[SQUEAKING]

[RUSTLING]

[CLICKING]

PROFESSOR: Welcome back. This is the second lecture on empirical models of demand. So let me give you a little bit of a roadmap for today.

So I'm going to first talk a little bit about the supply side of these static models of competition that we have been discussing, and that's actually quite useful even for informing the demand side. So I hope that this will become clear once I go through this. Then, I'll show you the empirical results in BLP-- Barry, Levinson, Pakes, the seminal paper that popularized the random coefficient model. And the results are actually quite nice to understand how these models work and what the case for these models is.

And then I asked you to read Nevo's paper on the cereal market. So we'll discuss that, and then we'll actually stay in this industry. And we'll talk a little bit about the debate that has captivated many people in IO, which is the common ownership hypothesis. So let's jump in.

So the supply side. So here is the objective function of a firm that owns potentially different products. This set of products here is calligraphic I, and it maximizes a sum of these individual level products. So this here is the margin. This is a quantity for product J, and this is a function of the vector of all the prices in the market.

So think of this q here as coming out of your demand model. So it's the vector of market shares that you get out of your demand model times the number of consumers that are in the market. And so then there's some fixed cost. The fixed cost is not going to be relevant for what we're going to do today. And later on, you'll learn about models that actually allow you to get at those fixed costs, as well.

So now, it's going to be very useful to define what is called an ownership matrix. So again, this is potentially a multi-product firm. The firm owns, let's say, several brands of cereals in the market, and so each row here tells us for that firm that indicates that row, which products does it own. So this is simply a matrix of binary entries.

And now you can do the following. You can define a new matrix, omega, that multiplies this ownership matrix by the derivatives of each product with regard to all prices in the market. This comes out of the demand model. You can only construct those derivatives once you have the demand model.

And so you get this new matrix, omega. And with this matrix omega, you can write the sets of first order conditions of all the firms in the market as this vector of quantities minus this omega matrix times prices minus marginal costs. These are all vectors, and this omega is this matrix that we defined here.

And this allows us to invert those first order conditions to back out marginal cost. So this is simply a matrix operation. So what this shows us is that with demand estimates and a conduct model that actually allows us to construct these first order conditions, I need to know how firms are actually competing. As well as observed prices, we can recover a vector of marginal cost.

What this also means is that the model is going to be always rationalized by a vector of marginal costs and is, therefore, not testable if we allow the marginal cost to be completely flexible and have no other way to discern whether the marginal costs that we back out are good marginal cost. So this comes back to the exercise that Nevo is doing. So we're always able to rationalize the model, but we cannot, without further assumptions, test the model. So again, just to reiterate the point, depending on what the true ownership structure looks like, or to what extent firms internalize profits of competitors, which we'll come to later today, will depend on how they want to price. And for any such conduct assumption, we can back out a vector of marginal cost.

Now, what you can alternatively do, instead of allowing marginal cost to be completely flexible-- I mean, we would call this non-parametric because we're not imposing any assumptions on those marginal costs. We can specify some functional form for the marginal cost. So, for instance, in the case of cereals, we might know that this varies with the market price of sugar, wheat, corn, and so on, or advertising prices, shelf prices, and so on. And so if I'm pretty confident that I can write my marginal cost as a function of these things, I might be willing to impose this.

So what this means now is that I can come from both sides, so to say. I can estimate my demand model with a conduct assumption, back out the vector of marginal cost. This is now data. For any given conduct assumption prices and market shares, this here is simply data. And this here is some linear function that I propose.

Now, why might this be helpful? So you have read Nevo's paper today where he first estimated demand, and then tested the recovered markups against observed markups. What you can alternatively do, if you're fairly confident in your conduct assumption-- you believe that these firms are maximizing their own profits, potentially accounting for these multi-product effects-- you can estimate the demand and the supply side jointly.

So you can-- from this equation here, which again, think of this here as data for a moment. For any given demand parameters, you can think of a vector of marginal cost that you back out. You could, from this, construct additional moments. Remember last time, we constructed moments interacting C with instruments on the demand side. We can construct a similar set of moments on the supply side, typically denoted omega, that we, again, interact with some instruments of things that are either exogenous, or-- included or excluded instruments on the supply side.

Now, If I estimate everything jointly, the supply side will actually also impose restrictions on the demand side parameters. So just to give an example, so what happens, for instance, if you find that firms price on the inelastic portion of the demand curve? You may recover negative marginal cost. I mean, for many strategic models, you will be recovering negative marginal costs. You have to somehow rationalize why firms priced it low, and then it says, if they need some transfers from somewhere else to rationalize what's happening.

