[SQUEAKING] [RUSTLING] [CLICKING]

GLENN ELLISON: OK, I guess I will start the tradition I started last time of starting at 9:04. So if you remember last time, I finished a lecture right in the middle of an example. I was talking about monopoly information design, which is about how monopolists can affect-- by changing the information that consumers receive about their product can affect the prices. And this is motivation, both for firms who are giving information about their product to consumers to try to maximize profits, or potentially for platforms like Amazon giving information to consumers to try to increase the consumer surplus and limit the monopoly power that the monopolist is going to exert.

I had been multiplying everything by 18 when I talked about the examples last class in my head. So I made this about the Sharan sandwich at Clover and said that if consumers had full information about the sandwich, their values would be uniform. On 0, 18, there's a cost of $6 of producing the sandwich. And I discussed a few different-- how consumer surplus and profit would change a few different information structures.

If you give the consumers full information, then their values are uniform 0, 18. When your cost is 6, the monopoly profit is halfway in between 6 and 18, which is 12.

Another example I discussed is suppose Clover puts the sandwich in a bag and tells you nothing other than it's a new sandwich. We're not telling you anything about it. Then all consumers would have values of 9. And so the monopoly price would be to set a price of 9. And in this case, they actually make more money, but consumer surplus is completely eliminated.

And then I gave this partial information example, where I gave the consumers partial information about the product, then ended up with a monopoly price of 10. And consumers and the firm were both better off than with full information. And this is what that example looked like.

I know people often ask me, How on earth could a firm do this in practice? And let's imagine-- you can think of this two ways. One is, is this a model of an omniscient firm like Amazon or Google or Apple that knows so much about you, they know your preferences better than you do because they know what people like you thought of the product? So they can actually tell you about what your value is going to be if you eat that sandwich.

The other example could be, let's imagine that the value for having a sandwich depends on how much you like cauliflower and how much you like the Indo-Chinese sauce that it comes in. And initially, you know nothing about the sandwich. But then one thing they can tell you is, I'm just telling you that the sandwich is made of cauliflower as the base.

And then people who don't like cauliflower are going to have values between 0 and 6, depending on how much they like the sauce. Those who have a medium preference for cauliflower would have these values. Those who have a high value for cauliflower would have those values. So it's like, we're not going to tell you what the sauce is like or the vegetables, but we'll tell you that cauliflower is the base of the sandwich.

Anyway, but what happened with this example was all the consumers, after getting the signal, had values of 3, 10, and 16. And that ended up making the monopoly price 10. So anyway, the monopoly information design literature is about how can we think more generally about what types of information we want to give consumers to either maximize profits or maximize consumer surplus or do something in between?

So the Roesler and Szentes paper that I put on the reading list assumes that the values are now distributed according to some distribution F on 0, 1. It specializes that the cost is 0. I'm going to write mu for the mean value of the consumers.

And what I'm going to do is give this firm tremendous ability to manipulate the signal distribution. What the firm can do is choose signals to choose any joint distribution whatsoever on v and s. So the most general set of signal structures is they're just some joint distribution on v and s.

Notice that when consumers get signal s, what they're going to be willing to pay is expectation of v given s. So rather than giving them a signal with a label like low, medium, high, I might as well just say, your signal is either 3 or 10 or 16. Given the information about the product I'm revealing, your expected value is 3, your expected value is 10, your expected value is 16. And so you might as well-- consumers are going to do the Bayesian updating of whatever letters or words you give them. So you might as well just do the direct signal structure, just tell them what their expected value is given what you've known.

Now, obviously, I can't pick any distribution over values of s I'm going to give them. v is going to be s plus epsilon, where expectation of epsilon given s is 0. And so that means two things. One thing that it means is that the expectation of the signals must be mu because the expectation of v is expectation of s plus epsilon given s-- iterated expectations. I guess you must've done [14.]380 already. And so the expectation of the signal distribution has to be the same as the expectation of the prior distribution.

And the other thing is that F is going to be the distribution of v, the true values, is a mean-preserving spread of the distribution of g. And so what you want to think about the monopolist is doing is just choosing a distribution g over possible signals, subject to those two constraints-- I've got to choose a distribution that has mean mu, and I have to choose a distribution such that F is a mean-preserving spread of the distribution that I pick.

So, for instance, if this is the true distribution, I can't give signals like this because this distribution of finite support is not a mean-preserving spread of that bigger distribution. I could choose something like this because this thing is a mean-preserving spread of that.

So let me say, the two problems here, one is the profit maximization problem. The other is the consumer surplus maximization problem. Profit maximization is trivial. It's economically probably more important.

No one ever writes about it because what do you do to maximize profits? You just tell consumers your value is greater than c or your value is less than c. And then if you tell them your value is greater than c, you sell it to them at the price of p star, which is just the expectation of v, given that v is greater than c.

So this simple pricing scheme, just tell them your value is above c or below c, when you hear your value is above c, every consumer who hears their value is above c is willing to pay at this price. So we get the fully efficient outcome, and we fully extract all the consumer surplus. So that's clearly the best thing you could do because you're maximizing social surplus, and the monopolist gets all of it.

What do we do if, instead, we're Amazon, and we're trying to maximize consumer surplus? Intuitively, what you're trying to do is get the monopolist to charge the lowest price it can and sell to the largest number of consumers that it can, because that's the way you generate consumer surplus is get everyone with a value greater than c to buy and do it at the lowest possible price. Could potentially be some trade-off between those two objectives. It turns out it's not. In this problem, you want to get everybody to buy.

So what do you do? The answer is this. I'm going to use some distribution that looks like that. And so the two-step argument to show that this is optimal is I'm going to first look at distributions of this form. I'm going to show that they work very well. I'm going to show what the maximum profit is I can get with the distribution of this form. And then I'm going to show that no other distribution can do better than this one. So it's like a guess and check answer, I guess.

So what is that distribution formally? So I let the cdf-- I guess I graphed the pdf here. So the cdf looks like this. It starts at some value, s lower bar, goes flat like that, goes like that, and then has a jump and then goes across at 1. So this is s upper bar. So this is my graph of G of s. So this is the cumulative distribution function that I'm choosing the signals from.

Notice that if the platform chooses this distribution, and the monopolist then picks any price p in this interval, the profit that the monopolist gets is p times 1 minus G of p. And p times 1 minus G of p is just p times s lower bar over p equals s lower bar.

So what this distribution does is I choose that distribution, and I get this graph of s lower bar, s upper bar. I get this distribution of profits. So profits are if I charge a price of anything less than s lower bar, everybody buys. And so my profit is linear in the price that I charge. So I'm putting price on this axis.

