing beliefs while ignoring information that contradicts their beliefs.

RE wants to construct a theory about social norms (we will get to the details in chapter 10). In chapter 8, he surveys and criticizes other people's theories about social norms. My personal interpretation is that in chapters 9 and 10, RE is attempting to build a bridge from rational actor assumptions (especially individual self interest) to functionalism (the idea that social norms promote aggregate or collective welfare).

One crucial question that any such theory must answer is: under what circumstances do people behave in cooperative ways even though they have individual incentives not to?

For about the last 50 years, there has been an enormous amount of research in the social sciences on questions of this kind. A key idea in this research is the prisoner's dilemma game, because it involves a dramatic conflict between individual self-interest and social

welfare. RE knows that his theory about norms will not be persuasive unless it applies to the PD game in some way.

Recall at this point our earlier definition of a PD game: it is any game where each player has a dominant strategy, so there is a dominant strategy equilibrium; but the DSE is not Pareto efficient (starting from the DSE, one can find a Pareto improvement).

RE uses essentially the same definition but he adds one more assumption: the sum of the payoffs to the players when they both cooperate is larger than the sum of the payoffs when one player cooperates while the other defects.

In order for this to make sense, it must be possible to add up the payoffs of the players in a meaningful way. For example, one could do this if all payoffs are measured in dollars, or some other objective unit.

A quick economic example: suppose Coke and Pepsi are competing. They play a PD game where 'cooperate' means 'charge the monopoly price' and 'defect' means 'cut your price a little bit to gain market share'. In a model like this, it would normally be true that if both charge the monopoly price, total profit for the two firms is higher than if one firm charges the monopoly price while the other cuts price a little bit. So this would fit RE's assumption about the payoffs.

Now consider RE's version of the PD game (look at Table 9.2 on p. 161 as you read these notes). We have $A > B > C > D$, which means the game satisfies my definition of a PD. The dominant strategy equilibrium is (defect, defect), which yields payoffs (C, C). This is not Pareto efficient because both players would be better off with (B, B).

Remark: RE's algebraic notation is confusing. People often abbreviate the PD strategies by C for cooperate and D for defect. However, when RE says the payoffs are (C, C), he is not saying that both people are cooperating, he is saying that both people are defecting.

As mentioned above, RE assumes that the sum of the payoffs if both cooperate is larger than the sum if one cooperates while the other defects. This implies $2B > A+D$.

Why does he do this? He wants to talk about compensation payments by one player to the other (these are sometimes called 'transfers', 'side payments', or 'gifts'). For example, if I defect today, I might make up for it by doing something nice for you tomorrow, and this may involve giving you money or doing you a favor. Such payments transfer utility from one person to the other, while keeping total utility constant.

In order to understand this, see my graph in Figure 1. The payoff or utility of person 1 is on the horizontal axis, and the payoff or utility of person 2 is on the vertical axis. I have indicated the payoffs for each of the four possible strategy pairs in the PD game. The line through each of these points shows a set of points where the sum of the utilities remains constant ($u_1 + u_2 = U$ where U is constant). The sum of the utilities U is what RE later calls 'aggregate welfare'.