I can, alternatively, impose that marginal costs are positive and use the supply side restrictions. And that means that when I guess a set of demand side parameters that tell me that the marginal costs are negative, but I force the marginal cost to be positive through a reasonable restriction, that this will lead to a violation of these moment conditions here. So it leads to larger deviations on the supply side moment conditions, and the supply side moments will, therefore, tell your minimization routine, hey, the demand side parameters might be such that I can't recover a good set of marginal cost.

So this can be very helpful to estimate both jointly. Now in practice, there are not all that many papers that actually do that. So I would say the majority of papers just estimates the demand side separately, but BLP have done this. And in part, we credit their good estimates for the random coefficient parameters to imposing these supply side restrictions. Yes?

AUDIENCE: Is this also informative to the extent to which you're making way too restrictive assumptions on your model? Like say you're trying to impose positive marginal costs, and then in your minimization of nu you're not getting the results you're expecting. Is that a sign that your results may be guided by your model extensions?

PROFESSOR: So that's a great question. So, in fact, at the very end of the lecture today, we'll discuss a case where you can use supply side restrictions as overidentifying restrictions on the model. So instead of, let's say, only estimating the demand side and being just identified, you typically get a lot of additional moments out of the supply side because you've not introduced all that many parameters. But maybe you have the price of wheat, and the price of corn, and the price of advertising, and all those things vary.

So that leads to overidentifying restrictions. And indeed, with those overidentifying restrictions, you can test different conduct models directly. Yeah, that's right. And this is somewhat-- I will comment on this later on-- it's somewhat different from the exercise that Nevo does. He directly compares to data.

So my comments on using the supply side, as I just said, this oftentimes introduces many new moment conditions. It's received wisdom that it improves the precision of demand estimates, and especially the random coefficients, again, through the logic that I walked you through. It's econometrically efficient to estimate the demand and supply side jointly if you're fairly confident that your conduct assumption is right. Of course, you may not feel comfortable making the assumptions that firms maximize their own profits or collude directly. So that seems a less innocuous assumption than saying, well, consumers just pick the best product that they like. And again, you can use overidentifying restrictions to test models of conduct, and we'll have an example of this at the end of class today.

So let me talk a little bit about the results of the BLP paper. I think they show nicely why you would want to favor random coefficients model. So they have data from 1971 to 1990. The market here is the entire US.

They use list prices, so not transaction prices. You could bicker about this. And an observation here is a model year, and they have 2,217 of those. As characteristics, they use the number of cylinders, doors, horsepower, length, width, weight, wheelbase, and measures of fuel efficiency, dummies for air conditioning, and so on.

So here are some interesting summary statistics on the market. So first, something that's striking is that the number of products just increased a lot over time. In fact, there's a recent paper by Grieco, Murry, and Yurukoglu who estimate mark-ups for the car market over many, many years. And the number of products, as you would not be surprised, goes up even more, and they actually find that consumer surplus in the car market has gone up over several decades.

Prices here go from $8,000 to about $10,000, they're denoted in 1983 prices. Something that's interesting is that the share of domestic brands goes down over time. This is mostly driven by increasing imports from Japan.

Another interesting observation here is that the size of cars is going down, very contrary to what we see today, where cars seem to get bigger and bigger. And the authors here speculate that this might actually be driven by the first oil price shock in 1970. So these are the summary stats.

This is the miles per gallon and miles per dollar. They use miles per dollar on the demand side. Consumers will care about how fuel efficient their car is in terms of the dollar expenses, but producers might not directly care about that. They will only care about how much more costly is it to produce a fuel efficient car, so they don't care how much the price of gasoline varies.

Here are some summary statistics where they denote different percentiles of their variables with actual models. So you see prices go from this famously crappy Yugo from $3,000 to almost $70,000, so a factor of 20. Horsepower is also varying widely. Somewhat ironically, the Gran Fury comes in at the lower end of the spectrum in terms of horsepower per weight. And you see here different summary statistics-- a lot of information on cars.

So here is the first table that I think is going to be useful to dissect a bit more. And in part, I want to make a point here that might be relevant for you practically for those of you who want to estimate these kinds of models. So you'll notice that it takes a while to code these models up, and debug them, and so on. So it is typically useful to do what these authors are doing here is before you estimate the big structural model to use this simple inversion that you get from the logit model without random coefficients. And remember, we can write this as a linear estimating equation and estimate it via OLS or two stage least squares-- things that you should be well familiar with at this point.

And you can actually see how sensitive your coefficient estimates are, let's say, to the inclusion of different instruments and so on. So I think that's practically something you should keep in mind that before you code up the big model and try to probe different specifications and different kinds of instruments, you probably want to do this Stata, whatever your preferred way of doing this is, and just run simple regressions. So I found this to be very useful.