Then here, everybody buys, and I get a price of-- I get a profit of exactly s lower bar. And then I've made people-- I've made you indifferent over all prices in this range. And then my profit drops to 0. So this, I'm graphing the profit that you get charging p.

So given that the monopolist is indifferent over all prices in this interval, one of the optimal prices is s lower bar. And so by setting this distribution, and then the monopolist is indifferent over all prices, it can charge price s lower bar. We observe we get profit s lower bar. And then the consumer surplus is the difference between each of these values and s lower bar.

So if you integrate under this distribution, you're going to get what the consumer surplus is. And so by putting s lower bar down here, the lower I can put it, the more consumer surplus I'm generating.

So within this class of G's, the consumer surplus optimal solution is clear. What I want to do is I just want to pick the smallest s lower bar that I can subject to the constraint that G has to have mean mu and F has to be a mean-preserving spread of G. How low can I go and still do that? Well, this distribution has the property that because it's only going down like 1/s, if I were to set s upper bar going towards infinity, the mean of this distribution would go to infinity.

So for any s lower bar, I can just pick a high enough s lower bar and make the mean mu. But the problem I run into is I need this distribution to be a mean-preserving spread of F. You guys-- I don't know. Do they still even do this in statistics? What does a mean-preserving spread look like?

So If this is G, and this is F, G is a mean-preserving spread of F. And it's a mean-preserving spread of F because it crosses the cdf of F once from below. If F looked like this, this thing would not be valid because-- oh, no, that would still be OK.

The problem is going to come in if let's suppose that F were to go like this, then it never crosses. And so if I'm having to use an s upper bar that's bigger than the upper bound of the F distribution, then this thing isn't a mean-preserving spread of F anymore. And so what I'm going to want to do is just pick an s lower bar and pick the smallest s lower bar such that when I put the upper bound below the upper bound 1, I'm getting a mean of mu. So anyway, so that's the best we can do within that class of distributions.

And then step 2, the harder step of the proof-- actually, I think I've done it pretty well on the slides if you want to look at the slides at home. But the basic idea in the slides is that no other distribution could possibly do better than this one.

And how does that argument go? Well, what I'm going to do is I'm going to say, let G be any other valid choice. If you give me G is a valid choice, I'm going to pick a distribution in this class and show that it does at least as well as G and is valid. So you give me any other valid distribution, let pi be the profit that you earn when you use G, I'm going to consider the distribution G of pi s upper bar where s upper bar is chosen such that s upper bar gives this thing a mean of mu.

This distribution, G of pi s upper bar, it also gives-- it gives the same profit as in G, because these distributions G of s lower bar s upper bar give profit s lower bar. So if I choose the profit I get from using this, this thing also gives the same profit pi. It gives the same profit pi. it's fully socially efficient. Everybody purchases in that distribution, so it has to give at least as much social surplus. Because it's maximizing social welfare, it has the same profit, consumer surplus has to be at least as high.

So if I switch from G to G of pi s upper bar, then I've gotten at least as much consumer surplus to complete the proof. I just then need to show that G of pi s upper bar is a valid distribution to choose. And to do that, I just need to show that F is a mean-preserving spread of G of pi s.

To do that-- I'm trying to show that F is a mean-preserving spread of G pi s upper bar. I do that in two steps. I show that-- we know that F is a mean-preserving spread of G because G was valid. And so I just need to show that any G is a mean-preserving spread of G of pi s upper bar. And the way I'm going to do that is show that any G looks like this. Any G that's valid has to start above G of pi s upper bar here-- has to start above it, and then it has to end below it.

Anyway, I've got the argument here. But basically, the argument is it has to be above it here because this one is 0. In this interval, it has to be above it here because this one, the maximized profit, is at most pi. This one, the maximized profit is exactly equal to pi. So you have to sell to fewer consumers. So it's above it all the way to here, and then it has to be below it here because that one's 1. That's the basic argument.

So it's a-- what are my takeaways-- let me get my takeaways first. Takeaways, profit-maximizing structure is simple. You just tell people that their value is at least c. A good way to think about information design is think about choosing the distribution of the consumer's posterior. And if you want to do that-- if you want to minimize-- if you want to maximize consumer surplus, what you want to do is get everyone to buy, and get everyone to buy at the lowest price possible.

The way you get people to buy at the lowest price possible is to use a distribution like this where there's a lot of low-value consumers. And if you raise your price at all, you lose a lot of demand. And demand is very steep here, so that gets the monopolist to be indifferent between charging this price and charging those higher prices.

And so that's the kind of thing you want to do with the information structure is create a lot of low-value consumers and create a distribution where every time the monopolist raises their price a little bit, they would be losing lots and lots of consumers. And distributions like this are the way you maximize consumer surplus.

Also, Paolo had actually asked after the last class, What's the full set of things we can do, and how do I show that? I added a slide that I won't talk about. But Roesler and Szentes just have this fairly simple argument that shows that this is the profit-maximizing point. This is the consumer surplus maximizing point. Everything in this triangle is possible. So if you want to achieve less efficiency, you can achieve less efficiency.

But basically, what you can do is give them-- you have to give a monopolist at least this much profit. If you were trying to hurt the monopolist, you can't do any better. And consumer surplus-- anyway, it's going to have to be in that triangle.

I really think it's probably mostly this diagonal that we're typically interested in, how do we maximize consumer surplus versus profit. And the way we do this is we choose these distributions G of s lower bar, s upper bar. And we just vary s lower bar. And the higher we make s lower bar, the higher is the monopolist's profit, the lower is consumer surplus.

In some sense, this is a special case of this argument where you just set s lower bar equals expectation of v, given that v is greater than c. That's the monopoly solution. That's the monopoly pricing solution. And so as we raise s upper bar to that point, that's what we're going to do. Any questions?

Let me say, I feel this is-- just in terms of thinking about topics, I don't really know. I looked. I couldn't find any good empirical papers on monopoly information design. It seems like that's something that I think there should be some empirical work on. And also, this question of translating this from theory to IO is this "I can choose any joint distribution over v and s that's possible" is a heroic assumption.

And I guess the question-- intriguing empirical question is, how can one do that in practice, and how close can one get to designing expectation of the distributions over expectation of v, given the signals, with a finite set of signals or signals about product characteristics or a more limited set of ways I can give signals than just this telling people their posteriors immediately?

Anyway, so rest of today's class, I'm going to start doing empirical papers for the first time. And I'm going to mostly talk about two papers. What I thought I would do first, though, is go back and give you the 8-minute history of empirical IO. I mean, I think the background is useful to understand. As you start reading IO papers, there are a lot of IO papers written about ketchup or cardboard boxes, and people often ask me why on earth do IO economists do that. And so I think giving some perspective on what IO people do and why could be useful.