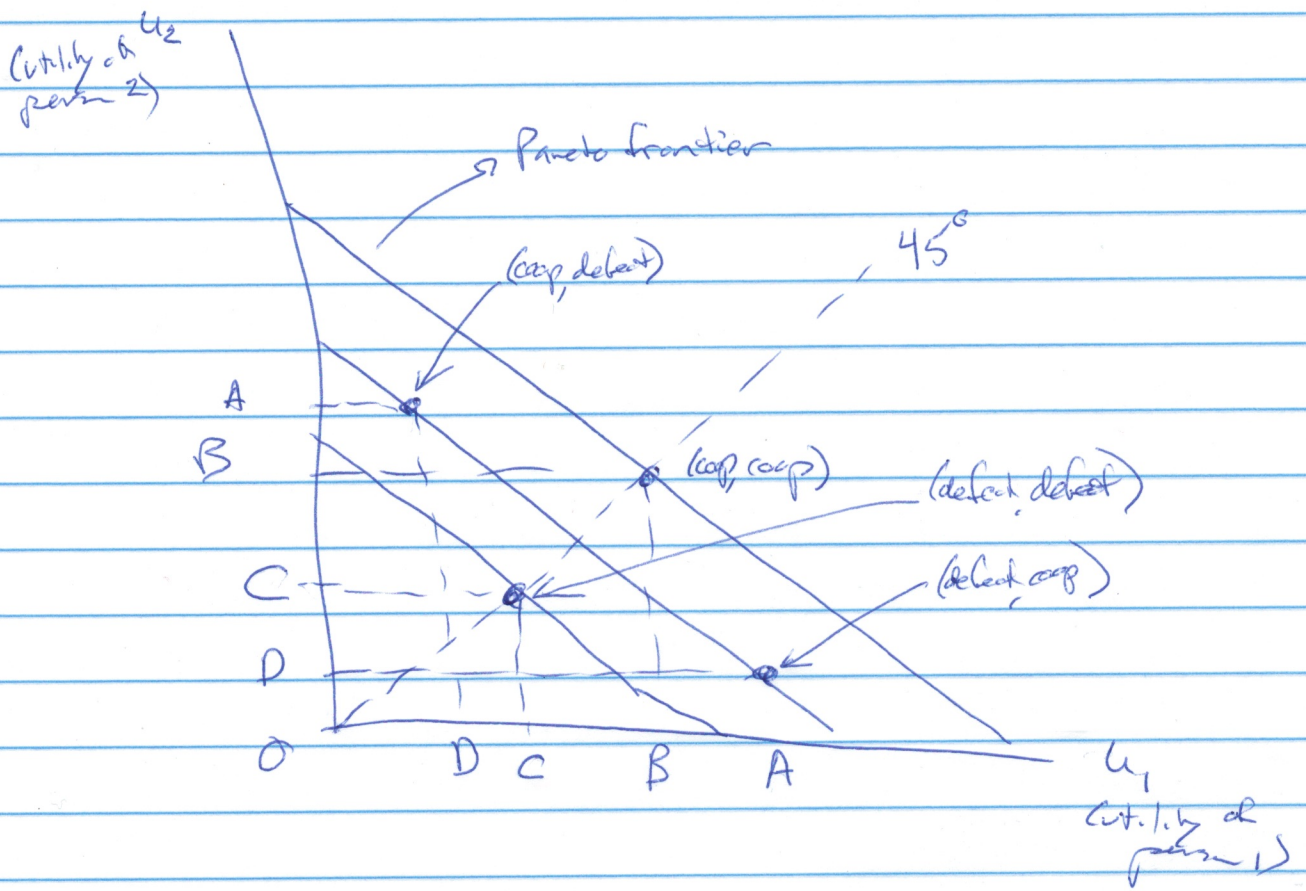


Figure 1

Lines of Constant ~~total~~ Aggregate welfare $(u_1 + u_2) = u$
 where u is a constant

For example, (coop, coop) yields a total payoff of $2B$. If we look at the line described by $u_1 + u_2 = 2B$ (or $u_2 = 2B - u_1$), we can see all of the points where the sum of the payoffs is equal to $2B$. We can achieve any point along this line by first going to (coop, coop) and then transferring utility units (for instance, dollars) from one person to the other. Notice that the slope of the line is -1 because when we give one more utility unit to one player, we have to give one less utility unit to the other player.

The same idea applies if one player cooperates while the other defects. In this case, the sum of the payoffs is $A+D$. Because (coop, defect) and (defect, coop) yield the same sum $A+D$, these two strategy pairs are located on the same line. Again the slope is -1 so this line is parallel to the one through (coop, coop). However, the $2B$ line is above the $A+D$ line because it involves a higher total payoff.

Finally, we have a line through (C, C) which is the payoff result when both defect. The total payoff is $2C$. I assume $A+D > 2C$ so the line for mutual defection is the lowest.

If the players could sign legally binding contracts and bargaining was easy, they would always agree to choose (coop, coop). This would ensure that they reach the highest line (this is called the *Pareto frontier*). Then they could bargain over side payments in order to determine which specific point along the $2B$ line they would go to.

No point below the $2B$ line can be Pareto efficient, because it is always possible to go to the $2B$ line instead, and then distribute the resulting total payoff in a way that achieves a Pareto improvement by making both people better off relative to the starting point.

Notice that there is a strong resemblance here to ideas from the Coase Theorem. The main difference is that in the Coase Theorem, legal rules (the rules of liability) set the starting point, and then bargaining takes the players to the Pareto frontier if necessary. To see this, look at the graphs I drew for the farmer and the rancher when I was talking about the Coase Theorem in my previous notes, and compare them with Figure 1 here.

RE thinks legal rules may not be relevant in determining the starting point for bargaining. He is vague about what does determine the starting point, but it is reasonable to think that in his PD game, it might be the payoff pair (C, C) from the dominant strategy equilibrium (defect, defect).

A closely related game is what RE calls the 'specialized labor game' (see pp. 162-164). Although he discusses this in chapter 9, I think it is better to come back to it after we understand his hypothesis about social norms in chapter 10.

Chapter 10.

RE's hypothesis about social norms is as follows:

"Members of a *close-knit group* develop and maintain norms whose content serves to maximize the *aggregate welfare* that members obtain in their *workaday affairs* with one another." (p. 167).

In order to understand what this means, we need to define and understand each of the three italicized concepts.

Close-knit group (pp. 177-178). RE defines this to be a group where informal power is broadly distributed among the individual members, and information relevant for social control circulates easily among the individual members.

Aggregate welfare (pp. 170-174). As mentioned earlier, RE thinks of this as the sum of the payoffs or utilities for the group. For this to make sense, we need an objective way of adding up the payoffs or utilities of the individual members. For example, payoffs may be measured in dollars (prices, costs, revenues, wages, profits, or rents); there might be 'shadow prices' of untraded goods, which give an approximate valuation or equivalence with things that are traded; or there might be some approximate equivalence between the favors exchanged between people. He is also willing to consider the value of time, and to compare the amounts of time people spend doing various different things. The details are a little vague, but RE would argue that his theory applies as long as there is some way of quantifying the sum of the payoffs and some way of transferring individual payoffs from one person to another.