And then usually, when you go to the random coefficients model, of course, you get very different substitution patterns. But you already have a pretty good idea what the average coefficient should look like. Any questions?

So let me tell you a little bit about what's going on on this table. So here, they have an OLS specification of demand-- again, based on the simple logit model inversion-- and here, an IV specification. What you can see here is something that you would expect if firms systematically price unobserved product attributes. Higher unobserved product quality leads to higher prices. So you would expect the price coefficient to be biased towards 0. And indeed, once we start using instruments, the price coefficient becomes more negative.

The other thing that you see, and this is actually of practical relevance again, that once you have a larger price coefficient, or a more negative price coefficient, you have a lot fewer elasticities that you get out of the model that are on the inelastic portion of demand, which would not be consistent with static profit maximization. In fact, oftentimes the problem can be even worse than that. And speaking from my own experience, you'll oftentimes find that elasticities are positive.

And then you have a problem. You know you have a problem. In the other case, you might not know that you have a problem. But in that case, you know you have a problem.

And so the last thing that they do here is they estimate a very simple supply side model of marginal cost pricing where they simply regress prices on a bunch of attributes. This is oftentimes called a hedonic regression. And to see these here can be interpreted as elasticities, how cost vary with different attributes. And so what is surprising here is that they can explain a large fraction of the variation with a simple model in terms of these attributes.

So this here is their main table. So I just want to point out a few things. So they have here the coefficient estimates, here the standard deviations of those coefficients. So this tells us how heterogeneous taste is for different attributes. And they estimate two different models which change something on the supply side whether or not there are scale economies in production.

And so a few things to point out here. They find that-- so first, they have significant estimates for all their standard deviations. So this is typically, again, not easy to achieve. And in part, this is because of the supply side restrictions that lead to more precise estimates even on the demand side.

The other thing that's interesting here, if you look at this, so it's here, this attribute, miles per dollar, is valued negative on average. So the average consumer doesn't seem to like fuel efficient cars. However, notice that the standard deviation on this is also very large. And so there will certainly be some consumers that value miles per dollar positively.

So what this model allows us to rationalize, for instance, is that if you, let's say, do this for the entire US, as they do, in much of the US, you see big cars. People seemingly don't care much about fuel efficiency. But then in, let's say, other parts of the country, you see a lot of Toyota Priuses, and Teslas, and so on. And a model that doesn't account for these heterogeneous substitution patterns with aggregate data wouldn't be able to get at this heterogeneity-- that different people have very different tastes for these attributes, and as a result of that, substitute differently.

So a few more tables that showcase that they pick up pretty good substitution patterns. Here, you see a bunch of different models that have in their estimation, and you can roughly think of this table as being ordered by how fancy the cars are. So this is, of course, not a purely vertical model, as we discussed last week, but think of this as these are more expensive, more fancy cars. And if you pick, let's say, a high end car like the Lexus, BMW, or Cadillac, you'll see that if you change the price, these are semi-elasticities. That means that how much do market shares change when I increase my price by $1,000.

You see that if you do this for the BMW, you increase the price by $1,000, there's almost no dent in the market share of lower end cars. And most of the action is here at the higher end models. And so similarly for Lexus, and Cadillac, and so on. The other thing that validates the model is that since we are looking at the same price increase for our low end and high end models, you'll find that the response for low end models is much, much stronger than for high models, as you would expect. If you're buying a $70,000 car, then an extra $1,000 shouldn't be as painful to you as if you're buying the $3,300 Yugo.

So this here is actually a quite striking table. This here shows how people substitute to the outside option if we base these diversion ratios on the logit model and compare that to the BLP model. So I'll let you stare at this for a second and think a little bit about what this comparison tells us, and in what sense the BLP model might be doing a better job here.

So I may just say something which is relevant here. Again, think of these as ordered in terms of how fancy they are. These are more expensive cars down here than up here.

AUDIENCE: The BLP is more capturing the fact that if you're going for an expensive car, increasing price won't induce you to move away from buying a car in general, but logit doesn't really get after that kind of aspect.

PROFESSOR: Exactly right. So we think that, again, cars are durable goods. If I already have a car and I'm considering buying a cheap car, I might be much closer to the margin where I drive my car for another year. Whereas somebody who is buying a $90,000 car, they'll probably not switch to the outside option if we raise the price of that car a little bit. They might instead opt for somewhat cheaper car.

And you see that the logit model cannot rationalize this at all. So it's quite striking. And so in the BLP model, you find that lower end consumers are much more likely to move towards the outset option.