Let me start my brief empirical history. So early empirical IO-- so before the 1960s, IO economists wrote books. And they wrote big, thick books. They wrote books-- an IO economist was an expert in some industry like the airline industry or the car industry or the coal industry or the steel industry or whatever. But IO economists were experts in industries. They wrote these big, thick books discussing lots of stuff about the industry.

We actually used to have an assignment in this course, which was try to get a Harvard library card, go into the stacks of Widener, pick out one of these many books, read it, and write something about it, because it was often they're full of facts and full of interesting things going on in the industry. And they often have lots of puzzles about what's going on and why is this happening when, with modern theory, you're like, OK, I could tell you why that's happening, or this is a good example of search affecting pricing or whatever, some kind of incomplete information story.

It was a very interesting field, in the sense that you would read books about these industries, and you would learn a tremendous amount about them. They were typically very light on theory. It was really a lot of descriptive work about how does this industry work, what are the firms, what are the prices like, what are the markups, what are all the practices that you see there.

So in the '60s and '70s, there was this increasing use of statistics in economics. And instead of just writing discussions of industries, people wanted to have regressions in their papers. And so regression started to come into the profession in the 1950s. I have this one anecdote here that my wife got from Zvi Griliches, who was a famous economist at Harvard, that his summer job at some point in the 1950s was to run a regression.

And you might ask, well, OK, was your summer job to run regression? It's like, no, he was given the data set, and his job was to compute x prime x and then invert it. And it was, like, a 7-variable regression or something like that. So his summer job was with a calculator, like, here's a 7 by 7 matrix. I want you to invert x prime x so that we can then report the results of that regression.

But anyway, obviously, this got easier in the '50s and '60s when computers got invented. And Joe Bain, who was actually a first one to do regressions in a within-industry study, had a view that IO shouldn't be developing these facts about the coal industry or the steel industry. We want broad truths that apply to all industries.

And so it became popular. And IO after that was cross-section regressions. And what happened in these cross-section regressions was people typically have a dependent variable like markups in industry i. And they would regress markup on industry i.

So it would be like markup in industry i is alpha 0 plus alpha 1 times the four-firm concentration ratio plus alpha 2 times the advertising to sales ratio, and there'd be a bunch of other variables. But you would just do-- I'm running a regression like this with six or seven or eight variables, and I'm going to try to talk about what is it that makes industries competitive. And they were looking for these broader truths that would give us-- instead of having to study industry by industry, what determines profit levels in general?

Anyway, so regressions here, there's a data set the Census of Manufacturers-- Census of Manufacturers, it's a census of manufacturing firms in the United States. So it has data on every individual firm in the United States, and they send them a big, long survey with hundreds of questions. If you ever want to write papers using the Census of Manufacturers, you can access the Census of Manufacturers individual data from the NBER, from a small, windowless, unpleasant room at the NBER. But there is a windowless room where you can access the firm-level data on every manufacturing firm in the US.

People in that era didn't really access the census microdata. They just got the published census data of means of these variables for each of the 153-digit manufacturing industries or 454-digit manufacturing industries. They'd run these regressions, and they'd talk about what determined-- how many firms did you need for a markup to be competitive, or how did competitiveness vary with the number of firms in the industry?

This was popular in the '60s and '70s, and then it started to become-- a number of critiques of this approach became very well accepted. One critique was that just all these studies are leaving out very important features of the industry. So for instance, if I'm going to study the steel industry, I can't talk about prices in the steel industry or profits in the steel industry without having price of iron ore on the right side. I can't think about prices in the steel industry without price of coke and coal on the right side. I can't think about steel industry without strength of union labor in the Midwestern United States on the right side. So for every industry I had, I would think of three or four or five things that just had to be on the right side to talk about that industry.

But if you've got 150 industries, and each industry brings in three or four or five unique things to put on the right side, and then you put them all in on the right side, you end up with a regression with 750 things on the right side and 150 observations. And obviously, now machine learning, people will tell you that's totally fine. 750 explanatory variables, 150 observations, it all works fine. But at least in the 1980s, that did not work. You couldn't run regressions like that.

The other problem that came to be accepted is that the census data, like when you're studying markups, markup in industry i was really price in industry i minus the cost of production in industry i. Or maybe it was price minus cost over price.

But you would need to use this data on what's the cost of production. And the cost of production that you get when managers put in for the cost what their accountant told them the cost is, is very different from what you would get if an economist asked what the cost is. For an economist, you really want the marginal cost, not the average cost.

But then even though when you think about-- even getting the marginal versus average, accounting costs of all these conventions in of how do you treat the cost of capital, and do you amortize it over time, and you put in a depreciation adjustment. And what do you do when the firm spends a lot of money on advertising? Is that a one-period cost, or is that a flow cost that depreciates over the number of times that the goodwill stock decays?

And so the costs-- a lot of the elements of the cost, how do you put the firm overhead and all the people who work in the management office or whatever? The costs have a lot of assumptions that go into defining them. And so the difference between the accounting cost and the true cost is large.

And it's probably systematically related to things like whether the firm is doing a lot of advertising or not, whether it's doing a lot of R&D, how old the product is, whether it's produced in big factories or little factories. So all the stuff that you're putting on the right side is correlated with the measurement error that you want. And so then are we really just finding correlations with measurement error, or are we finding correlations with actual firm conduct?

And so that was a big issue. Endogeneity is also a big issue. You're putting the number of firms on the right side. Something you get taught over and over again in classes like this is, well, if you have a business that's highly profitable, then more firms are going to enter, and that's going to change the profits. So this four-firm concentration ratio is also going to be endogenous, or the number of firms on the right side is going to be endogenous. There really wasn't much in the way of IV estimation in the structure conduct performance literature.

And then there was also interest in IO theory. This led to what was called the NEIO revolution. So NEIO stands for New Empirical Industrial Organization. I guess lots of fields run into this problem. You have the sort of new something, and then there's something comes after that. Like you have the new trade theory, and then you have to have the new, new trade theory or whatever, because-- anyway, so in the 19 early '80s, people started to call this the new empirical industrial organization. It's no longer very new.

But what happened in NEIO-- so one is, it went back to the '60s. It went back to people writing papers about single industries. Everyone thinks that there are a lot of specific things about every industry that matters, so we need to study it one industry at a time and control for all those factors that should be on the right side.

There was an emphasis on using what was well measured. So it's sort of funny. In macro, they think cost is well measured, and price is hard to measure, so they always want to use cost data. In IO, people think that price is well measured, cost is hard to measure, so we always use price data. So there's this emphasis on using quantities, prices, number of firms, market shares, those things as the explanatory variables-- as the variables you use rather than accounting variables.

