[SQUEAKING]

[RUSTLING]

[CLICKING]

GLENN ELLISON: OK, let me go ahead and get started. So today I'm starting my week on static competition. So think about what's going on here. So far, the first two weeks have mostly been monopoly models. Obviously, most industries or most things of interest have multiple firms in it. So I'm going to now start every lecture for the rest of the term, it's going to be multiple firms competing.

When Tirole goes through his book, he does this long-run to short-run chart where he talks about in the short-run firms can basically do nothing in competing with each other than setting prices. And then you get a sequence of variables that take longer and longer to adjust. And I think the way his book is written is, in some sense, the way we solve models is always by backward-- like multi-stage games by backward induction. You solve the last stage and the next to last stage and the next to last stage.

And so his book is written this way, where you go backward from here. And that's how my course is written. Although, these days it's really going to be an awful lot of discussion of this and then other things just as individual topics. So the next few weeks, it's going to be short-run things.

When we're discussing short-run competition, what do we talk about? One of the big things we talk about is markups. We talk about markups a lot because markups lead to deadweight loss. That's a major welfare concern. But they're also going to be, just like in price discrimination, you get that second welfare concern of the misallocation of the goods that are being sold. We'll also get mispurchasing of goods that becomes another consideration here as well.

So what I'm going to do in this lecture is I'm going to start with all of the classic stuff in I/O that people don't do anymore but do that as, I think, background for things that is how modern studies of competition are done. So anyway, I start-- and I'm talking about old historical stuff. And I added some pictures of the old, dead white guys who developed these things. So there's Cournot, Cournot competition 1838.

So obviously, Cournot was thinking about a world that's very different from ours, a world in which many of the things that people bought were completely undifferentiated goods, like wheat and corn and wool, or whatever, where just a product is a product. And there's a universal market for it. So Cournot thought about a model where you had n firms. There was some homogeneous good that they were all producing that had inverse demand P of X. The producers have cost functions, ci of xi that depend on what is the quantity produced by Firm I.

In some sense, Cournot invented Nash equilibrium. So Cournot thought of a model. He had this timing where the firms are simultaneously choosing outputs x1 through xn. When they choose those outputs, they're going to all bring their goods to the market. There's going to be a market-determined price. And all the firms are choosing their quantity, assuming that the quantities chosen by the other firms are all held fixed. And that's going to be the optimality condition, OK?

OK, what happens in Cournot competition? So in Cournot competition, the firm is maximizing over xi, xi times P of xi plus sum over j different from i xj minus ci of xi. So that's every firm's maximization problem, is maximize quantity times the market price minus the cost. Your quantity affects how much you produce. It ends up affecting the market price through its impact on the aggregate supply. And it obviously affects your cost of production.

So what do we get? Well, if a profile is a Nash equilibrium, it's got to satisfy first-order condition for this for each i. What does the first-order condition look like? So first, it's this times the derivative of this gives you, you get a P of the sum of the xi's minus the derivative of this is marginal cost. The third term in the thing is xi times the derivative of that. So we get Pi prime times xi star equals 0. So this is what the classic Cournot first-order condition looks like.

What are some implications? So first of all, price ends up greater than marginal cost. This is price minus marginal cost is the first term. If xi is a positive number, P prime is a negative number. So this is price minus marginal cost equals minus Pi prime xi is a positive number.

Production in a Cournot-- so we get a second welfare loss in a Cournot oligopoly. So the first welfare loss is there's deadweight loss. Second is production is inefficient in a Cournot oligopoly. Production is inefficient in an asymmetric oligopoly because if the firms are asymmetric in equilibrium, x1 is going to be different from x2. If x1 is different from x2 and P prime is the same for everyone, because it's evaluated at the aggregate demand, then that means-- then the P is the same for everyone. That means the c primes have to be different.

So in an asymmetric Cournot oligopoly, the realized marginal costs are different. So the quantity that's being produced is being produced inefficiently. OK, example to think about this is for-- if you think about the model for producing oil in the world today is something like a Cournot oligopoly. There's a set of countries. Every country, the ones we're centrally producing are choosing how much to produce. If each country is maximizing its production, then you're going to have countries like Saudi Arabia has oil that they could pump out of the ground for $5 a barrel or $10 a barrel that they're not pumping it out of the ground. Canada is getting oil out of tar sands for $50 a barrel. And there's an inefficiency there from low-cost producers not producing, while high-cost producers are producing, holding the quantity fixed.

OK, firm outputs are usually strategic substitutes. So if you look at a Cournot model, if I put q1 on this axis and do what's the best response for Firm 2 to an output of q1, you get this downward-sloping line. And the equilibrium looks like this. If I put q2 on this axis and graph BR1 of q2, Cournot equilibrium is typically this intersection of two downward-sloping lines.

OK, that's what it normally looks like. It doesn't have to look like that. And the Cournot equilibrium is not always unique. You could imagine if you could somehow draw these things with wiggles in them so that they intersect multiple times.

OK, fourth result of this, take this expression, price minus cost is minus xi star P prime, and then divide both sides by price. I get an expression that looks very much like the monopoly pricing expression, except it has an xi star on the right-hand side instead of having an x star on the right side. But in some sense this is the monopoly first-order condition. So because this is the monopoly first-order condition, except with an xi here instead of a capital X here, I get this expression, price minus marginal cost over price is minus xi over x star times 1 over epsilon.

So if we think about how does competition reduce markups in the Cournot model, obviously, elasticities-- if you don't have a constant elasticity demand curve, elasticities also change as output changes, because the elasticity is nonconstant. But if you think about the constant elasticity intuition, this is going to say that, in N firm oligopoly, prices are going to go down roughly like 1 over n. And this is one of the features that made Cournot popular is that people have this intuition.

With monopoly, prices are high. As you get more competition in, prices ought to decline, and they ought to decline in some decreasing returns to competition way. And they ought to asymptote to perfect competition as the number of firms goes to infinity. And Cournot has those properties, and that's one of the reasons why it became a popular model to use.

OK, notice that, like marginal costs are different with asymmetric firms, Lerner indexes are different with asymmetric firms because each firm's markup is xi over x star. And so the xi's are not all 1 over n at equilibrium. The markups are different. So we can think about what is the average markup in an industry across firms? So let's define the industry-wide Lerner index to be price minus average marginal cost over price, where we take a weighted average of the marginal costs, giving weight xi over x star. So we give more weight to the firms that are producing more. So this is price minus average marginal cost over price.