So I think this is all I wanted to say about those results. Let me comment on the BLP paper, and also just give you some practical leads here on where you can find additional information on this. So, you know, as I'm sure you may have noticed, we spent almost two lectures on this. This has been a very influential paper, led to countless empirical studies, a large methodological literature. It's one of the bread and butter pieces of what you do in empirical IO estimating demand, and this is the standard model to do that.

The empirical results, the papers, of course, mostly known for its methodological contribution, but the empirical results make a nice case for this model. There were many things at the time when this paper came out that people didn't really know about. So, for instance, nonparametric identification, to what extent are random coefficient models identified if we do not impose a type 1 extreme value assumption? So people have thought about that, and they are very general identification results for general models of this sort.

People have learned a lot about how to estimate this those things efficiently. And Aviv Nevo, whose paper we discuss after this, he has contributed to telling people how to estimate these models. There are now many good sources to dive deeper into all of this.

So the new IO handbook chapter just came out in 2021. If you want to learn about demand, chapter 1 and chapter 2 are good sources. If you're really interested into identification, somebody wrote me a question about the type 1 extreme value error and how restrictive this is, this would be the place to look. These are papers that show you how these models are non-parametrically identified. And I would probably start with the annual reviews article, which gives an overview of this econometrica paper.

There is also now out out-of-the-box estimation with Python that you can use. My view on this is this is super useful, and you may want to check your results with that. But I also believe that you should learn how to estimate those models and do it yourself because that's the real skill.

And especially if you stay in IO, part of writing an IO job market paper is to somehow push what you can do with these models, and show that you understand them deeply as opposed to just apply them. So this is a great resource. But, I hope that once you come to [14.]273, we teach you how to actually do those things.

So let me then also just say a few things empirically on this paper. We had already discussed this. The paper ignores the durable good aspect here-- that people hold cars, and that substitution to the offset option is, therefore, very heterogeneous, and potentially endogenous to the counterfactual that you run. Because if you change prices in a new car market, prices in the used car market might respond to that. And so your counterfactual might be misspecified as a result of that.

But at least, because they have the random coefficient on the intercept, you get much more heterogeneous substitution patterns to the outside option than you get from a simple logit model. So it captures that to some extent. Again, there's some great papers to look at this further. There's a paper by Gowrisankaran and Rysman that looks at people purchasing durable goods. This is, I think, for camcorders. And so they endogenize, basically, how people update across different generations of camcorders.

There is a paper on the used car market by Gavazza and Lizzeri which ignores the product differentiation aspect of cars. So cars here just basically described by their vintage. So thinks about the welfare effects of the used car market. And then there's a more recent paper that actually puts the two things together again-- so the differentiation part, and the fact that cars are aging, and that different consumers might hold different vintages.

This is a paper that abstracts from dealerships. This is actually quite problematic if you think about the US because the US has state franchise laws that prohibit manufacturers to sell directly to consumers. And so you might have heard about this in the context of Tesla. They have been rallying against this because they want to sell directly to consumers, and do not have large dealership networks. So that might be important for local availability of brands and substitution patterns, and could also lead to other problems such as double marginalization.

The model ignores that car prices are often negotiated. I think for what they're doing, that's not too problematic. But we know that in reality, the transaction price might be very different from the list price. And I think Glenn has taught you a paper where that has very different implications for different kinds of consumers. So to the extent that you care about these distributional effects, that might be an issue.

And lastly, it ignores something that is very important, and increasingly so, which is the financial transaction that is involved when you purchase a car. Because you're typically getting a loan, and maybe insurance, and other kinds of add-ons. So obviously, this is an older paper. There are many papers that now address all of these issues, so that's a good paper to write if there are many gaps to fill and people are willing to do that.

Any any questions? Otherwise, I move on to Nevo. OK.

So the next two papers are in the spirit of what Glenn has already started talking about-- Bresnahan in 1987 and Miller and Weinberg testing different models of conduct against each other. So the idea is using our estimates to figure out what game is it that firms are actually playing? And I think that's a very nice way to use this machinery. We will look at the ready to eat cereal market. You have read Nevo's paper, and then we'll look at Backus, Conlon, and Sinkinson, which talks about the common ownership hypothesis.

So the question in Nevo's paper is, why does the cereal industry sustain such high gross margins as measured from accounting data and the census of manufacturing? So they seem to be these very large margins. There are two competing explanations.

One is that the firms are somehow-- of which there are few-- are colluding with each other. And the other one-- and I'm sure you've noticed this-- this is a market in which there are a lot of different products. And it could be that consumers have very heterogeneous tastes for those products, and somehow, firms are catering to these very idiosyncratic tastes, and therefore, are able to sustain high margins in conjunction with internalizing their multi-product incentives-- the fact that they do not want to cannibalize their own products.