Note on this: when computing aggregate welfare, we have to subtract off the transaction costs of running the system (search costs, bargaining costs, and enforcement costs). It is the aggregate welfare net of transaction costs that matters. For instance, suppose it would be possible to have (coop, coop) but this would require a lot of time monitoring behavior. It might be better from a social point of view to choose (coop, defect) instead, spend less time monitoring behavior, and have person 2 compensate person 1 in some way.

Workaday affairs (pp. 174-177). This implies a focus on everyday life, especially the economic activities people engage in. It would include practical matters and ways of making a living. RE is not focusing on norms about family life, religious behavior, the way people behave in crisis situations, etc.

More generally, RE wants to exclude the following issues:

- (a) rules that enable exchange to occur (e.g. norms about basic property rights).
- (b) purely distributional norms (he is concerned with situations where members of a group have roughly similar incomes, power, status, and so on; he is not focusing on norms that support inequality or social stratification).

Now let's talk about the prisoner's dilemma game. Assume that legally binding contracts are impossible. People cannot call on the police, the courts, or government regulators to solve their problem (or it would be too expensive to use these solutions). Why does RE believe that a close-knit group would cooperate?

RE's answer: efficient norms evolve over time. He bases this idea on a famous research project by Robert Axelrod in a 1984 book called "The Evolution of Cooperation". What Axelrod did was to use a PD game that was repeated 200 times and then ended. Observe that this is a *finitely* repeated PD game, not the infinitely repeated game we discussed for the Ostrom book.

Axelrod invited game theorists, economists, and other social scientists to submit their favorite strategies for playing this game. Then he ran a tournament on a computer, where each strategy had to play against every other strategy. The strategy achieving the highest total number of points was the winner.

People could submit very simple strategies, such as "always defect", or very complicated strategies, such as "cooperate in the first round, and after that, cooperate if the opponent has always cooperated in every odd-numbered round, but defect otherwise" (and so on).

The winner was a simple strategy called "tit for tat" (TFT). In this strategy, the player starts by cooperating in the first round, and then copies the opponent's most recent past move in all subsequent periods. So if your opponent cooperated in the most recent past round, you cooperate in the present round. If your opponent defected in the most recent past round, you defect in the present round.

Why does this work well?

- (a) TFT is 'nice'. If you meet someone else who is willing to cooperate, then you can enjoy the gains from cooperation. TFT is never the first to defect.
- (b) TFT is 'provocable'. This means that if someone tries to take advantage of you by defecting, you retaliate by defecting. This imposes a cost on the opponent, which may get them to return to cooperation. The punishment is immediate (no delay).
- (c) TFT is 'forgiving'. Once you have made your point (you won't tolerate abuse), then you are willing to go back to cooperation. If the opponent is now willing to cooperate, you can return to getting large payoffs. Notice that this differs from a trigger strategy, which punishes one defection forever and therefore sacrifices all of the potential benefits from future cooperation.

Another good thing about TFT is that it is simple, so other people can easily figure out what you are doing and respond accordingly. This feature didn't matter in the computer tournament, but it could be important in real life.

After running the initial tournament, Axelrod did an evolutionary version, which can be interpreted either as natural selection or learning/imitation. In this version, he started off with a population of players using a variety of different strategies. Pairs of players would meet each other randomly and play for a while. Then Axelrod made more copies of the

most successful strategies, fewer of the less successful ones, and restarted the process. Again TFT did very well, in the sense that it often took over the population eventually.

Based on these results, RE is optimistic that a close-knit group can develop norms that support cooperation, even in a PD with strong individual incentives to defect.

Next I want to go back to the "specialized labor game" (look at pp. 162-164, especially Table 9.3 on p. 162). Now we have the payoffs $A > B > C > D > E$. The strategies have been relabeled as 'work' (contribute effort) and 'shirk' (don't contribute effort). There is still a dominant strategy equilibrium with payoffs (C, C), which is not Pareto efficient because (B, B) is better for both. Hence this still satisfies my definition of a PD game.

However, there are a couple of differences from the previous PD game. One change is that the payoffs of the players are no longer completely symmetric because $D \neq E$.

A second (more important) change is that in this game, RE assumes $A+D > 2B$. For this reason, aggregate welfare is maximized when player 1 works and player 2 shirks (notice that $A+D > 2B > 2C$ and also $A+D > A+E$).

The idea here is that in order to maximize aggregate welfare, it is best to have player 1 do some work while player 2 does nothing (maybe player 1 knows how to build a fence but 2 does not, or player 2 has high effort costs, or player 2 has other valuable things to do).

The problem is that if we do the aggregate-welfare-maximizing thing, we go to (work, shirk) where player 1 gets D and player 2 gets A. This is worse for player 1 than getting C, which is what would happen in the dominant strategy equilibrium. So player 1 does not want to do the work unless she receives some compensation from player 2.

Now look at my Figure 2. Let $t \geq 0$ be a transfer or a gift from player 2 to player 1. This transfer is paid if and only if the strategy choices are (work, shirk). No transfer occurs in any other situation.

