[SQUEAKING]

[RUSTLING]

[CLICKING]

IAN BALL: So today, I want to say just a little bit more about revenue equivalence, or the revenue formula that we wrote down. And I think this will be useful for the problem set and the quiz on Friday. And then I'm going to talk about an application of auctions to display advertising auctions, like the kinds of auctions-- the keyword auctions that Google does commonly today.

Just to remind you of this revenue formula that we computed, we were looking at a single item auction. And we had this formula-- maybe I'll write it down here-- where we said Ti hat of Vi, which was the equilibrium expected payment by a bidder with valuation Vi, was equal to this. So we said the expected payment of a bidder with valuation Vi in this equilibrium is equal to the expected payment of bidder i when she has the lowest possible valuation plus this integral of maybe this mysterious term of x times the derivative of this function, q hat-- where, remember, this is the derivative of the allocation probability.

So I just want to highlight a few assumptions that were required to get this result. And the problem set is going to explore some ways this assumption, this result can break. Yes.

AUDIENCE: Is there a reason why the lower bound is on the integral?

IAN BALL: Ah, yes. That should be V lower bar, sorry. Yeah, thanks. So a few assumptions we made-- one was the assumption of independent private values. Statistical independence, it says one bidder's valuation is statistically independent of another bidder's valuation.

Private values-- private is a confusing term. I think maybe known values is maybe a more descriptive term. It means that each bidder knows exactly how much they value the good. Information that another bidder has is not relevant to my valuation of the item. And I still think that's a bit abstract, so I'm going to go over an example where this fails in a second, and I think that should clarify the role of this assumption.

The other assumption we made-- notice we took an integral here over some range of values. So another important assumption is that the values-- maybe I'll say live in an interval. So if instead a player's valuation could only be, say, $1 or $2-- it's not an interval, it's just 2 points. Then we wouldn't be able to take an integral between $1 and $2, and this formula would no longer hold. And the problem set will show you an example of this.

One other maybe more technical assumption you might notice is that, well, for this to make sense, this function, q hat, has to be well behaved. We have to be able to take the derivative of Q hat. That's a more technical thing, so I'll just say technical Qi hat is differentiable.

I call this technical because it turns out if you use a more involved proof, you can get a substantially similar result without this assumption. So that's not something we're going to cover in the course, but I'll just highlight that our proof used differentiability of q hat, but that's not really necessary for the result. But these assumptions really are substantive and important for this revenue formula.

So to understand what private values mean, let's look at a classic example without private values, which is called a wallet auction. I don't think anyone actually does auctions like this, but it's a good illustration. So I think there's a few stories for this. The basic story is that I guess this happened at a bar. You have two people at a bar. They each get out their wallet, and they put their wallet in a bowl, say, or in a pile on the table.

So you have player one's wallet and player two's wallet. And then the players bid for the union of the two wallets. So we have both our wallets together in a bowl. We're now each going to submit a bid. Whoever bids more gets the bowl, but has to pay that amount.

Now, what's special about this auction? The story maybe is far-fetched, but the important thing is I know as player one what's in my wallet, but I don't know what's in the other player's wallet. The other player knows what's in their wallet, but they don't know what's in my wallet. So maybe we'll call the amount in my wallet theta 1 and the amount in the other person's wallet theta 2.

And just to make it simple, let's say theta 1 and theta 2-- just to clarify exactly which assumption we're relaxing, let's say theta 1 and theta 2 are independent and uniform. Independent and uniform over 0, 1. That's just a normalization. One could be $1 billion or something. But the key point is that their 1 knows theta 1 but not theta 2 because they don't know what's in the other person's wallet.

So if I'm bidder one, I say, I know my wallet, say, has 0.3. But all for all I know, the other person's wallet is uniformly distributed between 0 and 1. That's all I know because I can't see what's in their wallet. And symmetrically, bidder two knows theta 2, but not theta 1.

Now, why is this not a setting of private values? Well, I do know something. I know theta 1. I know it's in my wallet. But I don't know the joint value of our two wallets. And more specifically, something the other player knows is relevant for my valuation of the bowl, of the item. So what makes it not private values? So maybe I'll say here, it's a failure of private values.

Why? Because my willingness to pay, my valuation depends on information that another player has-- or maybe more precisely, another player and other players information influences or affects my valuation for the item, where here, the item is the combination of the two wallets. Maybe I'll say my willingness to pay.

If the other player's wallet contains $1, which only the other player knows, I'd be willing to pay more for this bowl that contains both of our wallets. Now, you might say, well, OK. I don't know exactly what's in the other wallet, but I can compute expectations. So the naive approach might be, well, let's try to treat this like a normal private values auction.

So here I'll write naive approach to say this is going to be wrong. But let's see what happens. To the naive approach, if I'm player one, what would I say is my willingness to pay for the bowl? If I just looked at it, I know theta 1. What's a reasonable way to say, what is my willingness to pay? What is my expected value from the union of the two wallets or the sum of the two wallets? Yeah.

AUDIENCE: Since the other player's value is uniformly distributed between 0 and 1, the expected value of the other player's value is 0.5. So the naive approach would be to add 0.5 to [INAUDIBLE].

IAN BALL: Exactly right. So you might say, OK, this is just a normal auction where player one's value is theta 1 plus 0.5 or plus 1/2. And player two's value is theta 2 plus 1/2. So I might say, look, we have some item, and I have some private information. I know my theta 1. I use that to compute my valuation. And let's just go ahead like this.

And what I want to show is that this auction is going to behave very differently from a private auction, where this is my value and this is the other player's value. So to see this, let's look at-- so fully solving this auction is quite involved, and I'll leave that to the notes. And we won't really cover that in this course. We will cover it in 1416 in the spring, if you're interested in that.

But let's just imagine we did a second price auction. And let's imagine that we looked for an equilibrium where each player bids truthfully. So let's look at a strategy profile where player one bids-- we could say b1 of theta 1 plus theta 1 plus 1/2 would be 2 of theta 2 equals theta 2, et cetera.

Now, if this were a private values auction, then this strategy profile would be in equilibrium. We already argued that in a classical private value single item auction, it was an equilibrium for each player to bid truthfully, to bid their valuation for the item. And in fact, we argued more. We said however the other player bids, it's optimal for me to bid my valuation in a second price auction.