Great deal of attention paid now to identification and endogeneity. And theory also came to take on a big role in empirical IO. And it's often that the-- two different reasons. One is that the empirical work is often designed to improve our understanding of some theoretical model, which were being recently developed in the '80s. And then the second is because even if you're not trying to comment on the theory, the theory is going to give you thoughts about what the functional form should be or what you should be regressing on what. And the papers are often designed to test some particular-- or enlightened by some particular theory when they're setting up what analyses they're doing.

More recently-- so at the beginning of the NEIO revolution, people would sometimes describe work as structural. And they would use the word "structural" in a way that's different from what structural means today. So in the '80s, structural really meant instead of estimating equilibrium objects like understanding price as a function of things, you're understanding supply as a function of things and demand as a function of things. And the underlying equations that determine the equilibrium instead of analyzing the equilibrium.

I think, over time, structural work has increasingly come to be used to describe papers that assume that some theoretical model is correct and then focus on estimating the parameters of that theoretical model. And the word "structural" is also increasingly used to talk about work that goes all the way to estimating the primitives. In some ways, every industry, the outcomes depend on the production functions of the firms, the utilities of the consumers, and then something about the strategies that are available to the firms. And so structural work increasingly is trying to go back to those primitives. Instead of just estimating supply and demand, you're trying to estimate distributions of consumer utility functions.

Something that's very popular in IO papers is that you-- a reason for getting all the way to the primitives is then you have these counterfactual simulations where you say, I observe this industry in this state. What would happen if there was entry in that industry? What would happen if we broke up these firms? What would happen if we put in a price cap regulation?

Once you know these primitives, you can then talk about here are all these alternate policies we might implement, and then what would that do. And in some sense, this is what's really become very popular in many fields like education and development, where you estimate a fundamental model, and then you talk about I did a trial of this one, I did a randomized trial of this policy. I can also do an empirical estimate of what other policies would have done because I've uncovered the primitives.

And I would say what's also happening at the same time-- while there's a lot of good structural work, over time, structural-empirical IO has become less and less connected to theory. And in some sense, structural-empirical IO models these days tend to use structures that previous structural-empirical papers have used rather than the structures of the theoretical models that theorists would use to describe the same situation.

And part of that is because repeated games really don't work very well for empirical work. Incomplete information doesn't really work very well for empirical work. And so empirical work tends to have these models that have this very limited sense of dynamic competition and of incomplete information and other factors because it's not what you can estimate well structurally-- although, obviously, there's a lot of work trying to do that better.

So as far as how do you come up with a topic for an empirical IO paper, people do multiple different things. One thing that you can always do, just like in any other applied field, is just come up with some applied or policy question and say, what can I teach people about that.

A second thing one can do is improve understanding of some theoretical model. There's recent work on this monopoly information design. Can I do some empirical work to understand, does information design work in practice, or how well does it work or does it not work?

The other third thing you can do is also do an improvement on existing empirical technique. If there are many previous papers trying to estimate empirical models of auctions, I have some better way to estimate empirical model of an auction. And you try to find a data set to improve on that empirical technique.

How do you find your topic? Obviously, if you have an applied question, you look for the data to find it. If you're trying to do something improving understanding of a theoretical model, this is where we get papers on ketchup and cardboard. It's like, it's very complicated to estimate anything in the mobile phone environment or the computer environment because there's so many different firms doing so many different things.

If you're trying to study some topic and you're like, I could study this in the cardboard box industry, the cardboard box industry is an incredibly simple industry. Basically, you take paper or you take leftover waste from wood mills, you mush it together, you add water, you dry it. There's just, like, three ingredients. There's nothing much there. You have shipping costs. You can estimate what it costs to ship cardboard from point A to point B.

If you find an industry that strips away all the complicated stuff going on, there aren't agreements with Google doing this for some other reason. It's just cardboard. So sometimes you can pick a simple industry in order to study a complicated model because you want to isolate that theoretical factor that you're interested in.

Third thing you can always do is find data. Data is always in short supply. IO papers tend to collect their own data. So you have a friend who works at a startup and is willing to give you individual-level consumer data on something or other. I tend to think, you look at all the topics in this course-- I'm going to talk about monopoly, I'm going to talk about price discrimination, I'm going to talk about entry, and I'm going to talk about search costs. And there must be some of those things that are involved in your friend's business.

So if you can get the really, really rich microdata that no one has, there must be some interesting comment about that industry that you can make. There must be something that that data set is good for studying. And so that's another way to think about paper topics.

So I would say that somehow, back to the 1970s, there has also been this recent revival of interest in IO in big questions, in part driven by macroeconomists asking big IO questions because the IO people haven't been doing it. And so there has been a big question in these questions like, Has competition declined in the US economy as a whole? And do we believe that there's been this increase in mark-ups or drop in concentration over the last decades?

So two papers are both-- these are both going to be classic NEIO papers. So they're going to have industry-specific questions of interest. And they're also going to have theoretical models that they're either trying to comment on or use as guidance.

The first one is, I would say, really-- it's an industry-specific question of interest. It's also, more than that, a paper that's about understanding theoretical models and whether they really apply or not. So the paper's, "Are durable goods consumers forward-looking-- evidence from college textbooks." To the extent there are policy questions here, these should be familiar with you.

One policy question is, Why is it that textbooks cost $300? We know that you can print a book for 20 bucks. Why do they cost 300? Is there something wrong with that?

And the other thing I don't know if you experienced as an undergraduate is, Why is it that this book that was only 2-and-1/2 years old is now already having a new edition come out, and I can't sell my old book back to the bookstore, which I had bought for $300, and it's now worthless because the 14th edition of somebody's book just came out replacing the 13th edition three years later? Is this some kind of scam or whatever the firms are doing that there's something about the durable goods sales where they bring out these new books over and over and over again to earn greater profits than they could if they just brought out books at an efficient time scale, whatever that means?

So I think I said most of these things. Yeah, so obviously, the theory-related question is, Are these durable goods models relevant in practice? In particular, are consumers forward-looking and well-designed and solving dynamic optimization problems when they're doing things, which always you think about students are so bad at answering questions on my exams about dynamic durable good pricing. But somehow, we always assume that they always know how to do that for themselves in real life and think about the future discounted value when they're making any purchases that they're making. Is that a good assumption?

The choice to study this, if you're trying to study durable goods and forward-looking consumers, this paper, you can see why they do what they do. First is, in addition to studying durable goods, they do get an industry-specific question of interest at the same time. It's good to do two things instead of one in a paper.

There are aspects of this textbook example where the simplicity of the business makes the modeling easier. So, for instance, all consumers have to either buy the textbook in the first week of the semester or the week before the semester starts or not at all. There's none of this "I can wait till 7 weeks into the semester and buy the book or buy it a semester after I take the course to take advantage of the price declines or things like that." It just-- it's a static decision.