What would be necessary in order for (work, shirk) to be an equilibrium? We need to be sure of two things. First, given that player 2 will shirk, player 1 must be willing to work. This is true whenever $D + t \geq C$. Second, given that player 1 will work, player 2 must be willing to shirk even though this requires that the transfer be paid. This will be true when $A - t \geq B$. Putting these inequalities together, we have $C-D \leq t \leq A-B$.

The question is whether such a value for t exists. The answer is yes. We know $C-D < A-B$ because $A+D > B+C$. The latter inequality is true because $A+D > 2B$ and $2B > B+C$. So there is some range of t values where (work, shirk) would be an equilibrium, even if the game is only played once.

Next look at my Figure 3. Starting from (work, shirk) with payoffs (D, A), a positive value of the transfer t gives a movement down and to the right along the A+D line. We have to move at least to point M or player 1 will be unwilling to do the work (player 1

②

①

	work	shirk
work	B, B	$D+t, A-t$
shirk	A, \cancel{B}	C, C

Figure 2

Specialized Labor Game ($A+D > 2B > 2C$)

where $t \geq 0$ is a transfer from Player 2 to Player 1
~~paid~~ if ^{and} only if the strategy pair (work, shirk) occurs.

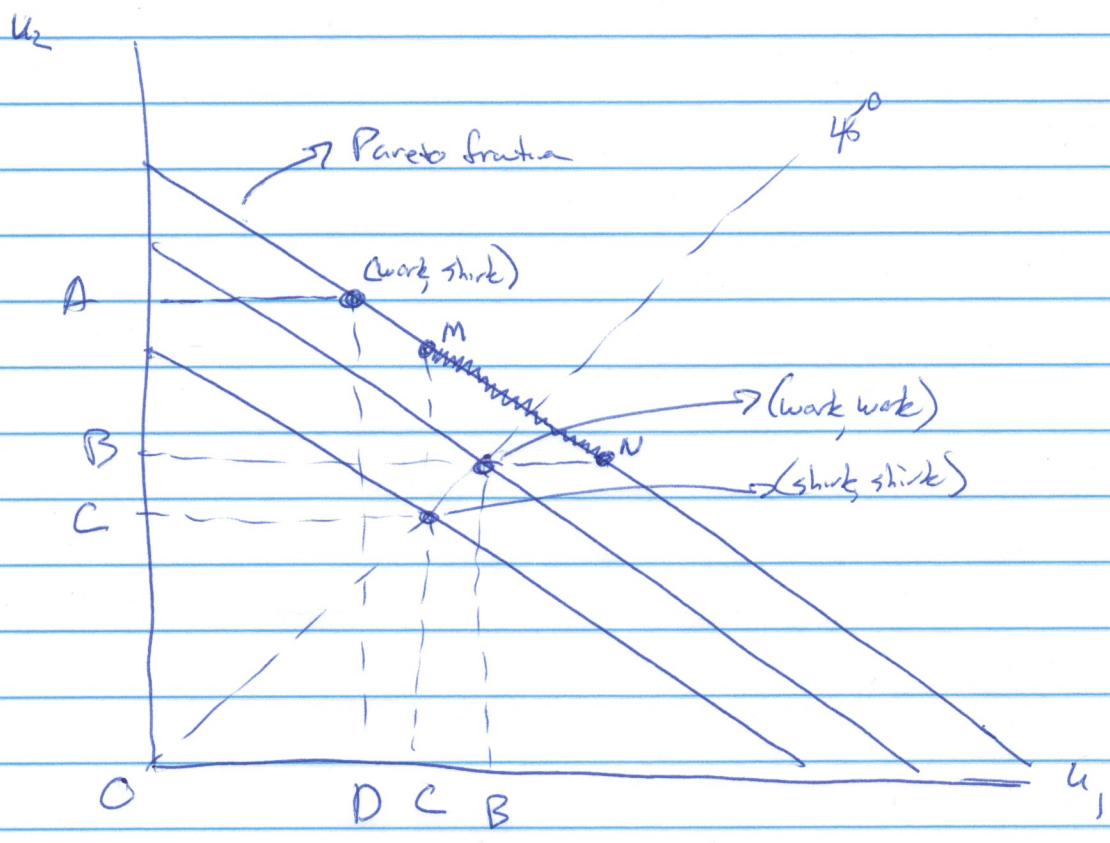


Figure 3

Using a transfer $t \geq 0$ to achieve (work, shirk)
in the Specialized Labor Game.

would prefer to stay at the DSE and get C). This corresponds to the lower bound for t (which is $C-D$). We cannot go any further than point N (where player 2 gets B) without violating the upper bound for t (which is $A-B$). Any value of t satisfying $C-D \leq t \leq A-B$ will put us somewhere along the line segment MN.

Note: RE has a favorite level of the transfer t , which he calls the "liquidated Kantian formula" (see p. 221). This involves $t = B-D$. I don't care about the reasons why RE likes this particular norm, but you can check that it does fall in the range of t values I derived above, so it would put us at a point along the MN line segment. Therefore it would be one way to solve the problem.