The approach is to use demand estimates to recover markups. A different alternative conduct assumptions, it should say here. So you're using only your demand estimates, imposing certain models of conduct, and then see whether those models of conduct with which you can recover marginal costs, and therefore compute markups, line up with the markups that you actually observe in the data. And he uses panel data on demand for cereals across different geographic markets.

So I thought I'd use this as an opportunity to tell you about the surprisingly entertaining origins of the breakfast cereal market. So this here is James Caleb Jackson. He lived in western New York. He was a very religious vegetarian, and he stumbled into something that resembled maybe not too much the modern breakfast cereals.

So it was kind of a hard thing to eat. Not much fun. You had to soak it overnight in milk to make it edible at all. He called it granula.

Not much later, there's somebody else. And I should have, said James Caleb Jackson ran a sanatorium in New York. So that John Harvey Kellogg in Michigan, his family's originally from Massachusetts. He was a surgeon, and he invented something very similar that he first wanted to call granula. Then he got into some legal trouble and he said, call it granola.

So you compare James Caleb Jackson, Kellogg, and you compare granula, granola. So you may already guess whose legacy was more enduring. So John Harvey Kellogg had a younger brother, and he was really onto something. Because for 10 years, they worked together and ran a company together until they had some differences about including sugar into cereals.

So, again, you have to think of the old cereals as there was a religious impetus, and there was a whole philosophy around eating healthy. And then we'll suggest, no, let's put some sugar in it. We'll make it more successful. And it did.

John Harvey Kellogg also had a patient, Charles William, or CW Post. So the brothers also had disagreements about the importance of trade secrets, and John Harvey Kellogg said, our patients should learn how to do those things. And that's what Charles William Post did. He basically copied the recipe, called it Grape Nuts, and became a competitor of the Kellogg company.

So little did they know, so this is what happened to the market. Now you have hundreds of different products. It's also known that there's a lot of churn in the market, and that the market is sustained with very high advertising expenditures.

And it's just funny to look at those, but it also illustrates a point that there's just this massive variety. You could be, let's say, interested in C-3PO cereals, or maybe Donkey Kong, or maybe Donkey Kong Junior if that's not catering precisely enough to your taste. And I should say in 1970, there was actually a complaint by the FTC that the companies in this market are colluding with each other and proliferating brands to deter entry from rivals. And so Schmalensee, who teaches at the Sloan school, has a paper on this in the Bell Journal of Economics.

And so what is striking is you see the number of products in the market, and then you look at how many firms are controlling those products. And it's very, very few. So with the C3 concentration ratio, so the market share of the top three firms, you get up to more than 70% of the market. So what this tells you is this is a market where you want to take into account these multi-product effects-- the fact that firms will internalize how much they will rival their own products when they introduce new products or price those. It would not make much sense to not have such a model when you have four firms and hundreds of different products.

So again, to remind you, the exercise that Nevo engages in is he tries to-- so he first measures markups directly in the data based off of accounting estimates. And the relevant estimate here is the retail margin of 46%. And then he does some bounding whether or not you include advertising as a-- whether advertising is a marginal or variable cost, and-- sorry, a marginal or a fixed cost.

And in the end uses model estimated markups to compare it to those directly observed markups in the data. So this comes from this report here. But he also has some data from the census of manufacturing.

So the specification that he uses is very, very similar to what we have discussed last week. So I just want to point out a few things here. So again, you have an indirect utility function that's linear in terms of product attributes-- these random coefficients for product attributes, prices for product J and firm T, and the random coefficient for prices.

Now, the one difference here is that he decomposes the unobserved product attributes into a fixed effect and a deviation from that. So I got some questions about this fixed effect. So why haven't people done this before, and what is the difficulty here?

So first of all, of course, this came not too much later after the introduction of BLP. So there were not that many people that actually learned how to estimate those models. That's kind of a perhaps somewhat disappointing answer.

But I think there's one important point here-- or maybe two points, actually. The first one is that if you want to include these fixed effects, you're introducing a lot of new parameters into your model. And I walked you last week through the steps of model estimation, where you had the outer loop and the inner loop.

Now, what is important here is that if you include fixed effects, you really want to make sure that in the inner loop, you're running a regression to back out those fixed effects. And you separate the linear parameters from the nonlinear parameters. To remind you, the linear parameters are all the things on observed or unobserved product attributes, in this case, with fixed effects that are only J specific and do not depend on the consumer. And so, again, fixed effects introduce a lot of parameters. If you can just optimize over your nonlinear parameters and then back out the fixed effects through a regression for any given guess of those nonlinear parameters, that's going to help you a lot.