The other thing is that they only buy the textbook that you're assigned. If you're assigned by Pindyck and Rubinfeld's intermediate micro textbook, you don't think about, OK, instead of buying Pindyck and Rubinfeld, I could buy Bernheim and Winston, or I could buy whoever and Snyder. In some sense, students only make the decision between two products or three products-- buy the new copy of Pindyck and Rubinfeld, buy the used copy of Pindyck and Rubinfeld, buy nothing. And so you get this much simpler demand system than you would get if students were choosing from this menu of 57 different things they could buy.

We also can think about things we can ignore, like in many used good markets, you'd want to study a lemons problem. What do consumers think about the quality of the used car versus the new car, and how is that impacting them, and how does that change over time?

Used textbooks, especially if you buy them in the college bookstore, you can just open them up, look at them, see how much highlighting there is in it, and then buy it or not. There isn't this asymmetric information we would want to model.

And then the other thing you're always trying to do is get enough observations to estimate your model. And here, what's really nice is that there's a very, very large number of textbooks. And so they have three subjects-- what is it? It's economics, psychology, and biology. They have 10 semesters in which they can study it, and then they have hundreds and hundreds and hundreds and hundreds of textbooks.

And so that combination of semesters times subjects times the textbooks is going to give you many, many observations. And you can-- there's always this thing of if you study 100 different products, are those 100 different products really identical? Can you use the same parameters for them? You can think textbooks are similar enough to each other that the elasticity of demand for Pindyck and Rubinfeld and the elasticity of demand for Bernheim and Winston are probably the same, and we can just put one parameter in there instead of having hundreds of parameters.

Some background facts before you start modeling-- it's always good to think about things that are true. Well, one is that new editions come out every three to five years. Second fact is that textbook prices seem to be fixed. Why this is true is kind of a puzzle.

Maybe you have to have some reputation model, but you might have thought, OK, here's how I'm going to do this test. The book is going to come out the first-- if I buy the book when it's new, I know I could rent it to five future students. If I buy the book one semester, well, I'm going to rent it to four future students. If I buy it, like, four semesters in, I'm going to rent it once to somebody. So there ought to be this real decline in the prices of a book. And a book that's five semesters old ought to be worth one-fifth of a book that's brand new because you can't keep renting it over and over again.

Well, that just doesn't happen. Textbook prices do not drop. The book comes out, it's $350. Five semesters later, it's $350. So for whatever reason, we're not going to think of optimal pricing. Maybe it's a pricing puzzle, but textbook prices just don't change in between editions.

I don't know to what extent these are still true in the internet era. But at the time, the used-to-new ratio was standardized at 50% to 75%. Does the Coop even sell-- you can still go to the Coop and buy all the textbooks for the courses? Is that right? So maybe no one does that, OK.

So anyway, if you wanted to buy your books, you would go to the MIT Coop. Down in the basement of the MIT Coop, there'd be a bunch of textbooks. MIT would have an agreement with the bookstore. In this case, we used to own the bookstore, but it's now Barnes and Noble who owns it.

But they would sell the used textbook at, at most, 75% of the price of the new textbook. And then you would also often have an agreement with the company that was selling books in your bookstore. If you want to buy back used books at the MIT Coop, you must pay 50% of the cover price when you buy it back.

And so there are these norms of people would show up in the bookstore at the end of the semester willing to buy back your books at 50% of the cover price. If you wanted to buy a used book instead of a new book, it would be sold at 75% of the cover price.

And from aggregate numbers, roughly 30% of the students in a course at the time would buy a textbook. About 20% would buy a used textbook. And then the other 50%, we hope, make extensive use of the library or some other resource because they don't seem to buy books at all.

The hazard rate for revision seems to peak in about the third year, that many books have a new version that comes out every three years. So if you're looking at John Gruber's public finance textbook, if it's been two years since it's come out, he's probably going to have a new edition next year.

Once you've gone a bit past three, the probability that it gets revised goes down and down and down, because it's probably the faculty member has just gotten exhausted with the 11th edition or the 12th or the 13th and has decided they're not making enough money off the new editions to write the new ones. They haven't got a new person to take over the book. But if the thing hasn't come out and it's a few years late, it's probably not coming out.

So papers, empirical approach is a two-step strategy. I'm going to estimate the probability that the book is going to be revised before the next semester. And then I'm going to see whether-- because prices don't change, if you're a student and you know a new book is coming out next semester, you shouldn't buy, or you should be willing to pay much less for the book because you know that the cost is really $350, not $175, so that, in some sense, the price is much, much higher in the semester when a book is about to be rendered obsolete than it is in other semesters. And so what we ought to see is a drop in price, a drop in demand in semesters when books are going to be obsolete or when people are going to think there's a higher probability that they will become obsolete.

And if consumers are rational and forward-looking, we can see-- we can figure out-- if we know the elasticity of demand for books, we know how much drop demand would go down with such a price increase. We can almost estimate how many consumers are rational and forward-looking by seeing how much does demand drop when that new book is coming out.

So step 1 is a very simple exercise. This is just regress-- you just estimate a hazard model where you just have the probability that the book is replaced is a function of the age of the book in semesters, whether it's the biology or psychology or economics book, whether it's an introductory textbook like for economics I or whether it's an intermediate micro textbook or intermediate macro textbook.

And here's the graph they get for the introductory books. So you can see in economics books, your book is five semesters old. There's, like, a 35% chance your book will be rendered obsolete one semester from now. If your book is six semesters old-- oh, actually, this is six semesters old. So there's actually only an 8% chance that your book is going to get rendered obsolete next semester. Or is it here? Anyway, seven semesters old, it's a 5% chance. But there's this big peak centered around buying five-semester-old books is a bad idea.

And the thing is less consistent in some of the other disciplines. Some of the biology books are being on a two-year cycle instead of a three-year cycle. Some of the psychology books are on a 2-and-1/2 year cycle, something like that. And again, your book is-- your economics book is eight years old, or any of your books are eight years since the edition has changed, they're not going to have a new one. So that book is very safe to buy.

How do we do demand? So can I ask-- well, actually, some of you haven't taken it yet. So how much discrete choice estimation is there left in [14.]381? Or has it all gone to causal inference these days? A lot of discrete choice or no?

AUDIENCE: Causal inference.

GLENN ELLISON: It's all causal inference. So is there any place where people are doing logit models and whatever? OK. [LAUGHS]

OK, anyway, let me say, Tobias has a full week of lectures planned on discrete choice estimation, which obviously, a full week is not so much. There's much more in [14.]273. But anyway, so it's week four of this course, Tobias is going to spend a week on discrete choice estimation.