Also note: I ignored the payoff E in the preceding discussion because it is irrelevant. We are assuming $D > E$ so $A+D > A+E$. This implies that player 1 should do the work if we want to maximize aggregate welfare. I didn't bother drawing a line for $A+E$ in Figure 3 because this would just complicate the graph without adding anything important.

How does all of this relate to RE's hypothesis about social norms, and to Axelrod's PD tournament? RE thinks a close-knit group can develop social norms that will solve the specialized labor game, but we need something more complex than the TFT strategy.

Suppose we play the specialized labor game many times, and after each round, player 2 has an opportunity to make a gift to player 1. This could be in the form of money, doing a favor, or something similar. If player 1 does the work, player 1 will get the gift, but otherwise not. We know from the analysis above that as long as the transfer t satisfies certain requirements, this will be enough to persuade player 1 to accept (work, shirk).

There is one problem though. Suppose 1 and 2 play the specialized labor game first, and then 2 provides the gift. Why is 2's promise to provide the gift credible? Why wouldn't 2 refuse to provide the gift after 1 does the work? If player 1 is concerned about this, then player 1 won't do the work, and everything falls apart.

The solution is to repeat the game. That way if 2 fails to provide the gift after 1 does the work, player 1 can shirk in the future as a punishment. If player 2 expects this (and cares enough about the future), then player 2 will provide the gift after 1 does the work. In this case, the promise to provide the gift is a credible commitment.

RE has a specific strategy in mind for the repeated game, which he calls "Even Up" (see pp. 225-229 at the end of chapter 12). He gets the idea for this strategy from his study of Shasta County in Part I of the book. It goes like this.

Suppose the members of the community will play many games over time. Some are PD games, some are specialized labor games where player 1 should do the work, some are specialized labor games where player 2 should do the work, and so on. Sometimes the payoffs are big, so it really matters what people do, and other times the payoffs are small so it doesn't matter very much. Between rounds of the game, people can make unilateral side payments or transfers (which RE tends to call 'gifts').

The Even Up strategy says

1. Each player needs to keep track of the size of the benefits people have provided to each other in the past.
2. When you owe another person significant compensation, make a gift to them in a way that gets your relationship back into balance.
3. If your accounts with others are roughly balanced, take whatever action would maximize aggregate welfare in the current period.
4. If anyone else deviates from the behavior that would maximize aggregate welfare, or does not pay a gift to others when one is owed, then go to self-help: defect the next time you play a PD game with that person, shirk when the other person was expecting you to work, and so on.
5. But don't escalate the conflict; try to impose a punishment on the other person that is less than what is owed to you. Attempt to restore the balance in the relationship without overshooting.

Some general comments:

- (a) In RE's view, it is important not to retaliate against a small violation by using a large punishment. This is likely to be destabilizing to the relationship and can lead to long feuds. This is similar to Ostrom's ideas about graduated penalties.
- (b) The side payment idea helps to avoid problems caused by mistakes, for instance where someone did not defect but others think they did. In such situations, a person who is believed to have violated a norm (even if they know they really didn't) can offer gifts to the others and smooth things over. This would not be possible with TFT. If we had two players using TFT and one incorrectly believed that the other defected, we would get a long series of (C, D), (D, C), (C, D), and so on, without ever getting back to (C, C).
- (c) In the real world, reputations are important. If there are more than two people, and someone violates a norm, the victim can tell other people about it. If the violator does not smooth things over with a gift, then everyone in the community could punish the violator. So at least up to a point, larger groups may have the advantage of being able to impose larger punishments. But eventually this could hit diminishing returns due to free rider problems (who will do the punishing?) and unreliable information about violations.
- (d) RE says he doesn't want to talk about distributive norms. But he has no choice. The social norms must single out some point along the Pareto frontier, so there must be some distributive principles built into these norms. How large is the gift for continued cooperation? If it is the minimum amount required, that leads to one distribution of the payoffs, while if it is some amount greater than the minimum that leads to a different

distribution. In fact, RE's "liquidated Kantian" norm mentioned above is a specific way of distributing the sum of the payoffs between the two players.

Closing Remarks on RE's Hypothesis

I am not going to provide detailed notes on chapters 11-16. However, you should read these chapters and there could be exam questions where they would be relevant.

RE's theory does not attempt to explain every kind of social norm, and it would not be fair to criticize it for failing to explain things it was not designed to explain. Examples include norms about social etiquette, gender roles, religion, nationalism, etc.

One restriction on RE's theory is that it only applies to workaday affairs. This includes practical decisions about how to make a living when there may be significant effects on other people (positive or negative externalities).

A second restriction is that his theory only applies to close-knit groups. He is specific that this means groups where informal power is widely distributed, and reliable info is readily available.

There are two other conditions I would want to include in the definition of a close-knit group, although RE does not say these things explicitly. I think his hypothesis will only apply to (i) groups that are relatively small and (ii) groups where there is a low rate of turnover among the individual members (think about the successful cases in chapter 3 of Ostrom). Point (i) is necessary in order to avoid large free rider problems, and point (ii) is necessary in order for people to care about what happens in their future interactions with the other members of the group.

I would also add a third limitation on the theory: it probably only applies in relatively stable environments (physically, socially, and economically). The reason is that good norms take time to evolve, and it will probably be hard to maintain good norms if the environment changes frequently in major ways.