So let's see what goes wrong in this case. So let's suppose that to make things simple, let's say theta 1 is 0.4. And theta 2 equals 0.3. So if that happens, then player 1 is supposed to bid 0.9. And player 2 is supposed to bid 0.80-- 0.5 plus 0.3. So this would be player two's bid. And this would be player one's bid.

Now, how would our reasoning for why bidding b1 is optimal go? Well, if this were private values, we would say, OK, I'm bidder one. My valuation is 0.9. If I bid truthfully, what happens? Well, I get the item and I pay 0.8. And 0.8 is less than 0.9, so I'm doing well.

And if I deviated, well, if I deviate to bid any amount more than 0.8, it's not going to make a difference. I'm still going to win. And I'm still going to play 0.8. And if I deviate to any amount less than 0.8, well, I'm going to lose, and I'm going to get a payoff of 0.

So the naive argument might be, well, I'm bidding optimally. Why? Because 0.1 is greater than 0. And maybe I should explain where this comes from-- 0.9 minus 0.8. So it looks like I'm doing well. I get a payoff of 0.1. I could do no better than 0.1. I'm bidding optimally. That would be the private values intuition. What's wrong with this? So everything I said was wrong. Yeah.

AUDIENCE: Your actual value is negative 0.1 because the value is 0.7. When you're bidding 0.8, you're paying 0.9.

IAN BALL: Exactly right. And this is getting at the key point that it's not really the case that my value is known to be 0.9. That was just my expected value without getting any additional information about the other player. The problem is if the other player bids 0.8, and the reason they bid 0.8 is that the value of their wallet is 0.3, that changes the amount that I value the wallet.

I no longer value it at 0.9. What do I value it at? Well, the true value of the wallet to me is the sum of theta 1 plus theta 2, which in this case is going to be 0.7. So if I bid truthfully here, I'm going to pay 0.8 for something that's only worth 0.7 to me.

What changed? And why is it different from the classical auction? It's that, well, this 0.9 wasn't my true valuation of the item. It was my belief or my expected value of the item. And unfortunately, my assessment of the value of the item is influenced by the other player's private information.

So the issue is, when I won, I'm winning precisely because the other player's wallet doesn't have much money in it. And this phenomenon, which this is a very simple example. Maybe you think, oh, this doesn't have relevance to the real world, but this is actually a really serious problem in a lot of real auctions, and it's something called the winner's curse.

So it sounds good for player one here. Player one won the auction. The problem is we have to think, why did player one win? Well, when will player one win this auction? Yeah, go ahead.

AUDIENCE: [INAUDIBLE]

IAN BALL: Yeah, they'll win the auction when the other person bids a low amount, which is precisely when the other person's valuation is low. And the other person's wallet doesn't have very much. So the winner's curse is that you tend to win when the item is less valuable.

And if you don't take this into account, people will tend to overbid. And the person who wins the auction will tend to come away a loser because they'll overpay for an item that's not worth very much. Now, just a few examples where this has happened in practice. And firms in the '70s and '80s lost tons of money making mistakes like this.

One example would be oil drilling. Could someone maybe walk me through how oil drilling has anything to do with this wallet auction? And why firms might be subject to the winner's curse? So here we're drilling for oil, and we're bidding for attractive land.

So oil drilling-- and let's say an auction for land. So it's for land rights, I'll say. So there's some big plot of land, say, in Alaska, in the Arctic. We are bidding for the right to drill for oil in that land and with different oil firms. Yeah.

AUDIENCE: Maybe a lot of firms ended up paying more than they would actually get out.

IAN BALL: So that's exactly what happened. And why did that happen?

AUDIENCE: I guess they thought it was worth more. They got in the auction, and were just trying to pay the most that person was selling it at.

IAN BALL: So I think we're along the right track. Any other thoughts about maybe a little more specificity. Yeah, in blue. Yeah.

AUDIENCE: The person who's going to win is going to be the one who thinks there's the most oil there. And assuming there's a range that's of people's beliefs, and it falls somewhere in the middle, they're going to have overvalued it.

IAN BALL: Exactly. So the crucial thing here is, what are we doing when we're a firm bidding on this? We're going to make a noisy estimate of the amount of oil in the land. So the issue here is noisy or imprecise estimates. So I'm going to do a little testing right. I'm going to try to guess or estimate how much oil is in this land. Every other firm is going to do the same thing.

If I win, what does that mean about my estimate relative to everyone else's? It means my estimate was higher than everyone else's. And if my estimate was higher than everyone else's, it probably means my estimate just was an overestimate of the true value of the land, and I'm going to end up paying too much. Another example where this happens is auctions for, say, government procurement.

So this is the analog of this. What happens there? The government says, we want to build a subway. And now the firms are going to bid on the amount that they need to get paid to build the subway. So I might say, I'll build it for $100 million. Someone else has outbid it for $110 million, someone else has outbid it for $90 million. And the government chooses the lowest bidder. So these, you always have to be careful here. So the winner is the lowest bidder because the government wants it built as cheaply as possible.

Now, I think something else-- there can be a similar winner's curse here. What's the story in a government procurement auction? Say we're building a subway in New York. Yeah.

AUDIENCE: To build a subway, there's unknown costs. You're not fully certain. And so, again, you're making a noisy estimate of exactly how much it's going to cost to make the subway. And then whoever wins is going to be whoever had the lowest estimate of cost, which is probably wrong.

IAN BALL: Exactly right. So it's exactly parallel to this. We're making noisy estimates. Here we're estimating the value of something. If I win, I probably overestimated the value. Here we're estimating the cost of doing something. If I won, I probably underestimated the cost. So I thought it would only cost me $100 million, then I build it, and it ends up costing a lot more. So it's the same basic idea that we have imprecise estimates that are all correlated.

There's actually a really nice story. So Paul Milgrom, who did some famous work on this and ultimately won the Nobel Prize, said his dad was a painter and would always complain to him that when he would bid on painting contracts, he'd always find that when he won, it cost more to paint the house than he expected. It was always a bit of a harder job than he thought.