Theorem is that-- and you can just-- this is just basically taking a weighted average of these, weighting them by xi. So the theorem is that the industry-average Lerner index is negative H over epsilon, where H is the sum of xi over x star squared. OK, so H is the sum of the squared market shares.

And so the Cournot model predicts that-- the Cournot model predicts that the markup in an industry depends on sum of squared market shares. And sum of squared markup shares is a summary statistic for how large markups are. And this was sufficiently influential that if you go back and read how US merger guidelines were designed in the 1950s, they explicitly refer to Herfindahl indexes. And they're explicitly referring to Herfindahl indexes in the sense of like if-- I forget what the numbers are, but they're something like this.

If the Herfindahl index in an industry is less than 0.18, then we will allow mergers. And if the Herfindahl index is between 0.18 and 0.28 and the delta H-- then what we'll do is look at whether delta H is less than or greater than 0.01. And they will tend to consider mergers as questionable or not, depending on whether the change in the Herfindahl index is bigger or smaller than 0.01. And then, for industries where the Herfindahl index is already above 0.28, then they're going to be-- then they're going to be-- then mergers are typically not going to be allowed.

But this is just-- because the Herfindahl index had this role in the theorem that it was a measure of how competitive an industry was, it took this very central role in competition. And if you look at the-- I discussed the industry cross section I/O from the 1960s, constant regressions of mark-ups on Herfindahl indexes because the Herfindahl index was supposed to be the correct measure of how concentrated an industry was.

One thing to note, though, you definitely should not be thinking of the Herfindahl index as a welfare proxy because you think about it, it's measuring mark-ups relative to the cost they're there. So, for instance, if you have two firms that are perfectly symmetric, the Herfindahl index is 0.5. It's 1/2 squared plus 1/2 squared is 0.5. Suppose one of those firms gets more efficient and lowers its cost. When it lowers its cost, everyone is better off. Social welfare goes up, but the Herfindahl index also goes up because the firm with a lower cost now starts producing a higher share of the output.

In the old days, one could-- and let me say, I think big mistake students still make today is trying to be too original. I think almost the best way to do research, especially when you're starting out, is always to take some existing paper and improve on it because if you take an existing paper, you improve on it. You've got something better than an existing published paper. Whereas, you start from scratch with some creative idea. You spend years on it. And then you realize two years later that what you've done has created a worse version of something that already exists. Whereas, if you start with the existing frontier paper, you know where the frontier is. You improve on it. You've got a new frontier paper, OK?

Anyway, back in the day, Cournot was 1838. Bertrand, 45 years later, picks up Cournot's paper and says, I can come up with an interesting comment on Cournot. So his comment on Cournot was that many firms don't choose the quantity that they're going to produce. What they do is instead set a price for their product and then sell however much at that price as people are willing to buy.

So you just flipped around this discussion from firms choosing quantities, treating everyone's quantity as fixed, to firms choosing prices. And his comment is basically that predictions of Cournot's model change dramatically if you make that simple flip from firms are choosing quantities, assuming quantities are fixed, to firms are choosing prices, assuming prices are fixed.

So anyway, Bertrand's model, let x be a market demand function. Assume that x is weakly decreasing and that P times x of P is bounded. Assume that all firms have unit cost c. Firms simultaneously announce prices. All demand goes to the lowest price firms. Then, we don't get this Cournot continuity. We get an immediate discontinuity, which is two firms is enough for perfect competition.

So the unique equilibrium of this model is P1 star equals P2 star equals c. And this is what Frank Fisher and his writing about economic theory would say. This is an exemplifying theory. You make these somewhat extreme assumptions, that is, firms announced prices, and all demand goes to the lowest price firm, even if it's one penny less than the other firm. But what you show is that, with these assumptions, you get of extreme outcome, which is competition works differently from Cournot, and it's a discontinuous thing. And any number of firms is enough to have perfect competition.

One version of this model that still does get used is if you think about firms that have asymmetric costs, you have two firms with cost c1 star and c2 star. What's the equilibrium? The equilibrium is the firms end up both pricing at c2 star with all of the demand going to the firm with the lower cost. If you want, you can think of it as Firm 1 is charging c2 minus epsilon and getting all the demand. But when you have asymmetric Cournot, prices are determined by the second lowest cost in a market. And I think that's the important intuition. That's the second lowest cost determines what happens in markets that are this very intense price competition or undifferentiated products.

AUDIENCE: Does this change if we don't have complete information on the cost?

GLENN ELLISON: Yeah, so let me see. So if you did this model where you had--

AUDIENCE: Like private info, like you don't have any knowledge on their distribution.

GLENN ELLISON: Yeah, so in some sense, if you have private information on costs, this is some sense that's what we normally refer to as an auction model. It's an auction. I don't know what the value of the painting is to you. You don't know what the value of the painting is to me. So I bid something, giving myself some expected markup because if I bid cost, then if I win or if I lose, I'm unhappy both ways. And so I bid something in between. In the buying something, I bid something in between my value and 0. And you do the same thing, and we get positive markups for that reason.

Yeah, so that becomes like in the-- values uniform 0, 1 model, you bid v over 2 is the equilibrium. So in some sense that is a model. You end up reconstructing something like Cournot that you get positive markups. And as the number of firms goes to infinity, the markups go to 0. So that would be another way to get a Cournot-like prediction.

Anyway, I do talk about-- I'll do auctions in like week eight or whatever. But we actually get that-- the place where that analogy is more direct is like procurement auctions. You have auctions for contracts to pave a road in the city of Cambridge. And those are allocated by auction. And exactly, you don't know what the rival's costs are going to be for that paving job. So both of you bid some markup, hoping to win, yeah.

OK, comment on-- again, go forward another 46 years. Comment on-- Cournot comment on Bertrand competition by Harold Hotelling. And I wanted to put him here because Harold Hotelling is my great, great grand advisor. So Hotelling was Arrow's advisor. Arrow was Maskin's advisor. Maskin was Fudenberg's advisor. Drew was my advisor.

So anyway, I have some fondness for Harold Hotelling. And he was actually-- so if you go look at my advisor family tree, they're all mathematicians beyond some point. So you get to all those French mathematicians, whose names you hear in probability theory or whatever, because Hotelling had his PhD in math rather than in economics. So my family tree goes back three branches in economics. And then everyone's a mathematician from that point forward, like for Hotelling's advisors.