For now, let me just say this is how everybody always seems to do demand estimation in IO these days. There are a lot of funny elements of it. But so what we're doing here is we have this model of consumers, and we specify a utility function that they're going to get from buying three objects-- a new book, a used book, and no book at all.

And what utility do you get if you buy a new book? First, there's a Xjt beta, which is something on book characteristics times some coefficients. Then there's the disutility of paying for the book. So alpha is your price sensitivity. And then we multiply alpha by the net price that you would pay. The net price that you would pay is the full price minus a discount that you apply to money that's going to come to you in the future times the money you expect to get back when you resell the book.

So this is if the book doesn't die, I can sell it back. And when I sell it back, I get mu times Pjt, where mu is maybe 50%. And so this discount factor delta is both I'm getting the money six months from now when I sell it back instead of getting it today. And it's also, am I thinking about it. Like, you could imagine totally myopic consumers are not thinking about selling it back, and they're just treating delta as 0. So they're just, Do I want to pay $350 today, not thinking about the future?

In order to explain the data, there are always-- for some reason, [INAUDIBLE] started this use of the letter xi. So this is the unobserved quality of the book. What is it that makes Pindyck and Rubinfeld sell many more copies than many other intermediate micro textbooks? The characteristics in terms of how many pages long it is and whether it's hardcover and whether it has a CD-ROM shrink-wrapped to it-- those, it's not very different. But it must have some unobserved-- or it may have some unobserved quality that makes it a higher fraction of students buy it.

And then you have this-- consumers have this sort of a type. In my previous lectures, consumers always had this type theta that determined like their taste for goods. Here, consumers have a three-dimensional type. They have a three-dimensional type epsilon ijt new, epsilon ijt used, and epsilon ijt none. So I have this three-dimensional type describing how much I like buying new books, how much I like buying used books, and how much I like buying no books at all for this particular book.

And we model this as being three independent extreme value errors, typically. So this would be my utility for buying a new book. Again, if my utility for buying a used book is going to be some function of the book characteristics plus some unobserved quality of the used book plus this idiosyncratic reaction, then I have an idiosyncratic utility for buying no book.

And so what we model many consumers doing is having these three choices, and then based-- different people make different choices because different people have different epsilons. So we have this utility specification, but then given any set of parameters-- beta, alpha, mu, et cetera-- and given assumptions about the joint distributions of these objects, we're going to be able to get demand as a function of the economic primitives. Yeah?

AUDIENCE: So just do the characteristics of the book enter differently, and your user is just yf or just another beta?

GLENN ELLISON: Yeah, no, so it potentially could be xjt times beta, or it could be some other function of xjt. In their estimation, it turns out it's-- under their estimation assumptions, it doesn't matter. In some sense, any function here, the estimation would work. So it's like we estimate this robustly. We don't need to assume this equals this, and so they don't.

So theorem-- and this is the magic of logit models. The magic of logit models-- and, you know, this may seem skeptical at first-- is like, basically, you have these quantities, and then you take the log of the quantity and you subtract something, and you have converted from quantity space to utility space. And again, Tobias explain more-- make this stronger argument for why that's not-- doesn't seem so crazy. Maybe it is crazy, but people do it all the time.

So anyway, the theorem is, if the epsilons are i.i.d. across consumers and products, and each of those epsilons has a type 1 extreme value distribution, then the log of the ratio between the number of people buying new books and the people buying no books is that utility specification. So it is book characteristics times effect of book characteristics on utility minus alpha times the price, plus alpha delta mu times the probability that it doesn't die times the price, plus this thing that was the unobserved product quality becomes the error term if you're regressing new book sales on book characteristics and prices.

So we make this functional form transformation, and we get to this thing that we could estimate these parameters, the utility function, just by running a regression. So we're going to regress log of the share-- log of the fraction of people who buy the book on book characteristics, prices, probability of death, and so on.

One thing that this specification points out is, the coefficient on the book characteristics are going to tell us how much people like big books versus small books. Coefficient on price is going to tell us how much people dislike buying books. The coefficient on the probability of survival times the price is going to give us alpha times delta times mu. So delta and mu are not separately identified.

So if I said that delta was the key parameter of are people forward-looking, the mu was how much money do they expect to get back when they resell the book, if they're allowed to resell it. We can't separately identify this product. All we can do is estimate those two things multiplied together.

So another observation, we're going to need an instrument for price. Why are we going to need an instrument for price? We always need an instrument for price.

But if the firm has a book that has a high unobserved quality, this is a book that many people like-- the same reason that just like when demand is high, price is high, to the extent that's true. But if the book has an epsilon new that makes the firm want to set a high price, then price is going to be correlated with this error term, so price is endogenous. So we need an instrument for price.

What do they do? The instruments they use for price are having a nonprofit publisher. So, for instance, Tirole's textbook is published by MIT Press rather than by Macmillan or somebody else. Tirole's book is cheaper than many other books because nonprofit publishers tend to set lower prices.

Other instruments they use are, What fraction of your close competitors are published by nonprofit publishers? So if intermediate macro books, a lot of them are sold by nonprofit publishers, then maybe the for-profit publishers also have to set lower prices. They also have a measure, which is what's the concentration of ownership? Are there 10 different major publishers who have intermediate machine learning books, or are there only two such publishers who have dominated the market?

Anyway, so this is basically the estimating equation. So it's just an IV regression. And it's an IV regression that, under the assumptions they've made, should let us estimate delta mu, which is what we care about.

And what do they find? So they define lambda equals delta times mu. What did we think mu is? We think that mu, you've got to be able to get at least-- normally, you'd get 50% back at the bookstore. If you're going to sell it back to your friends, maybe you could go to your-- well, not your closest friends-- and say, I'll sell my book to you for only 75% of the cover price, because if people want to go to the campus bookstore, they can buy the used book at 75% of the cover price. You're not going be able to sell it for more than that. Just posting a sign in the dorm saying, I've got a used copy to sell.

But so if we think mu is between 0.4 and 0.75, they estimate is-- they estimate lambda is 0.61. And so if you estimate that lambda is 0.61 and you think that mu is in this range, that's got to tell you that mu is not 0.4. Mu is bigger than 0.4, and delta is pretty close to 1. So the estimates are indicating that consumers must be fairly close to all being forward-looking and putting a very large weight on that revision probability because they're reacting a great deal to books that are about to die. And they could only react that much to books that are about to die if they recognize they're going to die, or there's a high probability of death and put a lot of weight on that.

So in some sense, we're estimating two things. We get this-- I have alpha and alpha delta mu. So the coefficient on the probability of death is giving me this, the coefficient on the price is giving me this, and so then I divide them, and that's where I get the lambda from. Any questions on that?