And Paul Milgrom said, oh, one day I wanted to tell him the answer. It turned out his father died before he could tell him, so it's a sad story. But I think it's a nice illustration of this phenomenon. Any questions on that? And then we'll move to another application. So now let's move to big application of auctions, which is going to be advertising auctions. And by advertising, what I really mean is keyword search auctions.

And you probably see this all the time. If you Google something, you get the organic results. So let's just see what it looks like on a page if I go to Google. Well, I'm going to get some organic search results. What do I mean by organic? Organic just means Google's algorithm is saying, these are the websites that are most relevant for what I typed in.

But then you're also going to have sponsored results. You might have something here-- I'll say S for sponsored one. Maybe something here, S2. Maybe you'll have a few over here, S3, S4. And these slots-- if you look, they'll be relevant to what you searched.

And you may not even notice-- I think you probably know, but a lot of people just think of all the results as the same. They don't quite recognize what's sponsored and what's not. What does it mean if it's sponsored? Do people know what's going on there? Once they've seen organic and sponsored? Yeah.

AUDIENCE: I don't know exactly, but [INAUDIBLE].

IAN BALL: They paid to have it there. So if you notice-- let's say you Google Uber, just as one example. What might you see in the first slot if you Google Uber? Yeah, you often see Lyft. And what's happening is Uber will probably show up as the first organic result. But the first sponsored result might actually be Lyft, because Lyft recognizes that someone who's searching for Uber is looking to rideshare. And maybe Lyft can undermine them and get them to buy that.

Now, Uber might actually buy a slot here as well to try to do that. And now we might see-- so Uber might actually appear both in your organic results and in the sponsored results. And there's a lot of strategy that goes on into this. And this is actually how Google makes a lot of their money-- that by having the best search engine, they have a lot of people searching. And firms are willing to pay a lot for these slots.

Now, a big insight, a big advance for Google was the fact that these slots are responsive to the keyword. So I'll say the sponsored slots depend on the keyword, the keyword meaning what you Googled. Before this, Google would just do banner advertisements, where they would just have something on the screen that was shown to everyone. But that's not very valuable because I might be showing an ad to someone who doesn't really care about it. These sponsored searches allow Google to hyper target their ads, and therefore, generate a lot more revenue from their ads.

Now, how does Google decide who gets to be in this slot? Who gets to be in this slot? Who gets to be in this slot? They run an auction among the advertisers. And this is actually, maybe in numerical terms, the most important auctions we see in the world. Hundreds of billions are raised every year doing this. And today we're going to analyze a very simple model of this sponsored search auction.

So just to be clear here, the entire auction is designed for a particular keyword. If people Google Lyft, then there's going to be an auction for showing slots to someone who googled Lyft. If someone googled Uber, there would be a different thing. If they googled hotel, there'd be something else. So everything we're going to be talking about is for a fixed keyword.

Now, why can't we just off-the-shelf apply what we've done so far? The challenge this is not a single item auction. We have many items. We have the first slot, we have the second slot, we have the third slot. And the other thing is these items are not the same. So it's not a single item auction. We have multiple heterogeneous items. Or maybe I'll say differentiated items.

Getting shown first is a lot better than being shown second, which is better than being shown third. So how do we model this? Well, we're going to write down a pretty simple model. And here's our general model. We're going to say that if your ad shows up in the first slot. It's more likely to be clicked.

So in our general model, we have n bidders. I'll call them i equals 1 to n. And then we're going to have m advertising slots. So in my example up here, we had four. In general, we'll have m advertising slots. We'll call these j equals 1 to m. And what distinguishes these slots is what we'll call the click-through rate.

So we'll say the Click-Through Rate, sometimes called CTR, are alpha 1, alpha 2, alpha m. So what does this mean? Alpha 1 is the click-through rate of the first position. If someone shows their ad in the first position, this is the probability that it's going to get clicked. If someone shows their ad in the second position, then the probability that it gets clicked is alpha 2 all the way down to alpha m if it's shown in the nth position.

Now, there's a big simplifying assumption here that the probability that it's clicked doesn't depend on which ad is shown there. If I showed a really bad ad for a really undesirable product, even if it's first, probably no one's going to click on it. But if we're looking at products that are pretty similar, like Uber and Lyft, maybe this is a reasonable assumption.

We're next going to assume that each of the bidders-- we have to specify, well, what do the bidders know? Well, we're going to assume the bidders have valuations per click. So the bidders-- so let's say bidder i has valuation Vi per click. The idea would be if someone clicks on my ad, well, maybe they'll buy it, maybe they won't. But the expected value to me of someone clicking on my ad is Vi.

Another simplifying assumption here is that we're saying if someone clicks on it when it's in the first position versus someone clicks on it when it's in the second position, that's equally valuable to me. That may not be a good assumption. Maybe if someone clicks on it even when it's in a bad position, maybe it's more valuable because those people really like it, but we're not going to focus on that. We're going to keep it simple.

So now I'm Google. I have these m slots. I have these i bidders. I'm going to ask them to submit bids. And I need to then choose-- so what do I need to choose? Let's go back over here.

So Google has-- maybe I'll call this auction design choices. So one first choice is, well, I'm going to ask people to pay me. But am I going to ask them to pay me every time someone clicks on the ad? Or am I going to ask them to pay me just every time the ad is shown? And these are two common models for these auctions.

One is called pay-per-click. And one is called pay-per-impression. An impression means the ad is shown to someone, but not necessarily clicked on-- may or may not be clicked on. In our model, this actually won't make a difference because we're assuming that the clickthrough rates are perfectly known. In a more realistic model, people might be uncertain about these clickthrough rates, and then it does make a big difference whether it's pay-per-click or pay-per-impression.

I might want pay-per-click if I'm really uncertain about how popular my ad is, because then if my ad turns out to be a failure and no one clicks on it, well, then at least I don't have to pay very much. Whereas if it's pay-for-impression, well, then I'm on the hook to keep showing my ad. No one's clicking on it, I'm still paying for it. So you can see why it's pretty common to use pay-per-click. You can think of this as ensuring advertisers against having a really bad ad that no one wants to click on. But in this model, that won't make a difference. Yes, Amy.