But anyway, Hotelling, through-- well, mostly through having advised Ken Arrow, Hotelling does have over 1,000 descendants, grand advisees or whatever. So anyway, so Hotelling-- and actually, I love Hotelling's writing also. You get him to-- he's got a very, in some sense, self-confident but trying-to-be-understated style of this, of-- "after the work of the late Professor Edgeworth, one may doubt that anything further can be said on the theory of competition.

However, one important feature seems to have escaped scrutiny. This is the fact that of all the purchases of commodities, some buy from one, some from another, in spite of the moderate differences in price." And that's saying that price reaction should be continuous rather than nondifferentiable as they are in Bertrand's model and then working out the [? direction. ?]

So anyway, Hotelling's model is-- and in some sense, Hotelling has opened the path that everyone has followed since, which is thinking about competition in terms of heterogeneous consumer populations and some underlying type space of heterogeneous consumers. So he had this sort of line, where he had consumers arranged on 0, 1.

And actually, Hotelling spent a lot of his career at UCLA and I think was thinking about this as a beach in Southern California. And so Hotelling did have this model you can think of as you've got the two stands on the beach, each one selling drinks and hot dogs and ice cream or whatever. And then you have all the customers, who are lying on blankets somewhere along the beach. And then when they want to get something to drink, they pick up, and they walk to the nearest stand. And then they buy something, and then they walk back to where they were sitting on the beach.

So his model is we have these consumers, and they have the utilities. So I put theta on this axis. So then if I graph the utility that you get buying consumer of type theta buying from Firm 1, I get some utility that looks like this, v minus t times theta minus P1. So that would be my utility of buying from Firm 1.

And then, consumers also have a utility of buying from Firm 2. So let me on this axis put v-- this is v minus t1 minus theta minus P2. So the argument is t is my disutility per unit of distance walking. And so my utility from buying the product from Firm 1 is v minus t times theta, because theta captures my distance, minus P1. And if I buy from Firm 2, it's v minus t times 1 minus theta minus P2.

And so he did this in a one-dimensional-- this one-dimensional type space captures your preference for both goods. And so we assume there's a constant marginal cost c. He considered, again, the Nash equilibrium before Nash of firms simultaneously announced prices P1 and P2. And then, firms announce those prices. It's a simultaneous move game. And then, consumers decide who to buy from, and then the firms earn their profits. And we just look for the static Nash equilibrium.

The standard case we discuss in this model is if v is large, then there's some consumer theta hat who's indifferent between buying the two products. So this consumer theta hat has v minus t theta hat minus P1 equals v minus t 1 minus theta hat minus P2. So at those prices, that consumer is indifferent between the two products.

You solve that equation, and it tells you that the indifferent consumer theta hat is 1/2 plus P2 minus P1 over 2t. With that indifferent consumer-- do I have it on the next slide or not? So with that indifferent consumer, then you get x1 of P1, P2 is 1/2 plus P2 minus P1 over 2t. And x2 of P1 P2 is 1/2 minus P2 minus P1 over 2t, OK?

The demands add up to 1 if-- as long as this consumer is willing to-- as long as v is large enough, this consumer is willing to buy. So everyone buys from one firm or the other. So there's no deadweight loss in this model. Demands end up being independent of v. So this is if, whatever, v minus t times 1/2 plus P2 minus P1 over 2t minus P1 greater than equal to 0.

So as long as this indifferent consumer is willing to buy, demands are completely independent of v. And just are these linear-- these are linear demand curves. So what he's done is he's just proposed this model of heterogeneous consumer populations and said, with that heterogeneous consumer population model, we just get price competition with these particular linear demand curves.

And what happens when you have those linear demand curves? Well, Firm I is just maximizing price minus cost times 1/2 plus pj minus p over 2t. So, if it sets a higher price, it gets a higher markup per consumer. But the number of consumers go down, with 1 over 2t being the slope of the downward-sloping individual demand curve facing it.

I take the first-order conditions for this. It's just algebra. And, given that it's a linear demand curve, we get a quadratic objective function. It has an interior maximum. And what we get is the best response for Firm I to a price pj is 1/2 of c plus t plus pj.

So, if I try to graph what's happening here, P1, P2, so if P1 is 0, the best response for Firm 2 is 1/2 c plus t is positive. And then, it's a line that's upward sloping. So this is the best response for Firm 2 to a price of P1. And this is 1/2 c plus t.

And then similarly, the best response for Firm 1 to a price of P2 is going to look like this. So I'm graphing Firm 1's best response. I'm graphing this thing backwards, as you do in game theory, where as P2 is higher, I'm graphing P1 is higher. But this line has slope 1/2 relative to this. This one has slope 1/2 if you view it as a function of this line. So look at these two upward-sloping lines. They intersect. That's the equilibrium. And we get P1 star equals P2 star equals c plus t.

So what are the economic implications here? One is that markups are proportional to the product differentiation parameter t. So it's the amount of product differentiation in a market that ends up determining how large are the markups? The hotter it is and the more unpleasant it is to walk across the sand, the higher will be the prices. And people today, I think, use this model not just for differentiation in physical space driving, but differentiation in taste space as well.

So if you think about competition between Coke and Pepsi, how much do you like Coke? How much do you like Pepsi? There are some people who love Coke and dislike Pepsi. There are some who love Pepsi and dislike Coke. There are some who are completely indifferent between the two. The amount of taste heterogeneity is what determines the markups. And numerically, with these linear demand curves, it's the amount by which people at one end of the line versus the other-- it's the relative preference of people at one end of the line versus the other. This person thinks Pepsi is t better than Coke. That person thinks Coke is t better than Pepsi. t is the equilibrium markup.

You can do an n firm circular version of this. So imagine that you have-- obviously, there are lots of parts of the United States-- well, obviously, you'd have to go all the way down to South America. And some parts, no one's going to be on the beach. But you could imagine a giant beach that's-- we have a resort island where the beach is just circle around the whole thing. If you put n firms around this, it's as if the competition between any of these two firms is now going to be like t times 1 over n is going to be the effective distance between them. And so if t times 1 over n becomes the effective distance between them, you're going to get P star equals c plus t times 1 over n.

So this is a model where it can be like Cournot. And, as you put more and more firms into the same finite dimensional product space, markups decline and are continuous with perfect competition as n goes to infinity.