Subject to these qualifications, I think RE's hypothesis is a reasonable starting point for an economic approach to social norms.

One big question: is RE's hypothesis testable? See chapter 15, where he talks about this.

RE would say yes, it is testable, but this is open to debate.

First, it would be necessary to determine whether his hypothesis applies to a particular situation. Therefore, we need to have some criteria for determining whether a specific group in the real world is close knit. How do we measure potential power asymmetries? How do we evaluate whether information circulates easily? We also need to have some criteria for deciding whether the behavior involved is 'workaday'.

Second, we would want to be able to observe norms in some objective way. What kind of evidence would count? What if deterrence is very effective, so the norms are rarely violated? Would we even know what they are in this case? Of course, RE could claim that observing norms is not that hard -- after all, he did it in Shasta County, and he gives a number of other examples in the later chapters of the book (whaling, bees, photocopying, landlord-tenant relations, and so on). I think the main issue here is replication. If other social scientists observed the same situation, would they come to the same conclusions about what the norms are? Or would researchers often disagree with each other?

Third, ideally we would like to observe the costs and benefits of various different norms, so we can see directly whether the norms that exist actually maximize aggregate welfare. The theory assumes that the people in the community know these costs and benefits, but that doesn't mean they would be easy for an outside researcher to observe and quantify.

There are a number of problems along these lines.

- (a) We need to measure net costs and benefits, so we have to measure not just direct costs and benefits but also the transaction costs associated with various alternative norms (costs of negotiation, monitoring, enforcement, etc.). The transaction costs associated with each norm have to be subtracted off when we look at total benefit minus total cost for that particular norm.
- (b) It will often be hard to identify the costs and benefits of social norms. Maybe we can do it in dollars, or time, but there must be some such metric. In many cases, people are trading favors or services that don't have explicit prices (like watching the neighbor's house when they are away). In Shasta, cash transfers would have violated the social norms, so we would not be able to use explicit dollar amounts! There may be multiple dimensions in the payoffs that are hard to compare. If the payoffs can only be expressed as subjective utilities, we are out of luck.
- (c) Even if we can observe the costs and benefits associated with norms that actually exist, how would we know the costs and benefits for norms that don't exist, but could have? The whole point of RE's hypothesis is that there are many possible norms, and we want to explain why we see particular norms in practice. How do we know whether the existing norms maximize aggregate welfare if we can't see the aggregate welfare that would have been obtained from alternative norms?

This may sound like a fatal series of objections, but now let me be a bit more optimistic.

Even if we can't observe costs and benefits directly, we could have a theory about what they are; then the theory will make predictions about the norms that should evolve; and we can look to see whether these predictions are correct.

Think of this in terms of comparative statics, which we discussed earlier in the semester. The social norms are endogenous, and the payoffs in the game are exogenous. Even if

we don't observe these payoffs completely, we could still make some predictions using a comparative static approach.

For example, maybe resource depletion problems tend to result in the normal kind of PD game, while problems involving the supply of public goods tend to result in specialized labor games. RE's hypothesis will predict different social norms in these two situations, and such predictions could be testable.

Economists often make predictions based on comparative static methods: if you change some exogenous (and observable) variable, what does your theory predict about changes in the endogenous (and observable) variables? Economists use this approach all the time even if some elements of the theory (like preferences or utility functions) are not directly observable.

One more issue: suppose we test RE's hypothesis and some prediction of the theory turns out to be incorrect. What do we do?

Someone who is skeptical about RE's approach might treat this as evidence that the entire theoretical framework is wrong (and say "I told you so").

Someone who is more sympathetic to RE's approach might say the hypothesis is still on the right track, but it did not apply to this particular case because the group was not really close knit, or the activities were not workaday affairs, or the researcher was making some incorrect assumptions about the payoffs in the game.

In general, social scientists do not usually toss an entire theoretical framework into the recycling bin just because it failed one test. Instead, researchers want to know whether the theory makes reasonably accurate predictions most of the time, and whether there are ways to patch it up in situations where it doesn't work so well. Of course, this is a matter of personal judgment and people will not always agree about whether a theory is working well enough to justify its continued use.

What we want to avoid are research programs where people look at some more-or-less unique situation involving social norms, and then dream up various arbitrary costs and benefits that rationalize these norms, where the costs and benefits are not observable. In this case, the theory is not testable, and we are back to all the vagueness that RE rightly criticizes in chapter 8 when he talks about functionalism, interest group theories, etc.

I want to close with a discussion of several larger theoretical issues.

1. What do we mean by a norm? Is it something internal that corresponds to a sense of duty, obligation, guilt, conscience, ethics, or morality? (e.g., don't steal, even if you are sure no one is watching and you will never be caught). Or is it something that involves enlightened self-interest (e.g. I obey the norm because if I don't, others will punish me; obeying the norm is part of my strategy in a game theory sense).

RE is ambiguous about this. His case study of Shasta has many statements from people along the lines of "I feel responsible," "good neighbors shouldn't do x", etc. This sounds like internalized morality. Also in chapter 7 he has a lot of discussion of ethics and self-control. But his game theory models rely on enlightened self-interest (along the lines of the rational actor models in chapter 9). In the game theory sections, people refrain from violating norms because they expect to be punished if they do. This is true for the Even Up strategy: what makes people give gifts in the specialized labor game is the threat of self-help by others (retaliation) if they don't.