AUDIENCE: But then does it close out on pay-per-click if all the ads are pretty unpopular?

IAN BALL: Yeah. I guess the firms would still have an incentive to have good ads. Yeah, I guess you might worry a little bit, maybe about moral hazard. If the pay-per-click was too high, then maybe firms wouldn't have an incentive to improve their ads. And I guess there is a bit of an incentive trade-off there, but I think the firms still have a-- Google can still tell the firms, no one's clicking on your ad, and the firms still want their ads to be clicked on. But I agree, their incentives may not be quite as strong in that case, yeah.

So there's a trade-off there between maybe insurance and incentives, but yeah. And over time, I think the firms learned a lot about how often the likelihood of their ad being clicked on. So today, we're going to use pay-per-click. Again, in our model, it won't matter. But this will just simplify things, right?

So what does the auction need to say? Well, the auction is going to collect the bids. And the bids are going to be payments per click. So my bid is going to say, think of it roughly as how much am I willing to pay each time my ad is clicked on? So we're going to collect these bids, b1 through bn.

And then what do we need to decide? Well, we need to have an allocation. So we have slot 1, slot 2, up to slot m. For each slot, we have to say who gets it? So which ad is shown in this shot? And then the payment per click.

So this is formally what we're choosing. We're choosing a function that says, if these are the n bids that the advertisers submit, who are we going to put in the first slot? And how much are we going to ask them to pay per click? Who are we going to put in the second slot? And how much are we going to ask them to pay per click? All the way down to the m slot.

So that's the choice we have to make. And of course, the design of the auction that we use is going to affect the bidding behavior of the advertisers. We could use this various auction format. So today, the first auction format we're going to look at is what's called a GSP, or generalized second price auction.

And I think we can maybe guess how this works. You know how a second price auction works? Why do we call it generalized? Well, the classic second price auction applies when there's a single item that we're auctioning off. Here, we have multiple items, so have to generalize it.

And I think this is the most intuitive generalization. Any ideas about if you had to think how-- let's say you're working at Google. Someone said, let's generalize the second price auction. What might you do here? Well, let's first think about the allocation. I think that's pretty easy. Who do you think you're going to show in the first slot? Who's that?

AUDIENCE: The highest bidder.

IAN BALL: The highest bidder, right? So the allocation rule makes sense. We'll show the highest bidder here, the second highest bidder here, all the way down to the mth highest bidder here. So in general, maybe I'll put it here.

We're in slot k. So slot k is going to go to the kth highest bidder, on the m. But now maybe the harder thing is, what's the payment per click? What's a reasonable amount to charge the kth highest bidder per click?

AUDIENCE: k plus first.

IAN BALL: Right. That's a pretty natural way of doing it. We say, OK, you're the kth highest bidder, so you get the k slot. But we're not going to charge you your bid. We're going to charge you the next highest bid. In a second price auction, you can think of the second price as the next highest bid after the winner. And here, we're charging the next highest bid after the person who wins the k slot.

So the payment-- maybe I'll call it pk-- is equal to k plus first. Highest bid. This seems pretty intuitive. And let's try to get a feel for how this works in the example.

AUDIENCE: What would happen in an event the last person?

IAN BALL: Good point. So there's a question of what's the relationship between m and n? Usually the last person would pay-- so m may be smaller than n, but if m is equal to n, the next highest bid, we would usually call 0. So we can imagine we have the n bids. And then imagine the n plus first bid is just 0.

AUDIENCE: So first of all, they--

IAN BALL: They wouldn't pay anything. Now in reality, m is probably much smaller than n. So if m is smaller than n, then the mth highest bidder is going to pay the m plus first highest bid, and then m plus 2, m plus 3, m plus 4, they don't get anything at all.

So what you're describing could only happen if there's more bidders than ad slots. And that makes sense if there's more ad slots than bidders. If there's as many or more ad slots than bidders, then we can't really charge much because we can just give them away. But good question.

OK. So let's go through a simple example to make sure we understand. So let's do an example with m equals maybe 2 and n equals 3. So we have three bidders and two ad slots. And let's say the ad slots-- we have ad slot 1, which has clickthrough rate alpha 1. And we have ad slot 2, which has clickthrough rate alpha 2. And we're always assuming that they're in descending order, so alpha 1 is strictly greater than alpha 2.

And let's say the bidders submit these bids. Maybe b1 is $7, b2 is $8, and bid 3 is $5. So now what's going to happen in this auction? Well, the first slot is going to go to the bidder who bids the most. In this case, that's bidder 2. So bidder 2's add is going to be-- they're going to show bidder 2's add.

And in the second slot, they're going to show the second highest bidder, which in this case is bidder 1. So they're going to show bidder 1's add. And what is the payment per click going to be? So maybe I'll call this p1 and p2. So I'll try to use superscripts for positions and subscripts for bidders.

So p1 is not necessarily bidder 1. It's the payment made per click for the first slot. So in this case, the highest bidder bid 8. The second highest bid was 7, so p1 is going to be 7. So bidder 2 is going to have to pay $7 per click.

And now if we go to 2, well, the second highest bidder is going to pay the third highest bid, which in this case is 5. So p2 is going to be 5. This all sounds good. But now we might ask, well, is it optimal to bid truthfully?

We know in a standard second price auction, it is. You should bid truthfully. But now things are a little more complicated. So let's take the perspective of bidder 2. Well, what is their expected profit going to be from bidding truthfully? That is bidding 8 in this auction.

I guess I haven't been clear. Let's assume that b2 equals 8. So let's suppose the true value is 8, and they bid 8. What are they going to get in expectation? Well, let's go through it. They're being shown in slot 1. The probability their ad gets clicked is alpha 1.

And each time it's clicked, their valuation per click is 8, but they end up having to pay 7. So we get alpha 1 times 8 minus 7, which is just alpha 1. But let's suppose that instead they bid a bit less. And let's suppose that instead of bidding 8, they bid 6. So this is a bid of 8, but let's suppose instead they bid 6.

Well, now, what's going to happen? Now what's their expected payment? Well, if they bid 6, where is their ad going to be shown? Yeah. It will be the second slot. If they bid 6, bidder 1 is going to get the first slot. Bidder 2 is now going to get the second slot, so they're going to get alpha 2.