Anyway, I'll just say this later. Actions are strategic. Complements just means that the reaction curves look different from how they looked in Cournot. Here, it's an intersection of upward-sloping lines. Cournot was an intersection of downward-sloping lines. Understanding some of the implications of the model, that's going to matter later, OK?

OK, any questions on Hotelling's model? All right, OK. OK, so vertical differentiation, final thing we think of is that there's many examples like Coke and Pepsi. Clearly, one is the better product, and one is-- sorry, two products are just a matter of taste, whether you a versus b. There are other examples we think of where differentiation consists of one product is better and one is worse. So if you think about do people want to buy the BMW, or do they want the Toyota Yaris or something? Everybody would rather have the BMW. Everyone admits that this product a is better than product b.

But the question is, how much are you willing to pay for product a versus product b? You could have, again, maybe some people still like flip phones. But you have the modern phone competing against the old flip phone, with just a keypad and a little screen like this. Everyone agrees that the modern phone is the better phone. But people just differ in how much they're willing to pay for the better product. So the consumer preferences here are going to be vertical rather than-- what we call vertical rather than horizontal.

So do this example. I've got two firms. They have goods of quality. SL is my flip phone. SH is my iPhone. And the low-quality phone, everyone would agree, is worse than the high-quality phone. We have consumers have types, again, distributed on a unit interval, or sorry, nonunit interval now, theta lower bar, theta upper bar. And the utility functions are going to be the consumer-- utility of a consumer of type theta buying product i is theta si minus pi. So the theta is your marginal value of a unit of quality.

You could also just think of this as a model where consumers are differentiated by income and have a different marginal value of income. And therefore, they have all the same utility that they get from higher quality. But just some people have more money than others. They have a greater disutility of spending money, OK?

OK, and so consumers have this utility if they buy from Firm I. They get 0 utility if they buy from neither firm at all. I'm going to, for simplicity, assume a unit mass of consumers, a unit density of consumers and an interval theta lower bar, theta upper bar.

OK, so what's going on here is, here, I'm typically going to have my theta space-- often, we have our theta space start at 0. But here, everybody in my model is going to like quality. So I'm going to think of a model. We have this sort of interval theta lower bar, theta upper bar. And so consumers are arrayed on this interval of positive values.

So again, consider the same game that Bertrand and Hotelling considered. Firms simultaneously choose prices PL and PH. Taking the other firm's price is fixed, consumers decide who to buy from. And then firms earn profits. That's the end of the game.

What do the utilities look like? Here, let me first draw a line like this. This is the line theta SL minus PL. So it starts at negative Pl, and then it goes up with a slope of SL. This is the utility that you get-- this is UL of theta. That's the utility you get buying from Firm L. Equalas UL of theta.

OK, then other line I'm going to draw is Firm H has the higher quality product. So if you had type 0, which nobody does, you'd get utility negative PH from buying that product. But then I'm going to assume that's a steeper-sloping line, like this, theta SH minus PH. It's a steeper-sloping line because SH is bigger than SL, and the slope is S, OK?

So what happens in this model? Again, there's some-- in equilibrium, there's be some indifferent consumer of type theta hat. Consumers with types above theta hat, so these guys all buy the high-quality good. These consumers, the upper envelope, is these guys buy the low-quality good.

As I've drawn it, there's another point that's of importance called theta-- I'll call it theta-- I named theta prime in the slides. Anyone with a type below theta prime is going to buy nothing because neither good is worth paying what's being charged. I've drawn this in the case where theta prime is below theta lower bar, and theta hat is between theta lower bar and theta upper bar. And in that case, what are the demands? The demand for the high-quality product is just theta upper bar minus theta hat. And the demand for the low-quality product is theta hat minus theta lower bar.

If theta prime were to the right of this, the demand for the low-quality product would be theta hat minus theta prime. But anyway, I drew it like this, so this becomes the equations. And then again, this theta hat satisfies theta hat SH minus PH equals theta hat SL minus PL. So that tells you that theta hat times SH minus SL equals PH minus PL. That gives you this expression here for what the demand is for product H.

So again, it's a story, and it's a story that just basically what is done is it's just given us, again, two linear demand curves. These are both-- they've got other parameters in them, and it's not so obvious. But you just look at it. It's like, yeah, that's just-- it's a linear function of PL and PH. It's a linear function of PL and PH.

Hotelling was a symmetric linear model. The two demand curves looked exactly the same. Here, the demand curves look different. In particular, if the prices were equal, everyone would buy from Firm H. So in some sense, Firm H's demand is higher than Firm L's holding the price is fixed. But that's what the equations look like.

But it's just a shift of demand curves. One demand curve is higher, but the demand curves are both linear with the same slope and the two variables. Any question on that, getting from the equations to the slopes?

So anyway, that's what our demand model looks like. And then again, the question is like, what happens in this competition? And again, it's just it's solving a-- it's just solving two equations and two unknowns to find what the demand curves are. It really is not so different here from what Hotelling did.

So if I plot P1 and P2 again, so the best response to Firm H of a price PL turns out to be-- if you remember, the other one was 1/2 p plus c plus t. So it turns out to be 1/2 p plus c plus, instead of t, we get theta upper bar minus SH minus SL, which is like the amount of differentiation that there is between the two products from the perspective of the theta upper bar consumer. So it's a lot like p plus c plus t but just with the amount of difference between the products here being from the perspective of the high theta consumer.

And so what that looks like is we start again here. So we can put PL here, PH here. So it starts at 1/2 PL plus c plus something or other. So it starts at some positive thing, and it goes up with a slope of 1/2. So this is the best response for Firm H to a price of PL. And this is 1/2 of c plus theta upper bar times SH minus SL.

Now, Firm L's best response is 1/2 of p plus c minus something. And so the minus something is again related to how much differentiation there is. Notice because when PH is 0, this becomes 1/2 of c minus something. This thing could even be negative, or it's certainly below c. And it could even be a negative number if differentiate, depending on how big differentiation is. But anyway, I'll draw it as starting as a positive number, but it starts from some smaller amount. And then it also goes up at a slope of 1/2.

So we get BR H-- BRL of PH. Looks like that. And it starts from 1/2 of c minus theta lower bar times SH minus SL. And, given these things both have slope of 1/2, again, there's a unique intersection. And it's right there. And you can think of this as PL star comma PH star. And what we get is this asymmetric equilibrium. The formulas look the same, but this one's got twice the bigger number minus the smaller number. And this one has 1 times the bigger number minus twice the smaller number. So this is a bigger number than this one.