Probably the best way to think about this is that both things are going on simultaneously. Most people have some degree of conscience. Clearly the human capability for engaging in moral reasoning is genetically hard-wired, even if the details of the substantive norms are not. But conscience operates with varying strength across different people and across different situations. Most people are neither saints nor sociopaths. They have varying capacities to resist temptation. Appeals to conscience or ethics work for some people in some situations, but this needs to be backed up by a system of penalties to guard against opportunists, predators, and con artists, in order to ensure that the conscientious types are not exploited or abused. You will see more about these ideas in the Putterman book.

2. How do we know that Even Up would be evolutionarily successful in the real world, in the same way that TFT was in Axelrod's repeated PD tournament? The answer is we don't know. RE makes a reasonable hypothesis about what might happen in a complex real-world environment. But the arguments based on the success of TFT in a computer tournament are only suggestive, and are not an adequate substitute for serious empirical research. Of course, RE would say that he has already done empirical research (he went to Shasta County), and the results supported his hypothesis. But a more accurate version of the story is that he went to Shasta County first, and afterwards developed a hypothesis to explain what he observed. We need other researchers to look at other cases and see if the hypothesis applies in a more general way.

3. RE is trying to construct a hypothesis based solely on economic efficiency and wants to avoid distributional issues. In principle, he is predicting that we will get to the Pareto frontier, but he is not saying how the gains from cooperation are divided (remember that we had a range of possible transfers that could solve the specialized labor game). This is similar to the Coase Theorem: in the example with the farmer and the rancher, the Coase Theorem predicted that the fence would be built by someone (assuming zero transaction costs), but there was an interval of potential payoff points along a line segment.

As I argued earlier, in practice any pattern of reciprocity or gift exchange involves some distributional outcome. The existing norms take us to a particular point on the frontier, but other norms could have taken us somewhere else. So there is an indeterminacy or incompleteness in RE's theory.

Coase limits the range of potential distributional outcomes by having the legal rules pick a starting point from which bargaining begins. RE could do something similar by starting from a particular game theory equilibrium (such as the dominant strategy equilibrium in a

PD) and then having people split the gains from cooperation equally. However, he does not do anything explicit like this.

4. Suppose RE's hypothesis is right. Should we relax and sleep well, knowing that social norms ensure economic efficiency? I would say no for several reasons.

- (a) The distributional issue has not been resolved, as I discussed above.
- (b) The theory doesn't apply to numerous important issues like basic property rights, or public goods in non-close-knit groups.
- (c) In the real world, there are failures as well as successes (Ostrom). We could have an equilibrium with bad social norms (try to think of some potential examples).
- (d) When the stakes are high, or groups are large, or the group members have a low probability of interacting again in the future, we often need to rely on the law.
- (e) Norms may benefit a particular community or society at the expense of other communities or societies. Such problems tend to arise in the context of resource depletion issues, global environmental issues, or wars among nations. RE doesn't discuss the question of whether larger-scale institutions could evolve to overcome problems of this kind. His theory is limited to institutions at a more local level.

I'll stop there. As you read the rest of the book, you may want to ask yourself whether you find Ellickson's arguments convincing. If you don't, is there some other theory of social norms you would prefer? Why?

What are we doing in this course?

Trying to study economic institutions: what are they, where do they come from?

If we want a theory of institutions, we want institutions to be endogenous, and we want to know how a change in exogenous variables would lead to a change in institutions.

We also want to define institutions in a way that makes them observable, so we can test theories about how they are determined.

I haven't given an "official" definition of institutions yet (will do it for LP book).

But we have seen many examples of institutions in the JE and EO books, and we have seen theories about how certain types of institutions might arise.

For RE, we can think of social norms as a type of institution. He gives some examples of observable social norms, and develops a theory about how they arise.

What is a social norm?

"A social norm is a widespread pattern of behavior in a community, followed by most people most of the time, with penalties for people who deviate from this pattern."

This seems to capture what RE means by the term.

According to this definition, social norms are potentially observable. This gives some hope that we might be able to test RE's hypothesis (see chapter 15; discussed last time).

Today I want to dig a little deeper into the interpretation of what a social norm is, and how norms influence behavior.

Systems of social control

Go back to chapter 7, where RE describes alternative systems of social control:

- | | | | |
|----|-----------------------|-----------------|-------------------------------|
| 1. | First-Party Control: | personal ethics | (conscience) |
| 2. | Second-Party Control: | contracts | (self-enforcing agreements) |
| 3. | Third-Party Control: | norms | (decentralized social forces) |

plus organizations, governments [will ignore these]

He also describes five kinds of rules. I will focus on two:

1. Substantive rules (conduct that is rewarded, punished, or left alone)
2. Remedial rules (what happens when a substantive rule is violated)

Let's ignore the other three, they are not important for what I want to say.

Are social norms maintained through conscience?

Example: norms about honesty.