Their value is still 8, but now how much do they pay? Yeah. They're going to pay 5, because now they're getting the second slot. They're going to pay the third highest bid. The highest bid is 7. The second highest bid is 6. The third highest bid is 5. So we get 3 alpha 2.

Well, what's better? Well, it's going to depend on the relative values of alpha 1 and alpha 2. And in particular, if alpha 2 is pretty close to alpha 1, they're going to be much better off shading their bit. So this is going to be preferred if alpha 2 is close to alpha 1. It's smaller than alpha 1, but it's close.

And here, we see a fundamental difference between a standard second price auction and the generalized second price auction. In the standard second price auction, you want to bid truthfully, because if you don't get the item, you don't get anything. But in a multi-item auction, now it's less clear. If you shade your bid instead of getting nothing, you get the next slot.

And if the next slot is almost as good as the first slot, you might as well shade your bid. You might as well reduce your bid. So we now have a much stronger temptation to reduce your bid when the penalty for losing is lower. Instead of getting nothing, I just go one slot down. And in particular, when these slots are very close, going one slot down is not very costly.

So now we need to ask, well, we could use this bidding format, but then bidders have to spend a lot of money trying to figure out how much other people are going to bid. And their bid is going to depend on other people's bids, and it's going to be complicated. So maybe we want an auction where it's optimal for people to just bid truthfully. So we want a different generalization of the second price auction than what's called a generalized second price auction. Yes.

AUDIENCE: This jsp, does this decrease revenue for the auctioneer because everyone's going to shade?

IAN BALL: It's a subtle question. So there is an equilibrium of this auction that will give-- I guess you have to say, what is the comparison with? The auction format that we're going to describe next, the VCG auction, there will be an equilibrium of this auction that will give the same revenue in this. But that's in this very stylized example where there's no uncertainty. If you added some more realistic features, then you have to be a little more careful, and it's actually not so clear.

AUDIENCE: So the auctioneer doesn't have a preference between the style of the second price auction?

IAN BALL: So I think the question is-- some of the theorems say it doesn't matter. But those theorems make some very strong assumptions that people bid exactly according to the equilibrium. And we know in practice that doesn't really happen. So in fact, one concern there was, going back historically, is people were using basically a version of a first price auction of this.

And there was a suggestion that we should move to a second price auction. But the concern was, well, if we jump right away to a second price auction and firms don't adjust their bidding behavior, now we're going to lose a lot of money. So the second price auction should give you as much revenue if we play the new equilibrium where people bid much differently. But if firms don't know that, and they don't react quickly enough, then there's really a strong disincentive to move to a second price auction. I think that's one issue that went on.

So there's a gap between the theory and the applications here. I guess I should say maybe the fundamental issue that motivated a lot of this auction work was that firms are able to modify their bids quickly in real time. So the firm submits some bids, their ad is shown, and then they can change their bids over time. And what was happening is firms were constantly changing their bids, and these bids were going in cycles.

So we're modeling this as a one-off interaction. If we use the auction format we're going to describe, I think that's a good model. But if people really care about how other people bid, then it can make things a lot more complicated when people can change their bids over time. Good question. So let's think about how else we can generalize the second price auction. And I think the key is to think about the second price auction a bit differently.

So I think the most obvious way to generalize the second price auction is to say, well, what is a second price auction? You pay the second highest bid, so let's generalize it. You pay the next highest bid. But let's think about it. What's really going on in the second price auction a bit differently?

Another way of thinking about a second price auction is that each bidder is asked to pay their externality. Let's understand what we mean by this. So let's say the valuations are ordered. It's v2 to vn.

And let's say I'm this guy. And I win. So I'm the highest bidder. Let's suppose everyone's bidding truthfully, I guess, to make this argument. I win. I get my valuation of v1, and these guys all lose. Now, what if I weren't here and we just wanted to allocate the good in a way that was efficient for everyone else?

What would we do if bidder 1 were not here? Who would we allocate it to? Bidder 2. And what would they get? Their evaluation would be v2. So we can say if bidder 1 were gone, then the total value that all the other bidders would get would be v2. If bidder 1 were gone, the item would go to bidder 2. And the value would be v2.

So we can think of it as, well, the total value that everyone other than bidder 1 gets is 0 when bidder 1 is here and would have been 2 if bidder 1 were not present. So the externality that bidder 1's presence or bitter 1's bid imposes on everyone else is exactly v2, because they take the item away from everyone else. And the most efficient allocation of it to everyone else would have been to bidder v2.

So this is the aspect of the second price auction that we're next going to try to generalize. We're going to try to charge people the externality that their presence imposes on everyone else. In this case, that externality happens to be the second highest valuation, the second highest bid. But in general, it's not necessarily. And that's what we'll see with the next auction format. So let's look at this.

So let's say that we bid b1, b2, b3. And let's suppose that these are ordered like this. And we have these slots-- 1, 2, 3. Now, if I'm the highest bidder, then I get slot 1. The second highest bidder gets slot 2 and the third highest bidder gets slot 3.

So here, I'm just assuming bidder 1 is the highest bidder, 2 is the second highest, and bidder 3 is next. Well, what would happen if bidder 1 were not present? If bidder 1 weren't here-- or in other words, if bidder 1 just bid 0 and didn't win anything, what would happen to bidder 2?

Which slot would they get? They'd move up to the first slot. So they'd go up here. And bidder 3, similarly, would move up to the second slot. So let's now compute bidder 1's-- the highest bidder's externality. Well, bidder 2-- the externality is that they're pushed down from alpha 1 to alpha 2 as their clickthrough rate. So we have b2.

So what happens? Bidder 2 gets alpha 2 instead of alpha 1 as their clickthrough rate. And their value per click-- let's treat it as their bid, b2. This is the externality imposed on bidder 2, and this is the externality imposed on bidder 3.

Notice here, we have to be a little careful. To directly compute the externality, we'd have to know how much other people actually value the item. So when we compute the externality, we compute it as if people are bidding truthfully. So I don't actually know how much bidder 2 values the slot, but we're going to compute the externality as if they were bidding truthfully.