So, obviously, as with any empirical paper, I can think of lots of concerns. And I think one of the hurdles to getting research done is you just-- there are always going to be problems with your paper. You just go ahead and write them.

In our profession, the people who get ahead are the people who write the papers, and then other people criticize them. That's OK. Just writing the papers and having people talk about your paper is what you have to do to get ahead in this world. And so you write the paper, there may be flaws, there are things people can criticize. But as with any paper, there can be lots of things to criticize, and that's OK as long as you get it out there.

So what can I criticize about this one? They have no data on the non-bookstore sales. They have very good data on the bookstore sales they get through the bookstores themselves. They really don't know much about the used market. They have a couple different treatments of the used market, but, in some sense, they're kind of making up the used book sales and just going with that.

Instruments are important. I think the non-profit publisher instrument is a good instrument. But let's suppose you haven't fully instrumented for this thing, the lambda is-- lambda hat is going to be like alpha delta mu hat over alpha hat. If the alpha hat ends up biased toward zero, you're going to be inflating the lambda hat. So imperfect job of instrumenting for alpha is going to make you think that lambda is bigger than it is. You worry, could that be part of what's going on?

Distribution of errors-- again, we have that theorem. If the epsilons are i.i.d. extreme value-type errors, then log of s over log of s equals this. There's a converse of that theorem, which says if the errors are not i.i.d. extreme value type one, then log of s new over s0 not equal to x times beta minus alpha times p minus alpha times p1 minus die. So obviously, we're relying on this functional form identification in that sense.

We're also relying on functional form identification in the second sense, which is the probability that a book will die is a function of its age and the field it's in and other things. So by using that probability of death as an instrument on the right side-- so the way the estimation is actually done is they put 1 minus whether it actually died here on the right side and then use the probability of death as an instrument for whether it actually died.

But still, using that-- you're using as an instrument something that's a function of variables that are all in your equation. And so it has to be that somehow, that nonlinear function of semesters and whatever is uncorrelated with the errors in the demand equation, even though you put the variables themselves also on the right side just linearly. So there's a use of functional forms to identify that.

I should say, one nice thing that they did at the back of the paper to try to make it more convincing is this paper was written before books were sold on Amazon. They then got some more recent data of used books sold on Amazon.

And they also find this effect there on the price side that if you look at, at what price can you sell your used book on Amazon, it goes down dramatically in semester 5. And sometimes, you go on Amazon, and you see the 12th edition is already available, not yet available to buy, but it's there. I own the 11th-- I'm being asked to buy the 11th edition. The 12th edition is listed on Amazon as available October 15. In that situation, price is really-- price of used books really, really drop. Any further questions on this paper?

So let me go to final paper-- and again, this is a-- the final paper I thought I'd talk about is-- it's 10 years old. It's 10 years old. It almost seems-- or 12 years old. It almost seems more topical than it did when it was written. So this is about media slant. This one, the question is, why do we see biased news media?

One reason why you might think we see biased news media is maybe Rupert Murdoch is very conservative and wants to spread conservative values across the United States. Or maybe Newsmax is secretly controlled by Russia, and they're out to display misinformation that will destroy American democracy.

But another explanation would simply be that running Fox News and getting people angry about things is a very, very profitable business. And you don't even-- Rupert Murdoch doesn't even have to be conservative to want to do this because he's just making lots and lots of money because people watch Fox News. In this case, I said in Chevalier and Goolsbee, it was like there was a public policy question about books and why they're so expensive and why they get revised.

But then there's an understanding the theory better. Here, I think it's less of an understanding of the theory better than just recognizing the theory, which is the theory is the optimal product quality theorem from monopoly pricing, that we think monopolists ought to choose the quality of their product that is optimal to maximize profits. And so the question here is, does that optimal quality theorem apply to help us understand Fox News, and does it look like Fox News is profit maximizing, and do we understand things are profit maximizing?

Now, turns out this paper is not about Fox News. This is about newspapers. And again, things do age. There was a time in the United States not so long ago when people read newspapers. Newspapers were things printed on physical paper, delivered to people's houses by carriers. And then you would get up in the morning, pick up your newspaper, and read it while drinking your coffee and eating breakfast.

There were many, many local newspapers in the United States. Every city had multiple newspapers. So in Boston, instead of just having the *Boston Globe,* we had the *Boston Globe* and something called the *Boston Herald. Boston Herald* was smaller and easier to read while riding on the subway. But many towns across the United States had newspapers. Those newspapers would convey both local news and national news.

So let's see, motivation-- so anyway, media slant remains a very big topic. I should say that-- I mentioned Fox News, I think another very interesting topic, there's a company flying more under the radar, Sinclair Media. Sinclair Media is basically the Fox News of local television. So Sinclair has bought hundreds and hundreds of local stations across the United States and made the 10:00 and 11:00 news that you receive free over the air more Fox News-like.

Anyway, what Gentzkow and Shapiro were trying to do is ideally, they would like to have this two-step investigation of understanding motivations for biased news. Step one is estimate the slant that would be profit maximizing. And then step two is compare the actual and profit-maximizing slant and see whether there-- are there differences between the two? If there are differences, are they related to the ideology of the person who owns the newspaper?

I think the paper falls a little short of the self-description, or it's a little heroic in the assumptions that you need to think that it's estimating profit-maximizing slants. But it has been very influential for the questions it raised and also, actually, for the ways it's quantifying media bias. And techniques developed here get used in many other papers on media bias.

OK, so quantifying media slant-- this was really one of the first papers to use-- you wouldn't even these days call them machine learning techniques, but automated text processing as a way to get data. And so what they're trying to do is they're trying to figure out, is a newspaper right skewed or is a newspaper left skewed? How do you make a newspaper skewed one way or the other?

Well, one thing you do is just with the words you use. And in the early 2000s, they give examples like, do you call something the Iraq War or the War on Terror? And generally, people who favored the Iraq War would call it the War on Terror. People who opposed the War in Iraq would call it the Iraq War.

Estate tax and death tax is a very nice linguistic thing. In the United States, if you have less than $5 million when you die, you can pass it on to your heirs tax free. If you have billions of dollars, it seems like you can often pass it on to your heirs tax free, never having paid the capital gains taxes either. But there are taxes on the books that you're supposed to pay when you pass your billions of dollars on to your children.

But anyway, the Republicans tend to refer to it as the death tax, because who wants to pile on someone's misery when someone has died and tax them for it, versus the estate tax. Everyone can agree on taxing estates. If you have wealthy people, estate-- wealthy estates should pay taxes. So there's a question of what words do you use in your newspaper?