So that's one comment on this. The second comment is that it's a little bit against the spirit of these characteristic space models to estimate fixed effects. Not to say you shouldn't do this, or that this is not a good idea. This can be very useful for identification, but you are introducing new parameters for each new product, which is something that we try to avoid.

So one thing that we had discussed last week, but that's new relative to BLP that I want to point out here, is that the random coefficients are now not just-- there's not just the standard deviation of some mysterious unobserved distribution. They're now written in terms of demographic variables of distributions that you actually observe for each market-- so the distribution of income, distribution of family size, and so on. So I'll come back to this once more when we discuss the estimates later on.

The market here is defined as a city quarter combination. I had some questions about market definitions. Vat, you asked about this, right? That's a good question. So the question was, why do we define it this way? Why don't we take commuting zones, or, let's say, a quarter, or a half a year, a year?

That's typically a difficult choice to make. What could happen is that if you don't draw the right boundaries, that firms at the edge of your market are actually competing with firms in other markets. And you mis-specify the number of competitors that they're responding to. And so there's some papers that use clustering procedures to figure out how to separate markets, but in general, this is a bit more an art than a science to think about how you how you designate markets. But yes, that's an important question.

So one more issue that comes up with including those fixed effects is that brands are not-- or products are not changing their attributes much over time. Which means that those instruments are completely absorbing the variation that you would be using when you build the BLP instruments. So remember, the BLP instruments, it was this idea that if I'm offering a very sugary product, and I'm in a market where all the other products are also very sugary, then I face more competition. I'm probably going to respond to that in my prices.

So what he does to get around this is to introduce two other instruments. First, the Hausman instruments, which we had discussed last week, which is basically the idea of using average prices in other markets for the same firm as proxies for costs. So I've got some questions about this, so I just want to ask, does somebody want to explain again what the idea behind those instruments exactly is, and what are some of the issues that could come up here?

So again, the problem is that in many cases, typically, costs are not observed, or it's hard to get at those. Remember also that-- just picture a very simple linear supply and demand picture, and the identification problem that we had discussed at the very beginning of the last lecture, that you ideally want some cost shifters that shift around the supply function to identify your demand function. Again, cost side instruments that you believe are excluded from the demand side are typically the most convincing types of instruments.

So instead of using directly observed costs, the idea, again, behind these Hausman instruments is to proxy for those costs with average prices of the same firm in other markets. And the idea is, if I'm, let's say, the only cereal producer that uses rice as an ingredient, and the price of rice goes up, then that's going to shift my average price. And to the extent that the shock to the cost of rice is uncorrelated to demand shocks across different geographic regions, I can use this as a proxy for a cost shifter.

Now, this already leads to where the error where this might go wrong is if you have things in those prices that are actually correlated across different markets. So think about advertising campaigns. If I'm running a big advertising campaign that leads to more demand for my brands, and that's going to be potentially reflected in the prices, I mean, to the extent that we think that it is a marginal cost. And then it's as if I'm instrumenting my endogenous-- as if I'm instrumenting prices with demand. That's not a good idea.

So the other instrument that he uses are what he calls cost side instruments. They have a somewhat similar logic. So he uses region dummies to pick up things like transportation costs, city density. Again, this is an issue to the extent that these things might also be picking up systematic variation in preferences across regions.

Any other questions on this? You have read the paper, so you're curious. Yes?

AUDIENCE: So one of the things that sort of-- that utility formulation is that the price sensitivity varies across consumers, but it doesn't vary by type of good. And so when they present the table with price elasticities between-- cross-price elasticity between types of goods, that's only driven by the type of consumer that's buying that good, right?

PROFESSOR: OK.

AUDIENCE: So I guess my question is, how reasonable of an assumption is it that all of the prices, cross elasticities that we estimate are only driven by differences in consumers, and not in types of goods that we're buying?

PROFESSOR: OK that's an interesting question. So let me try and say a few things. So first of all, I mean, think first about the idea behind the characteristic space approach is that people value characteristics. And so if you really believe that that's the way you construct your models, then you shouldn't actually have random coefficients be product specific. So that's one answer to that.

Now, it might be that they're a more flexible, functional form where, let's say, your arrow is not additively separable leads you to a specification like this where your random coefficient is essentially J specific because you're multiplying either alpha or beta with this unobserved term. So I think you would probably need-- I would have to think about, you would probably need some, again, variation over time to estimate such a model because you would supposedly want to see how does the same consumer vary attributes across different products. And you might need microdata for that, but I'm not exactly sure.

AUDIENCE: Thank you.