So if bidder 2 were bidding truthfully, then the amount they value the slot per click is exactly b2, and their loss is the difference in clickthrough rate times the valuation per click. And similarly, b3-- their loss is the difference in clickthrough rate times their valuation here. So in general, the idea or the goal is to charge each bidder per externality. Motivated by this simple example, let's see if we can give a general formula for this auction format.

And this is called a VCG auction, and this is named after three people. Vickrey won the Nobel Prize for this-- Clark and Groves. And we can think of this as a different or maybe the true generalization of the second price auction.

So now we get some bids, b1 through bn. And we have to decide what to do with them. It's helpful to have some notation for what's the highest bid, what's the second highest bid, and so on. It's not necessarily the first bidder is the highest.

So just some notation-- I'm going to write b1. In parentheses, this is going to say, which among these bids is the highest? And then I'll write b2 in parentheses to say, which among these bids is the second highest? All the way down to b in parentheses n for the nth highest among the bits.

Now I'm going to have my slots-- say, slot 1 down to slot m. And instead of writing it many times, let me just describe what happens to the kth slot, where k is an arbitrary. Or I guess I was using j before. Maybe I'll call it the j slot.

So who gets the j slot? Well, the j slot is going to go to the jth highest bidder, as before. So the allocation rule is going to be the same as the second price auction. I allocate the jth best good to the jth highest bidder. So it's allocated to jth high square.

But the tricky thing is the payment. So maybe I'll call it pj. And remember, this is a payment per click. Well, it's easier, I think-- I guess a little more space-- to think about alpha j pj. Because remember, I want each bidder to be charged for externality.

So what I want to set equal to the bidders' externality is the actual amount they pay. And the actual amount they pay is not pj. That's the amount they pay per click. So they pay pj per click times the probability that they're actually clicked on. That's the expected amount that they pay. So this is going to be the expected payment. So later we can divide by alpha j and get a formula for pj, but it's just when we're thinking of the externality, it's easier to think this way. Yes.

AUDIENCE: [INAUDIBLE] right?

IAN BALL: No. Well, I guess it depends what you mean. If alpha j is the probability per click of the probability of a click, then this is just the expected payment per click.

AUDIENCE: Per click.

IAN BALL: No, just pj is the payment per click, right?

AUDIENCE: Yeah.

IAN BALL: So if I multiply the payment per click times my probability of a click, that gives me my full expected payment. If I pay 100 per click, and there's a 1/2 chance of getting clicked, then my expected payment is 50 because there's a half chance I get clicked and pay 100 and a half chance I don't get clicked and pay nothing. If there's multiple clicks, then we could more generally think of alpha j as the expected number of clicks. And maybe the expected number of clicks is more than 1, but in reality, each time a keyword is searched, there's a different option.

Well, that is how it used to be. Now people have a lot of automated bidding. Now people are using robots to bid on their behalf, so it's a little more complicated. But originally, yeah. Maybe they were batched a little bit, but roughly, yeah. You can change your bid very, very frequently.

AUDIENCE: If two people are searching for the same keyword at the same time.

IAN BALL: Yeah, OK. There's going to be some lag. I think formerly what happened-- I don't know, some of this may be proprietary. Firms are able to change their bid very quickly. I know firms are able to change their bid maybe at a second interval. So maybe all the keywords that were searched in a given second would use those bids. And then when your bid changed, all the auctions in the next second would use the new bid.

And there was actually-- when they used to use things that were not strategy proof and people were gaming it, this is what would happen. People were investing in technology to change their bids and update their bids very, very quickly. But yeah, probably not literally per auction, but every auction in a second, which is a lot.

AUDIENCE: So that's why we're just considering, is there one click or not? Instead of the--

IAN BALL: Yeah. Mathematically, it doesn't matter at all. If you want to think of alpha j, one way of thinking of this is this is the probability per click. And maybe our bid is relevant for a batch of 100 auctions. And then you could just multiply every alpha by 100, and then you get the expected number of clicks in this batch of auctions that's relevant for a bit.

Great. OK, so here's our expected payment. And I was trying to be consistent about superscripts, but I wasn't-- oh, no. I was over here. OK, sorry. This should be alpha j, I guess. I don't know if I've been consistent. I wasn't consistent over here. I'll change this.

So this is actually right. b3 is actually the third bidder, so that's a subscript. Alpha 2 is the second bid slot, so it's superscript. So I think my notation is good.

So now we want to compute, what is the externality that bidder j is imposing on the other bidders? So the first question is, who is bidder j affecting? So over here, we thought bidder 1 was affecting bidder 2 and bidder 3. Which bidders is bidder j's bid affecting? If bidder j were gone, whose allocation would change? Everyone afterwards? OK, yeah.

So what we need to compute is, what is the externality that bidder j is imposing on everyone after bidder j? Because these people above him are not going to be changed. If bidder j weren't there, those people are still going to win. So it's going to be a sum. And let's sum it over everyone after bidder j. So this is the jth highest bidder. We want to sum from the j plus first highest bidder to n.

So within our sum, we want to say, what is the externality imposed on the kth highest bidder? So if I'm the kth highest bidder, where k is bigger than j, when j is here, I get the k slot. But if j were not here, what slot would I get? The k minus 1 slot. So the externality, the change in my clickthrough rate is alpha k minus.

This is the externality. Without j, I would have become the k minus first highest bidder, and I would have gotten the k minus first highest slot. With j, I'm only the kth highest bidder, and I get the k slot. But I have to multiply that change in the clickthrough rate times my value per click.

Again, we don't know that bidder's value, but the way we compute externality is as if people are bidding truthfully. And we have notation. What is our notation for the bid of the kth highest bidder? It's just b to the superscript k.

So this is our formula. And this fully specifies the auction because we've specified which bidder each slot is assigned to. And we've specified the payment per click for each of these slots. But now let's try to understand a little bit what this formula looks like. So let's dig in a little more. Let's go over here.

Let's try to interpret this. Let's divide both sides by. alpha J. So we're going to see that pj, the payment per click in the j slot, is going to equal the sum of k equals j plus 1 to n of bk times alpha k minus 1 divided by alpha j. Good.