- (a) You could steal something, where very unlikely to be caught. But you don't.
- (b) You could lie about something, where very unlikely to be caught. But you don't.

In each case there is a conflict between self-interest and conscience. The guilt you would feel from behaving in an immoral way exceeds the benefit from violating the norm.

Real world examples: people often return missing wallets, even though they don't have to and there is some cost of doing it. But many people think this is "good behavior".

There are big economic benefits when norms of honesty are widespread. If you can trust other people, transaction costs are much lower. You can make a verbal agreement, shake hands, no contracts, fewer lawyers, fewer lawsuits, less monitoring and enforcement.

There is some evidence from lab experiments that if you allow people to talk before they play a prisoner's dilemma game, there is more cooperation than if they can't talk first. A hard-core game theorist would predict that this makes no difference (each player still has a dominant strategy). But some people don't like to break promises, and letting them talk first is like giving them an opportunity to create a contract.

How far can we push this?

Suppose parents taught their kids from a very young age to recognize prisoner's dilemma situations, and taught them that good people always cooperate. Would this eliminate all problems with prisoner's dilemma games?

If parental teaching caused everyone to cooperate, a game theorist might say that we have changed the payoffs from the game. If we just look at material self-interest it is a PD, but if we include guilt from defection, cooperation is a dominant strategy, so it is not a PD.

Anyway, this doesn't seem practical in the real world. People don't learn about PD games at the age of 6, they learn about them in universities.

Furthermore, the evidence from EO about tragedy of the commons suggests that often we need complex institutions with monitoring and penalties; conscience alone is not enough.

Are social norms maintained through enlightened self-interest?

Economists and game theorists often assume people behave in self-interested ways. Can social norms still work when this is true?

RE talks about "rational actor" theories, and he includes self-interest in the definition.

Also: RE talks about remedial norms. Wouldn't need them if people always followed the substantive norms. This suggests penalties play a significant role (not just conscience).

Let's go back to "Second-Party Control" in Chapter 7.

RE thinks about this as involving contracts (he is a law professor). But if we really want to think about "order without law", we should ignore lawyers and courts.

Instead, let's think about "second-party control" as similar to "self-enforcing agreements" in the EO book.

Imagine two people playing an infinitely repeated PD game. They can have a discussion at the beginning and agree to use trigger strategies as a way of ensuring cooperation.

We know this will work if they each put enough weight on future payoffs.

I would call this "second-party control", and it is a possible way to maintain social norms.

The substantive norm: cooperate if no one has ever defected in the past.

The remedial norm: defect if anyone has ever defected in the past.

This is an example of enlightened self-interest. We are not relying on conscience to bring about cooperation. Instead, we assume each person understands what is in their own self-interest, given the behavior of the other person.

What about third-party control?

Can we extend this to an entire community with hundreds or thousands of people?

Maybe. But now we might need third parties (people not directly involved in the original dispute) to bear some costs of enforcing a norm.

For example: in Shasta, self-help might involve gossiping to friends about someone who violated a norm. But then the friends have to impose some kind of penalty on the person who broke the rules (maybe not do business with them anymore). This could be costly.

Are third parties following their conscience, obeying enlightened self-interest, or what?

Not clear. Maybe if there is gossip saying that someone is unreliable, it is in your own self-interest not to do business with them. On the other hand, maybe you incur a cost by punishing someone even though there is no personal benefit, because your conscience says it is important to uphold the community norms (the norms are like a public good).

In the LP book, we will see lab experiments where people often punish others for not maximizing aggregate welfare, even if this is costly and the punisher gets no personal benefit by imposing a punishment.

Closing thoughts on Ellickson

1. Most of the time, RE seems to be talking about enlightened self-interest. That is the basis for his argument about Axelrod's PD tournament, TFT, etc. His claim is that TFT provides high payoffs, people tend to imitate successful strategies and so on. His hypothesis does not rely on conscience or morality as the basis for norms.
2. But it is an incomplete theory of self-interest. He doesn't talk about how large are the temptations to defect, how large are the collective gains from cooperation, or how much weight people put on future payoffs. We found with trigger strategies that these factors are highly relevant to the question of whether cooperation is an equilibrium outcome. You might think they would play a role in his hypothesis about norms, but they don't.
3. He is also vague about how we get from 2-person repeated games, where the role of self-interest is easy to study, to decentralized social forces involving hundreds of people in a community. In this case, we might need to go back to conscience or morality to explain why third parties follow remedial norms.
4. In particular, he relies on "Even-Up" as an extension of TFT. Is Even-Up purely about enlightened self-interest (people follow it because it is not in their interest to deviate), or does it also require widespread community beliefs about how people ought to behave, and reliance on morality to punish deviations from norms?
5. My view is that in the real world, both factors are important in maintaining social norms. If only conscience was important, we wouldn't need remedial norms at all. But if only enlightened self-interest was important, how would we explain why third parties punish violations of norms, even if they get no benefit from doing so?
6. People are heterogeneous (an idea from LP). For some people, appeal to morality will work, and we don't need threats of punishment (as long as temptations are not too large). For other people morality is a weaker factor and it is necessary to have systems of monitoring and punishment. The balance depends on the situation.

We will see more about all of this when LP discusses human nature.