PROFESSOR: OK, any other questions? So let me briefly talk about his table where he estimates, again, as a logit model, as I explained before, and that's a good idea. And he has a lot of different specifications.

The first three here are without any use of instruments. These are all instrumental variable specifications. And I think these two here are including fixed effects as controls-- so product fixed effects, whereas the first one does not.

So what you can see is with fixed effects, you explain a lot more of the variation. And the price coefficients get more negative. And even more so when you include consumer demographics.

So this here is an interesting instrumental variable specification that some of you pick this up what this is doing-- the inclusion of brand dummies as instruments. What is this close to in spirit?

AUDIENCE: BLP.

PROFESSOR: BLP instruments, right? Could you think of a story where this is preferable to, let's say, the simple BLP instruments? Or why would you want to do this? OK, so--

AUDIENCE: Is this what's captured on, like, advertising effects with a brand?

PROFESSOR: Exactly. OK, so let me answer this in two ways. This is exactly right to the extent that advertising is not captured in the utility specification, which it isn't here. You may be worried that the BLP instruments are-- they're picking up that I'm responding to the fact that the average sugar that's in my competitor's brands is very high-- that I'm responding to that in some form, or the mushiness.

But it might be that I offer the Donkey Kong cereals, and my competitor offers the Donkey Kong Junior cereals. And advertising, kids see the ads. And, of course, in all the markets where I'm competing with Donkey Kong Junior, I have to compete much harder. And those brand dummies would pick that up.

So they capture not just the way that firms are responding to observed product attributes. They could potentially also capture other things that the researcher does not directly observe. So that can be useful.

But of course, you cannot include those and also directly control for them. So that's why he uses here, these are indicating the Hausman instruments. So average prices in other markets, as well as these cost side instruments.

And so if you look at these specifications here, you get even more negative price coefficient. I think starting somewhere here, he finds no longer that there are any elasticities on the inelastic portion of demand. So I think, again, this is something useful. It gives the reader a sense of what is driving your estimates, and it gives you a sense of what is driving your estimates. And that's producing these tables can be very useful.

So let me say something briefly about domains table of the main specification. Now, these are the estimates from the BLP model, what he calls the full model. So here is a similar observation as before. Somebody was disappointed that people didn't like mushiness on average. I was also a bit disappointed by that because I also like mushy cereals, but it seems like the top 15% in the distribution of that.

So that's actually somewhat interesting story here that he had to actually measure that. And it was one of his classmates, Sandy Black, who suggested including this as a characteristic. But again, the point here-- the point that is important-- is that you see here the demographic effects on this are large. And depending in which demographic you are, you actually value-- the sign on this attribute flips. And again, shows that people can value different attributes differently.

The one substantive point that I want to make here is that if you look at what drives the random coefficients, you see that many of the demographic interaction effects are significant. But many of the standard deviations which capture this unobserved term in the random coefficient-- which, in the BLP random coefficients, that was the only source that led to variation-- are insignificant. And so this suggests that much of this variation, which was completely unobserved variation in BLP, is, in fact, in this market, driven by demographics.

Now, I'll let you be the judge whether you find these interaction effects convincing. It seems like higher income people like mushy cereals more. I don't have very strong priors on that.

So what does he get out of this at the end of the day? So let's focus here on his preferred BLP specification. So to remind you, the exercise that he engages in is I see some mark ups in accounting data from the census of manufacturing.

And for a second, let's say these are the true markups and they do not have the kinds of issues that we talked about last week that they don't account for things like opportunity cost and so on. So these are not necessarily the correct economic costs that are measured in this accounting data. But let's say for a moment these are the true markups. So the idea is I estimate my demand model, use different conduct assumptions, and then see which conduct assumption replicates the true markups in the data. That's the essence of this exercise.

And so what he finds, so he has this range of markup estimates depending on whether or not you include advertising in marginal cost, which ranges from 31% to 46%, he finds that you can soundly reject this collusive model where, for the 25 brands that he includes in his estimation, firms jointly maximize profits. And he cannot reject any of the alternative models, which is either firms-- or that each brand competes separately. Prices are set at the brand level without internalizing these cannibalization effects as well as the current ownership, where he uses the actual ownership matrix that he observes.

One interesting thing here is that with the logit model, you get, in general, much, much lower markups. And you would reach a very different conclusion. You wouldn't reject the collusive model. So this, again, shows you that for this substantive question, which is, are firms in this market colluding? It seems to be important to account for these more realistic substitution patterns where different people value product attributes differently.