And we can see that the amount that we pay in the j slot is a sum over all the lower bids. But each of the lower bids is multiplied by some coefficient here. So let's introduce notation for this. Let's call this wk. And let's see what we can say about wk. Well first, wk is greater than or equal to 0-- in fact, strictly greater than 0.

Why? Well, alpha k minus 1 is strictly higher than alpha k. That's our assumption. The higher slots-- higher slot meaning lower numbered slot. The higher slots have a higher clickthrough rate, so each of these numbers is strictly positive. Now let's try to see what the sum of these numbers are. So let's look at wj plus 1, all the way up to wn.

Well, let's go through it. wj plus 1-- let's plug in j plus 1 for k. We're going to get alpha j minus alpha j plus 1 divided by alpha j. Plus-- well, the next term is going to be alpha j plus 1 minus alpha j plus 2 over alpha j. All the way-- well, what's the last term going to be? We have wn here, so we're going to have alpha n minus 1 minus alpha n over alpha j.

Well, alpha j is underneath everything, so we can make this into a single fraction. And we add alpha j plus 1-- sorry, we subtract alpha j plus 1, and then we add it back. We subtract alpha j plus 2, and then we add it back. So this is going to be what's called a telescoping sum. It's going to simplify a lot. And we're only going to get the first term and the last term. Every intermediate term is going to cancel. So what we're going to get is alpha j minus alpha n over alpha j. Yes.

AUDIENCE: What's the significance of taking this sum if each of the terms is multiplied by a different value so it's not a convenience factor.

IAN BALL: Just as an interpretation. so let's see in a second, and then I think it'll make sense. Yeah. So what can we say about this value? Well, in general, this value is less than or equal to 1. And in fact, it's going to equal 1 if alpha n equals 0. So if we have fewer slots than bidders, this is going to equal 1. You might worry if there's fewer slots than bidders.

Are we ever going to divide by 0? It turns out that top's going to become 0 before the bottom does. and we're not really going to have to worry about it, but I will go through the details. So it turns out this is going to be less than or equal to 1 with equality if m is strictly less than n-- if there's fewer slots than bidders, which is a pretty standard phenomenon.

You have many advertisers wanting to advertise on just a few limited slots. So generally, n is going to be bigger than m. And I guess the trick I've used here is if m is bigger than n, well, let's just create null slots beyond the m. I guess I should have said this.

We could have just stopped or we could say, yeah, well there's actually an n plus first slot, but it just has 0 clickthrough rate. It's just a fake slot. And if we do that, then all these alphas are well-defined. They're just all 0 after the m. And if we do that, then we see that this sum is going to be exactly 1.

So why did we do this sum? The sum was for interpretation. What is the payment per click that I make if I'm the jth highest bidder? It's a weighted average of all the bids that are lower than mine. The weights are non-negative, and the sum of the weights-- well, in this general case here, it's exactly 1. So it's exactly a weighted average. And now we have a nice comparison because-- I'll go back over here.

If we looked at the GSP, or the Generalized Second Price auction, pj was what? Just using our notation, what was pj? What was the payment per click for the jth slot? Right. It was the j plus first highest bid. Under the VCG mechanism, pj-- well, we have this form over here. But let's just check.

K ranges from j plus 1 to n, so these range from bj plus 1 to n. So we get a weighted average of all the lower bids. And we can exactly see why the VCG-- or at least there's some intuition for why, under the GSP, we see bid shading. And under the VCG, bidding truthfully is optimal.

In short, it turned out in the GSP-- while it seemed like the right generalization to have you pay the next highest bid, that was actually too high. And as a result, people wanted to shade their bids. And in a VCG, you're actually ask to pay less relative to your bids. Instead of paying the next highest bid, you pay a weighted average of the next highest bid, the next highest, the next highest, and so on. And these later bids are lower than this bid here.

In particular, let's look at a special case I think is instructive. What if alpha 1-- so this technically isn't covered by our assumptions because we assumed it was strict. But we can think of this as a limit. So let's say there's n bidders, and the n slots are all basically the same.

So we have n bidders and we have n identical slots. How would people bid in the GSP? Would you want to bid your evaluation for the slot? How would you bid if you were in the GSPN and you knew there were n slots? Yeah.

AUDIENCE: You could end up bidding nothing and still you could just end up-- even if you're in the last slot, it doesn't really matter because you'll have the same value.

IAN BALL: Right. So here, there's really no scarcity. There's n slots. They're all equally good. There's no competition between the bidders at all. It's not like we're fighting for that top slot because the next slot is just as good. So we see that the GSP is going to induce really intense bid shading. People are going to shade their bids a lot. In fact, people are generally going to bid 0.

There's no shortage? I can just bid 0. I'll still get one of the n slots, and I'm happy with that. So you might think, well, how is the VCG going to get people to be truthful at all? How could we possibly get people to pay anything if there's no scarcity? Well, what is the payment rule going to be in the VCG? Well, what are these weights going to be if all the alphas are the same?

AUDIENCE: 0.

IAN BALL: 0. So in fact, in the VCG mechanism in this special case, people are always just going to pay 0. So you might think, oh, VCG-- how do we get something for nothing? Well, it's like if people had a really strong incentive to shade their bid in the GSP, then the VCG is going to shade on their behalf and just charge everyone 0 anyway, and then people are going to be willing to be truthful.

So nothing comes for free. How do we get people to bid more? They're willing to bid more because it's a function of their bids. They're asked to pay less. And that's, I think, the key idea or key insight of the VCG mechanism.

So maybe in the last 15 minutes, we'll try to-- I'd like to go over maybe the general formula for VCG beyond this setting. So it turns out the VCG mechanism, it was applied to this setting of ad auctions. But it makes sense, and it works in a much more abstract setting. And we're just going to go over how that works.

So now I'll say abstract VCG. So the general setting of VCG is-- well, we're having to make some decision for all the bidders. Here, the decision was how we match the bidders to the ad slots. Maybe the decision is, do we build a bridge? There are many, many things it could be. How we allocate some item among the bidders, there could be many, many things.