And so what we get is, again, vertical differentiation, just like horizontal differentiation, it creates finite demand elasticities. It creates positive markups. If you put uniform distributions on the vertical tastes, it can be linear demand curves, just like the horizontal model is.

In equilibrium, the differentiation, again, leads to positive markups. Leads to positive markups, but it leads to positive markups where Firm H is favored. Firm L is disfavored. Notice also that if theta upper bar and theta lower bar are too close together, this doesn't work, and Firm L just has to drop out of the market because the equilibrium looks like theta upper bar minus 2 theta lower bar over 3.

As theta lower bar approaches 1/2 of theta upper bar, Firm 2 is just pricing at cost. If firm theta lower bar was even closer to theta upper bar, what you'd get in equilibrium is Firm L would just be shut out. It would be driven down to pricing-- to play [INAUDIBLE], that would have to be pricing below cost. It eventually doesn't price below cost. It can't compete in the market, OK?

So vertical differentiation, like horizontal differentiation, produces finite demand elasticities. It can give us, again, standard Cournot-like markups. It doesn't give us a model where Cournot is-- it's like Cournot had that continuity n goes to infinity. Price heads to cost. Here, this doesn't happen because there's only room for a finite number of firms. You start taking this space, and you try to have three or four or five or 10 different firms selling to this.

At some point, the firms get too close together. The firms that are below just get shut out of the market and can't compete. And so it's only the high-quality firms that are making any sales. And so there's some minimum level of markups you're going to have to get there. Although, obviously, as the firms crowd the space, the difference between the SH firm and the SH minus epsilon firm, you're only going to get epsilon profits. But you're going to get it with just a few firms active, OK?

OK, and obviously, I think modern IO-- in reality, we think that vertical differentiation and horizontal differentiation are present in most markets. We think most products, there's some extent to which we think product a is better than product b that everyone would agree on or almost everyone would agree on. And then there's other horizontal aspects. Everyone wants a phone that has better battery life and a better screen resolution and a better camera. But then there are other things, people like the look and feel of a versus b and are going to have-- know the operating system of a versus b and are going to have horizontal and vertical tastes.

So let me go back to horizontal differentiation now. And this is motivated by how the empirical literature and IO has developed. So generally, the standard-- Hotelling is a very nice model of competition on a line. Hotelling is not as good a model once you have three firms instead of two firms. Because you have three firms, you have to think about where do we put the three firms and product space.

One thing you could do is-- what did I do with my eraser? There it is. You could, if you're someone who likes high school geometry, you could go back to your youth and say, OK, here's what I'm going to do. I'm going to have the three-firm Hotelling model. And I have Firm 1 here, Firm 2 here, Firm 3 here. And there's going to be this set of consumers who are indifferent between Firm 1 and Firm 2 based on their prices and a set of firms indifferent between Firm 1 and 3 based on their prices and a set of firms indifferent between Firm 1 and 2 based on their prices. I think that has to go through the same point.

And so we're going to get these three different regions. So in equilibrium, these guys buy from 1. These guys buy from 2. And these guys buy from Firm 3. I'm going to use everything I remember from high school geometry to figure out the areas of those things and how the areas of those things are functions of P1 and P2 and P3. And then, when I get to 4, I'm going to make it a tetrahedron.

And I'm going to draw all the volumes of those things. But it quickly becomes just an intractable mess with lots of corners. And you also think, why were we doing this in the first place? It wasn't like we felt that people were uniformly-- uniformly distributed on the beach seemed OK. Once you get between competition between Google and Samsung and Apple to sell modern phones, is it really uniform on a triangle? Why do we believe that anyway?

So what people do instead is they've switched to this kind of utility spectrum, where in some sense what Hotelling was trying to do was fit two firms into a one-dimensional space. And fitting two firms into a one-dimensional space was a great idea to make it more tractable. Fitting three firms into a two-dimensional space, it just gets awkward. And so the observation is, why not just stick with an n-dimensional space for n firms instead of an n minus 1-dimensional space for n firms?

And so what people do is assume the utility that Firm I gets from by-- utility that consumer I gets buying from Firm J is vj, which is like the vertical quality of product j, minus alpha times pj, the price of Firm J, plus epsilon ij. That's my uniform taste for how much I like iPhones or how much I like Samsung phones or how much I like Google Pixel phones, OK?

And then people typically also actually with n firms will now typically go to n plus 1 dimensional uncertainty. And they'll add epsilon i0, which is how much do I like the outside good of not purchasing anything? So that could be how much sentimental value do I have for my current phone, which is often the outside good is just not buying anything and keeping the phone that I've already got.

And people can have different-- so in some sense, in that phone example, I've got a four-dimensional preference space. How much do I like new iphones? How much do I like new Samsung phones? How much do I like new Pixel phones? And how much do I like the old phone that's already in my pocket? And people can have different views about these things, both for taste-based reasons and like the one in my pocket is broken or not broken or whatever condition it's in, OK?

Demand in this model in general case is just an n-plus-1-dimensional integral to figure out. The demand for product j, when these are the prices, I just integrate over that n-plus-1-dimensional space. It's just the integral over the subset of that space, where product j gives higher utility than product k for all k different from j. And I'm just integrating the density of the joint distribution of types over that region.

So it's like here I'm integrating the density of types to find demand for good 1 in this two-dimensional space. I just integrate the density over this particular region of space. And so I'm still integrating over these regions of space that look kind of like that. Actually, the two-dimensional model here-- let me do two products and no outside good. It would just be like I've got epsilon i1 on this axis, epsilon i2 on this axis. And I would just be this region. So these people buy from good 1. These people buy good 2. It's just entirely is epsilon i2 greater than epsilon i1? So it's an easier integral to do, typically, than those ones with the things stuck into a lower dimensional space because you're just integrating over quadrants or half spaces.

Sometimes, in empirical papers, if people want to do this, they're just simulating a bunch of-- drawing a bunch of epsilons at random, figuring out how many of those consumers buy each product. That's how you do this, practically speaking. This model is tractable in the special case. When there's no outside goods, the v's are all equal, and the epsilons are i.i.d. with some density f.