But then much more-- if you look at what Fox News does, some of it may be outright lying about voting machine tampering or something like that to get people angry. But much more, it's a question of, do you have a lot of news stories about undocumented immigrants assaulting and killing people? Or do you have a lot of stories about people invading the US capitol and beating police officers or whatever? So in some sense, the extent to which you talk about tax breaks being given to corporations or the extent you talk about illegal immigrants committing crimes is going to be one way in which you can slant your news coverage and change political opinions of the people who live in the area that you're serving.

And so what they try to do is-- the question is, how do you pick out a Democratic from a Republican issue or a Democratic from a Republican talking point or Democratic from Republican phrasing? And what they do is they go through the full text of all speeches given on the floor of the US House of Representatives and the House of Senate. And they pick up two- and three-word phrases that are being used by Democrats and Republicans. And they actually classify each congressperson as a where they are on a continuum spectrum by what's the vote share for George W. Bush in their congressional district.

And so they use this for what is being spoken by far left people, what's being spoken by far right people. And so they identify basically-- they select 1,000 two- and three-word phrases that are used significantly different frequencies by people on the left versus the right of the political spectrum. And then for each newspaper, they regress the excess frequency with which it uses each phrase on that phrase's Republican-ness. And they define the slant of the newspaper as the coefficient on that variable.

So if it's really using almost exclusively right Republican phrases, it's going to be a very high coefficient. If it's using mostly Democratic phrases-- so if the distribution looks like this, it's going to be a very left-skewed newspaper. If the distribution of phrases looks like that, it's going to be a very right-skewed newspaper. And they measure the skew by the slope of those estimated lines.

Now, obviously, part of that in these examples sound like things where you're doing media slant. There are also going to be all kinds of other things picked up in this. So, for instance, some Congresspeople talk about agriculture policy, and some talk about guns in their urban neighborhoods.

Talking about agriculture policy may just mean that you're talking about agricultural policy. But generally, Republicans tend to represent agricultural states, and so you may get corn as being a Republican phrase when it's not-- talking about corn prices is not a political statement. They didn't pick the phrases to have political content. They just picked phrases automatically. And so there's going to be some contamination of political slant with other differences.

Consumer demand, again, is going to be derived from utility function. We assume that the utility that consumer I in ZIP code Z gets from buying newspaper N is some function of the average quality of the newspaper minus disutility that the consumer has from consuming something that's different from the ideal slant. So Iz is the ideal slant of all people in ZIP code Z. yn is the actual slant of newspaper n. So as the slant of the newspaper differs from what's ideal for people in that ZIP code, utility of people in that ZIP code goes down. And then there's a random error term.

They assume that the utility-maximizing slant in ZIP code z is just alpha plus beta times rz, where rz is the fraction of people in ZIP code z who are Republican. So it's a very specific assumption in several ways. So there's no heterogeneity. All people in ZIP code z have exactly the same preferred slant.

And there is no error in this equation. All people in-- preferred slant in ZIP code z is this linear function of the fraction Republican, with no heterogeneity across ZIP codes. There are no ZIP codes where a left slant is better than you would think, given how many Republicans live there.

Once again, they also assume now that these epsilon izn's are i.i.d. extreme value, and then they make this magical transformation from sales of the newspaper to utility of the purchasing consumers. And they get that the log of the market share of purchasers versus nonpurchasers is the average utility minus gamma times the difference between the newspaper slant and that linear function of the fraction Republican in the ZIP code.

And obviously, we can just expand this quadratic. And when you expand this quadratic, you're just going to get things like-- you're going to get y on the right side. You're going to get y squared. You're going to get r squared, and you're going to get y times r. So you're just going to have this thing that's this right-hand side term that's got y's and r's and y squareds and r squareds and y times r.

And in this model, if this model is correct, the utility-maximizing slant is alpha plus beta rz. And we can estimate how the utility-maximizing slant varies with rz by just looking at the coefficient of-- the coefficient of yn times rz is going to be-- it's going to be gamma minus 2 times gamma times beta.

So anyway, we're going to get that-- yeah, the coefficient on yn times rz is going to be minus 2 times gamma times beta. And so we're going to be estimating-- we're going to get gamma times-- identify gamma times beta by just looking at the coefficient of this when I regress log of share of purchases of the newspaper on the right-hand side variables.

What did they find? The clear finding is that this ZIP share Republican times slant has a clear, positive, significant coefficient. So this is evidence that the ideal slant is to put more Republican news in a Republican ZIP code from a profit-maximizing sense.

Let me say that my view on this is-- it's a very, very strong assumption to assume that there's a profit-maximizing slant, and it doesn't differ across consumers within a ZIP code, and it has no error term across ZIP codes. I find it convincing evidence that in more right-leaning areas, you do better with a more right-leaning newspaper. The fact that this slant is the profit-maximizing slant-- that relies on those no errors, homogeneity assumptions.

So it takes pretty heroic assumptions to say that this really is the profit-maximizing slant. But I do think it's pretty clear evidence that right-slanted news does sell better in right-slanted areas, which is fundamentally-- we're regressing log of q on the share Republican times slant interaction and getting a positive effect, saying in more Republican areas, more right-slanted news is selling more copies.

And then is slant profit maximizing? They show this graph. And this graph is supposed to show you, are newspapers more right slanted in more Republican areas? And the answer, unsurprisingly, is yes.

So we get this positive slope. The more Republican is the area in which your newspaper is being sold, the more right you do make the slant of your newspaper. The demand estimates aren't separately identifying gamma and beta, so it's a little hard to say whether this slope is the slope we would expect given the profit maximization.

So then the final thing they do is they say, well, let's look at-- regress the slant on the share Republican in the market and then ownership fixed effects to see whether it looks like-- they do this two ways. One is they look at owner political contributions and put owner political contributions on the right side. For some reason, it's not in the same table.

And the other thing they just do is put in ownership fixed effects and see whether the owner fixed effects increase the r squared of the regression. And if you put the owner fixed effects in, it does make a big increase in the r squared of the regression. It more than doubles the r squared from 0.18 to 0.44.

But obviously, you're putting in an awful lot of fixed effects on the right side. And so what they note is that if you put in state fixed effects instead of owner fixed effects and then put in state fixed effects and then an ownership effect, then the ownership effect is not significant. So while the ownership fixed effects do increase the fit, you put in owner fixed effects and state fixed effects, suddenly you can't tell whether the owner effects are actually there or not at all.

So they take this to argue that it looks like there isn't much evidence that owner slant-- it looks like the "Rupert Murdoch is just making money" story works pretty well here, and it's hard to say that there's a lot of media bias other than that. They also put in-- they do a version where they put in owners' political contributions on the right side.

And it's like, well, this share Republican in the ZIP code explains 20% of the variance, it's like the ownership political contributions explain 5% of the variance. So it looks like the ownership political contributions are just not a powerful explanatory variable compared to the Republican-ness of the ZIP code that they're serving.