So we're going to be totally abstract. And the decision is just going to be some abstract x in some set X. But if you want to make it concrete, think of x as the allocation of bidders to slots. Maybe another classic example would be the allocation of planes to airport landing slots. This is a classic application of BCG. Or maybe routing internet protocols, routing packets through the network-- that might be what x represents.

But the crucial thing is x is a collective decision that affects everyone. x does not represent bidder 1 gets slot 1. It represents the complete mapping of bidders to slots or the complete allocation of airplanes to landing slots, and so on. We now need to describe the bidder's preferences or the player's preferences.

So we have-- let's say player i has valuation vi. Now, before, vi was just a number. When we were in a setting where we were allocating a good, my valuation was just, how much am I willing to pay for that good? But now my valuation needs to specify how much I value every possible social decision that's made. Let me say social or collective decision.

So now my valuation is a function, vi, from x to r. So I'm just going to call it vi, but remember, it's a much richer object now. It's not just one number that says, how much do I value an item? It's an entire mapping that says, if this is the allocation of bidders to auction slots, that's how much utility I get. If we have a different allocation of bidders to auction slots, here's how much utility I get, and so on. Yeah.

AUDIENCE: What do you to consider how other bidders are put in different slots? If you're in a fixed auction slot, can it just be your auction slot to your utility?

IAN BALL: It's a good point. So if we're willing to impose more structure on x, maybe x says, this is what you get, this is what you get, this is what you got. And my valuation only depends on my component of x. But at this level of generality, we're not even assuming that kind of structure.

And it may be that we want to be-- the point is, this is more general, so it may be that if we're allocating planes to landing slots, I also really like if my competitor gets a bad slot. I don't just care about getting a good slot. So in principle, that could be the case. Maybe there are ways we can restrict it, but this is more general and it makes the notation cleaner, so we're not going to impose that assumption. But I think that's a very reasonable assumption.

OK. So what are we going to do? Well, before, in a second price auction, we had this idea that people were bidding truthfully. We'd also like people to bid truthfully. But if we want people to bid truthfully, well, we have to just ask them what their valuation function is. And that's what the mechanism is going to do.

We're going to say, everyone, tell me what your valuation function is. So instead of telling me just how much you value one item, we're going to say, tell me how much you value every possible social decision. So the mechanism, we'll say, elicits reports. And we'll often denote these v1 hat through vn hat.

So what's the interpretation here? v1 is bidder 1's true valuation function. v1 hat is effectively a bid. It's the valuation function that bidder 1 reports to the mechanism. And that may not be truthful. We want to make sure that people have the right incentives to report truthfully. But they're reporting an entire function, which may be a little bit confusing and abstract at first. So what does the mechanism do?

We have 5 minutes. We're just going to write two formulas, but they're a little long. So just like an auction, a mechanism is going to need to specify two things. It's going to have to specify an allocation rule and a payment rule. We need an allocation rule and a payment rule.

The allocation rule says, if these are the reported valuation functions that people send to me, what decision am I going to choose? And the payment rule says, if these are the valuation functions that are reported to me, how much am I going to ask each agent to pay? So the allocation rule, we need to define xv1 hat vn hat.

These are the valuation functions that are reported. x of this is going to be the allocation or the decision-- allocation or decision rule. I'll use these interchangeably. Well, we want to allocate things efficiently. So what does that mean? This is going to be in the argmax-- I think the notation here is a little tricky-- of x prime in x of the sum.

So let's just understand what this is. This says, if we choose decision x prime, what is the sum of the players' utilities when we compute them, with respect to what they report? Taking the argmax of this says we're going to choose the decision that maximizes the sum of the reported valuations. What we'd really like to do is maximize the sum of the true valuations, but we don't know those. We only know the reports. So all we can do is take in the reported valuation function and choose the decision that makes the sum of the valuations as large as possible.

And that's what this is. And sometimes we call this x star to denote that this is efficient. I think this is a bit abstract, so any questions on this? No? Good. OK. There could be multiple maximizers. And we'll just choose one of them. It doesn't matter. If there's two things, that would be efficient.

So as an example, if we were in the auction context, then this would be the allocation that gives the highest bidder the first slot, the second highest bidder the second slot, and so on. That's what x star would represent. And v1 hat would be the report of how much I value each slot, each click.

Now, the payment rule is exactly what we said before. We want to charge each bidder their externality. So ti of v1 through vn hat is saying, if these are the valuation functions that are reported by the bidders, how much are we going to charge bidder i? We're going to charge bidder i informally the externality imposed on all others.

So let's write that more formally. What is the externality imposed on all others? It's how well everyone else would do if bidder i weren't there minus how well they actually do. That's what the externality is. So how well would they do if bidder i weren't there? Well, let's just look at the best they could possibly do. So the best they could possibly do if bidder i weren't there would be this.

I'm looking at everyone other than bidder i. I'm using their reported valuation functions. And I'm saying, if we chose x prime, this is how well everyone but player i would do. And let me choose the allocation that makes those people as well off as possible. So I want to maximize the welfare of everyone other than bidder i. That's what we could do if bitter i weren't here-- or in other words, if we didn't care about bitter i, this is how well everyone else could do.

Now I want to subtract from that how well they actually do, how well they do given that bidder i is present. How well do they actually do? Well, it's the sum over j not equal to i-- vj hat. But now what I want to plug in here is the decision that's actually chosen. And the decision that's actually chosen is this one up here.

And this is the formula. It's a little abstract. I'll put this in parentheses. And if you go to the notes, you can see a proof as to why using this mechanism, it's always optimal for bidders to bid truthfully. So let me just write that fact. This is a fact that you can prove.

So for each bidder i and each valuation function vi, bidding vi hat equals vi. What I mean by that is reporting my true valuation to the mechanism is optimal, no matter how everyone else bids. And when I say no matter what others bid, well, that's basically vj hat for j not equal to i.

So whatever everyone else other than me reports to the mechanism, if my true value is vi, bidding or reporting my true value vi is going to be best for me. It's going to be weakly better than any other bid. And this is an incredible property of the VCG mechanism. The proof is three lines, but it's a little dense, so I encourage you to check the notes to see the proof. And maybe I'll stop here a few minutes early. And I'm happy to stick around to go over any questions people have about VCG.