In general, you're trying to solve a game, an n player game, you have to solve an n equations and n unknowns problem. But if you have looking for a symmetric equilibrium, then you're back to a one equation and one unknown problem. Because you have a one unknown equation and one unknown problem, I'm just trying to max-- you know, I'm solving the fact that when is there an equilibrium of a symmetric model where all the firms set price p star. That's going to be if p star belongs to the arg max over p of profit of Firm J when it sets price pj and the other firms set price p star.

So, sorry, if I make it symmetric so this profit doesn't depend on j, this is just one equation in one unknown that lets me find p star. It's just each firm is maximizing its profit, holding everybody else's profits at p star. And so, where asymmetric models are an n equation problem, one equation-- symmetric equilibria you can often find by just solving one equation and one unknown. And so people often do that to give insights.

So let's look at what is the-- and obviously, the pi of pj comma p star is going to be xj of pj comma p star times pj minus c. So your demand is just price minus cost times demand, like it was in all the other models.

OK, what does demand look like? So if I'm thinking about every other firm is charging price p or p star and I'm considering charging my price pj, when is someone going to buy from me? They're going to buy from me if their preference for me-- so they're going to buy from me if epsilon ij is at least the maximum k different from j epsilon ik plus pj minus p star.

If everyone else is charging p star and I charge pj, which is greater than p star, I have to be that much-- they have to have that much higher idiosyncratic preference for me than they have for anybody else. And if I charge a lower price, they can be this much lower than their best preference for everybody else and still buy from me.

So what you get is the demand for me is this is-- let's suppose that there-- suppose that this, I'm going to call theta, their greatest utility they get from buying from any other firm or greatest idiosyncratic taste for any other firm, then they're going to buy from me-- this is the probability that their type is bigger than theta plus pj minus p. So that's a probability that-- this is the probability that epsilon ij is bigger than theta plus this. And then, this is the density of theta.

So if you're choosing-- there are n other draws or n minus 1 other draws, there are n minus 1 draws from some distribution f, what's the probability that the largest of n minus 1 draws from f is in some tiny interval theta plus d theta? And the probability that this is-- it's an order statistic calculation. I don't know if the order statistics have disappeared from [14.]380.

But so, the density of the highest of n draws-- so the probability that someone is in the highest of n minus 1 draws in that interval is you pick some firm to be the highest, their value-- the epsilon for them has to be in this interval, which happens with probability f of theta and all other epsilons have to be less than that theta, which happens with probability capital F of theta to the n minus 2.

So this is the density for the highest of the n minus 1 rival draws. But so anyway, that's what my demand looks like, OK? If this is what my demand looks like, then again, I'm just going to have the first-order condition that I'm going to have to have. p star satisfies the first-order condition d pi j, dpj evaluated at pj equals p star equals 0. And so I'm going to get first-order condition. The first-order condition is just going to look like pj minus c times dxj dpj plus xj equals 0.

And so it turns out you can just-- maybe I should do that right. Again, so I have this first-order condition. It's going to look like p star minus c times dxj dpj evaluated at p star plus xj of p star equals 0. So that's what the first-order condition is going to look like, as it has always looked in this class.

And then the xj of p star is going to be 1 over n in equilibrium with n symmetric firms. So you just get p star minus c is just equal to minus 1 over n dxj dpj. One other thing I meant to mention before is look at this utility function, v minus alpha times p plus epsilon. Utilities are only defined up to an affine transformation. I think probably hear that in [14.]121 a lot. So you can also think of this as like v over alpha minus p plus 1 over alpha times epsilon ij.

So this model, when people put in a price coefficient alpha, is a lot like the Hotelling model, where one puts 1 over alpha where the t used to be because 1 over alpha is like the size of the transportation cost, size of the heterogeneity relative to having a unit coefficient on the price. So when people do empirical IO, they always put an alpha here. But you can think of it as fixing this as 1 and putting 1 over alpha in front of the epsilons.

So anyway, so what do you get as the markup when you do this? What you get as a markup is price. So this is price equals cost plus negative 1 over n 1 over dxj dpj. So anyway, the equilibrium formula is price equals c plus t times something. So it's c plus t times 1 over m of n, and the 1 over m of n is just what you get differentiating-- m of n is like what you get differentiating this thing with respect to pj.

And you can see it's got things like f to the n minus 2 and F and n's in it and stuff. But anyway, if you do the calculation, you differentiate on the integral sign, this is what you get. So price equals c plus 1 over m of n times 1 over alpha, where m of n depends on the assumed density of the distribution of the epsilons, OK? So some corollaries.

Suppose F is uniform 0, 1. So certainly, Hotelling could have done, instead of the beach, he could have done the people living on a square, where this was utility epsilon i1. This is epsilon i2, OK? It turns out that if Hotelling had done his model on a square, he would have found the neat result that price equals c plus 1/2t. He would have gotten exactly the same thing except with 1/2 in it.

So with the linear model, you can extend this to people living on a hypercube-- or cube or a hypercube or whatever. And the formula is exactly correct, that price equals c plus 1 over n times t. So the thing that everyone loved about Cournot is true of this hypercube model of product differentiation. Price is just c plus t, where t is the length of a side of the cube. How intense is your preference for Coke versus non-Coke or whatever? And things decline like 1 over n if there are n firms in the model.

What can we say more generally going outside the uniform distribution? Suppose that uniform distribution is bounded above. Suppose that the epsilon is bounded above or that the tails are thin. Tails are thin like the normal distribution in this sense of limit as epsilon goes to infinity F prime over F goes to negative infinity.

Then, as the number of firms goes up, prices drop to cost. So if we take bounded distributions or we take distributions that are not bounded but have very thin tails, then you do get the result that competition drives prices to 0. Does that happen for all distributions? The answer is no. Most important example where that doesn't happen is the logit error distribution, which people often assume because it makes algebra easier.

But it turns out for the logit model-- so this is the logit cdf, where gamma is the Euler-Mascheroni constant, like 0.577. So in the logit model, what you actually get is price equals cost plus some constant times 1 plus 1 over n minus 1 times 1 over alpha. So as n goes to infinity, p star is going to c plus 1 over alpha. So in the logit model, prices--

AUDIENCE: [INAUDIBLE]

GLENN ELLISON: What?

AUDIENCE: k over alpha?

GLENN ELLISON: Oh, yes, k over alpha. Thank you, where k is some--

AUDIENCE: [INAUDIBLE]

GLENN ELLISON: k, it's--

AUDIENCE: It's some arbitrary--

GLENN ELLISON: It's some constant. I forget exactly. You can think of it as like 0.2378. I forget what the number is, but it's just a number.

[LAUGHTER]

So in the logit model, you can compute what it is. Prices are just like c plus-- as the number of firms goes to infinity, competition doesn't eliminate markups. And how is this plausible? The idea is that these models have a somewhat unreasonable assumption that I snuck over on you, which was that, as I said, in two dimensions, it's like, how much do I like iphones versus do I like Samsung phones? It seems like, OK, it's reasonable to assume there are people with those two separate differences of preferences.

But what I snuck over on you is that-- here I said it-- epsilons are i.i.d., independent and identically distributed. The more products you have, it gets harder and harder to say it's really plausible that my taste for a BMW and a Mercedes and a whatever are all completely independent of each other. As the number of firms gets large, the products get awfully similar to each other. So if you're a person who likes Mercedes styling, you probably also like BMW styling. And if you're one who likes Toyota Lexus styling, maybe you also like Acura styling, or whatever.

So, in some sense, what the logit model is assuming, if you assume we keep having this-- every new product has this completely new idiosyncratic reaction to it, suppose there are a million products out there and you tell me, OK, yeah, you're one of a million products in the market. But this person likes your product better than every other product. Well, to think about-- the markups depend on what this dxj dpj. What do I conclude when I've just been told they like my product more than everybody else's product? I'm like, wow. They like my product a lot, OK?

Now, in the hypercube model, if the epsilon ij is distributed uniform 0, 1, I'm like, they like me best out of a million things. It must be my-- utility for me is like 0.999999 something or other, or they wouldn't have been buying from me. But while they like me this much, there's also somebody else who they like 0.999998 or something. And so I think the amount by which they like my product over anybody else's product is tiny. Therefore, my demand is very price sensitive. And if I don't price really low, they're going to switch. So I have this incredibly price-sensitive demand.

But if you have a-- maybe you know this. If you have an exponential distribution and you're told that you're the largest of n exponentials, how far are you above the next highest draw? The answer is 1. That's a property of the exponential distribution, that if you're the highest of n draws from a standard exponential, your posterior for their value for you is n. And your posterior for the product-- their value for the next highest one has high expectation n minus 1. So you think you have one of space to price relative to your rival.

And so with an exponential distribution, we actually-- you get-- exponential distribution, you're going to get like p star equals c plus t doesn't depend on n at all. The more exponential draws that are out there, they're just, yeah, they probably like me a lot more than somebody else. They like me a million. They like the next guy 999,999.

Logit isn't quite like the exponential, but it turns out to be just very similar to the exponential and that approximately-- it ends up with an approximately constant markup. And then markup depends on what weight you stick in front of the epsilon ij in the logit model, the 1 over alpha.

And there are distributions for which-- if you take a really fat-tailed distribution, one that's fatter tailed than the exponential, prices can go up as n goes to infinity. And I don't know that we think that's often realistic. Where do these-- you keep getting these independent tastes from? But in these models, it's there, and you should be aware of it, OK? Questions on that?

So, OK, I think that's what I had to say. Yeah, so I think I'm going to skip this slide. One thing I wanted to say in this slide is suppose a student comes to me and says, I want to write a paper about mergers and the effect mergers are having on markups in industry x. This is a know-your-theory cautionary tale. If a student says, OK, I'm going to estimate logit demand in this model and see what's happened when you merge from five firms down to four firms if phone maker A buys phone maker B, or whatever.

So you can estimate your logit model and say, I'm going to allow the price sensitivity to vary and have it be v minus alpha p plus epsilon ij utility. And I'm going to estimate the alpha and see how price sensitive things are. And then when I find out how price sensitive things are, see what the merger is going to do. If you then use that model to say when you go from five to four firms, how do markups go up? The answer is markups go up by 1.067. With five symmetric firms, the markup is c plus k over thing times 5/4. When you go to four firms, it's c plus k over alpha times 4/3.

So on the question of how do markups change when you go from four firms to three firms-- five firms to four firms, the answer is-- that question has an answer that's independent of all the parameters you estimate in the model. And the answer is just this. And so if you do want to estimate how mergers would affect markups in models with horizontal differentiation, you have to make sure you're correctly estimating this object dxj dpi versus dxi dpi. And the models, like the logit model, can have built-in assumptions about what those things are. And if these things both just have an alpha in them and those alphas cancel, you're not actually estimating what you think you're estimating. But I didn't want to take time on that.

So going to more modern pictures, I stuck Jidong's picture on this slide. So discuss a recent paper discussing the effects of bundling on competition. So the motivation for this paper is there are many firms where there is this antitrust concern about firms bundling and selling their products together instead of selling their products separately. So firms, for instance, if you're-- maybe you think they don't really matter. But Apple is selling phones and they're selling earbuds.

And you can think of, I can design the phones and the earbuds so they only work together and have to be bought as a bundle. Or I can design the phones and earbuds so they work separately. And I can be competing with Samsung separately in the phone market and separately in the earbud market. With operating systems, this often comes up as what features do I build into an operating system so that people have to buy the bundle of the operating system containing these 37 different components, all from me? Or do I allow plug-and-play compatibility so we have separate competition product by product? OK?

OK, and so this is a potential antitrust remedy and industries with large market power is forcing them to be compatible and not use proprietary connectors or whatever to extend their monopolies. And so Joe's paper is about this question of, when will markups-- will bundling products together make markups higher? Will it make markups lower? How will it affect consumer utility?

So in his model, you have these n firms, and the n firms all sell m distinct products. And he looks at utilities of the following form. You get utility v minus pjm plus epsilon ijm if you buy product j-- if you buy-- sorry-- product m from Firm J. And he compares two separate models. One model is where the firms compete separately in this whole matrix of prices. So Firm 1 sets prices p11, p12, up to p1m. And then Firm 2 sets prices p21 up to p2m. And then Firm 3-- I'll just have three firms-- post its bundle of prices p3m.

He compares that to a bundled system where the firms just design their products to be all mutually incompatible. So Firm 1 just sets a price capital p1. Firm 2 sets of price, capital P2. And Firm 3 sets of price capital P3 for the bundle. So this is the bundle of getting all of these goods from me. This is the bundle of getting all of these things from Firm 2. And this is the bundle of getting all of these things from Firm 3. And he thinks about which results in greater consumer welfare and greater profits, having unbundled pricing or having bundled pricing?

OK, first observation is when do consumers prefer buying bundle 1 from bundle 2? Well, that's if m times v minus the price of the bundle plus the sum across all products within the bundle of epsilon ijm is better than buying mv minus the price of bundle k times the sum over all things in the bundle of epsilon ikm.

And again, normalizing this by dividing by the number of firms-- so if I convert this to-- this is m products. So p1 over m I'm going to think of as the per product price. This is what's going to be equivalent to p11 or something. How much are we charging per product in the bundle?

So I like this one better if and only if v minus 1 over mpj plus 1 over m the sum of epsilon ijm is bigger than this. So basically, bundling is like converting the model where you have utilities epsilon i1 up to epsilon n1 into a model where you have 1 over m sum of epsilon i1 and 1 over m sum of epsilon im, right? You're just basically comparing the model with utilities from the original-- random draws from the original distribution to the model with utilities that are random draws from this distribution.

So that's what I do when I bundle, is I just change the distribution of idiosyncratic consumer preferences and make it an average of these. And obviously, the central limit theorem is going to tell us that as m goes to infinity, these things are going to have smaller-- this is going to have smaller variance than that. And it's going to look approximately normal as m goes to infinity, but normal with a smaller variance than the first one had.

OK, so let's think about whether prices are higher. Again, same price, first-order condition I've done over and over again. So the first-order condition for price competition is pj minus c times minus dxj dpj equals xj. And again, if I'm looking for a symmetric equilibrium of this model, this is going to be p star minus c times minus dxj dpj equals 1 over n.

And what is dxj dpj going to be? dxj dpj is just the density of consumers who have utility for me equal to the maximum of their idiosyncratic reaction to me equal to the maximum of their idiosyncratic reaction to everybody else. And so, fixing n, this first-order condition-- where-- I don't have it written down anymore. So my first-order condition was p star equals c minus 1 over n times 1 over dxj dpj. All right, let me do this, OK? That's what my first-order conditions always keep looking like.

So equilibrium prices are low if this density is high. And equilibrium prices are high, if this density is low. So in this dxj dpj is the density of the set of consumers who have my epsilon just about everybody else's epsilon. Notice that the distribution of sum of the epsilons, it's more concentrated about its mean. It's thinner in the tails.

OK, so consumer surplus in this model depends on two main welfare effects for consumers. One is price levels. Consumers like low prices. And there's less deadweight loss when prices are low. The other thing that welfare depends on is match quality. Consumers are better if they're getting their favorite phone and their favorite earbuds and their favorite charger than if they're getting their favorite of one and their second favorite of the other and their third favorite of another.

So the match quality is are consumers getting what's the best product for them on every subcomponent? And then there's the price level thing that they care about. OK, so what's the result? So first result in the paper is that bundling reduces per good prices when n equals 2. So if you have two firms and they bundle their goods together, equilibrium per product prices go down. Noting prices go down, meaning that the firms make less profit.

And why is that? The answer is that density of 1 over m epsilon im becomes more concentrated. When n equals 2, what we have is just what's the probability that ui1 minus ui2 is less than delta? And that's just the probability that we're close to the 45-degree line. That is that my epsilon for Firm 1 is within epsilon of my-- I'm sorry, within, let's say, this small delta. What's the probability that my preference for Firm 1 and my preference for Firm 2 are within delta?

Well, if you take the distribution of tastes that used to be uniform and you sort of-- like in the uniform distribution, this thing is not all that big. But then you add n goods, and you make this distribution look like this. Almost everybody has a preference of exactly 1/2 for both products. Then, once they have a utility of exactly 1/2 for both products, they have a huge probability they're within some tiny delta of each other. Therefore, price competition becomes very intense. So when n equals two, it turns out that whatever m is, bundling just always makes the prices lower. So, in some sense, firms should not want to bundle in a two-firm model.

What's noteworthy, though, is while the firms don't want to bundle, the consumers also don't want the firms to bundle. So the prices go down, but the consumers are worse off. And that's because the worst match quality overwhelms the price drop. So there's a price drop, but it also makes the consumers worse off because they get worse match quality. And it doesn't compensate-- and the price drop isn't big enough to compensate them for the drop in match quality.

Now, when n equals 2, That's. True. When n is large, when n is above some threshold, bundling starts to increase profits. And why is that? Well, again, the value-- this sum of epsilon ijm starts getting to be a very tight distribution. And so, what's the density of consumers who have very close values for the best and second best thing? That's a density measured out here. If you like me best, what's the probability you like someone else? What's the probability the other person is close to you?

When n gets large, holding m fixed, letting n go large, looking further out in that distribution, it gets thinner tailed. And the dxj dpj is small. And therefore, the p star is high. So bundling is good when the number of firms is large. And a third result is when you fix the number of firms and the number of goods goes to infinity, then eventually bundling drives prices to 0 as m is going to infinity.

So I think in some ways I wanted to do it because I wanted to cover some recent paper. And it's a nice example of how that basic Perloff-Salop intuition of price is about differentiation. And it's about dxj dpj is a very good way to think about a lot-- it's a good way to think about a lot of issues, bundling included.

AUDIENCE: I don't understand why 4 is a result.

GLENN ELLISON: So--

AUDIENCE: Because we're just putting infinity in the denominator. Would they ever charge an infinite price?

GLENN ELLISON: Sorry, I should have said-- sorry, I should have said this is-- that's with c equals 0. So it's going to be that-- I think I had a c equals 0 assumption. So it's the limit as-- it probably should have written the limit as n goes to infinity pb over m minus c goes to 0. I think maybe I had c equals 0 throughout the paper. Yeah. OK?

OK, let me end there then. I didn't think I would get to these last slides, but-- so anyway, on Monday, what I'm going to do is discuss oligopoly price discrimination, where you have oligopolies who are competing to sell multiple products, and therefore, are going to be price discriminating selling multiple products. And then I'm going to do two empirical papers, Bresnahan and Miller-Weinberg. Bresnahan, I took the star off it this year because Bresnahan, it's original. It's a real classic.

It really looks 40 years old when you read it. And you're like, wow, I can't believe you do things that way. So I'll discuss the ideas in Bresnahan not going into the details, which now look very old fashioned. And then I'll discuss Miller-Weinberg. And I'm doing less empirical work here because Tobias is talking about demand estimation here. And an awful lot of demand estimation is really about estimating static models of product competition with those epsilon ij kinds of models that I talked about.