So I think this is a well written and explained paper. In part, I asked you to read it because it provides a good explanation of the BLP model without many of the technical aspects that we will discuss down the line, and you know, in part, Aviv Nevo is also known in the profession because he has written this practitioner's guide at the very beginning that actually taught graduate students how to estimate these models. And he put his code out there, so it was very useful for people to understand what is going on.

This is, in terms of thinking about picking an industry and the right model for this industry, what I really like about the paper is it would be a bit crazy to estimate a model for this industry without accounting for these multi-product considerations. At least as a prior, you'd think that it would matter a lot. You have four large firms that account for almost the entire market, and, there are hundreds of different products. So you have to do something about that.

Collusion is, obviously, one of the important topics in IO. So it might not speak to your preferred industry, but it speaks to a substantive economic question. You could bicker that advertising is not a strategic variable. We know, actually, from the paper that advertising is very important in this market, and I imagine it's particularly important when you introduce new brands. You want to make them known to kids that there's a new Donkey Kong-- the new Donkey Kong cereal.

So I guess another potential criticism is would firms actually be able to implement this, from their point, first best solution of full collusion across all the different products? And so is that the right comparison point? And then, as I already said, the testing approach is predicated on observing the right accounting cost. And that's not something that-- you know, once you use that as a benchmark, you can obviously no longer say anything about that.

Any other comments on the paper? Yeah?

AUDIENCE: So in some part of the paper, he argues that one shortcoming of the BLP literature is that you take observed product characteristics as given or as exogenous, and then you should internalize those because firms actually choose which characteristics will [INAUDIBLE] target.

[INAUDIBLE]

So in his model, is he implicitly allowing for firms to choose certain product characteristics? Because I don't think he's improving it.

PROFESSOR: No, he's not doing that. The only endogenous characteristic is price in this case. And again, advertising would be a very good candidate for that because that's something that we potentially think varies from, let's say, quarter to quarter, city to city. And there are papers that look at advertising effects, and do use advertising as an endogenous variable. Then, of course, you have to think about additional instruments. But yeah, that's not something that he tackles in this paper.

So let me use the last 20 minutes to talk about something that-- we're staying in this industry, but moving to a new topic. Which it's a fairly recent debate that came up in IO, which is called the common ownership hypothesis. It's something that's fairly controversial and hotly debated with strong proponents on either side of the argument.

So what's the idea? The idea is that we have seen this big trend towards diversified investments where people hold large portfolios of the entire, let's say, S&P 500, or stock indices. And typically, they purchase those through large institutional investors like Vanguard, BlackRock, and Fidelity. And so a lot of people use those to save for retirement and so on.

And the idea is that these owners, they might hold-- or these institutional investors, they might hold significant shares in a number of different companies that are actually competing with each other. So here is an example of the US airline market. You have American Airlines, Delta, Southwest and United, the major airlines. And you look, who are the biggest shareholders of these companies?

Vanguard holds 6% of American Airlines. It holds 6% of Delta Airlines. It holds 6% of Southwest, and 7% of United Airlines. And so the idea here, this hypothesis, is that if I'm holding stocks in a bunch of companies that should be competing with each other, who says that. I wouldn't be very interested in making these companies compete a little less hard and raise my profits? Moving the incentives closer to monopoly incentives.

So this is something that people had thought about in literature at the intersection of corporate finance and IO. There are some theory papers that I'll talk briefly about just to clarify what this hypothesis is about, but it wasn't really until very recently that this was-- that people paid a lot of attention to this. And of course, in part, this attention is driven by the fact that we have, now, these large diversified investments that have been trending up.

So there was a series of papers in 2016, 2018, now there are many more, that document that the growth of large, diversified common owners led to increase in prices in various industries, for instance, for banking services or for airlines. So these were actually people that-- this was first researched in finance. And so it was a little bit awkward for IO people. Could we have possibly missed this, that this big structural force that led the whole economy to become less competitive? And do we really think that the price that we pay for these cheap, diversified investments is that monopoly power as a whole has increased in the US economy?

So these papers suggest that it did. Now, they use an approach which I think Glenn has talked about, which is a structure conduct performance approach where you essentially regress prices on measures of concentration. And we know that this has well-known issues, in part because these concentration measures are themselves driven by quantities which are endogenous. So today, we're going to talk about one paper that tries to do this in a structural way, and a companion paper to that that documents a little bit to what extent this is an actual issue in the US economy.

So just very briefly, what is this idea here? Formally, the common ownership hypothesis, according to this paper by Rotemberg-- and here, I'm using the notation of O'Brien and Salop-- is the idea that we have an investor, S, or different investors, that own a portfolio of shares in different companies. And according to the amount of shares that they hold, they are entitled to some fraction of the profits of each of the firms that they hold.

And you have firms that, in this model, maximizing the value of the shareholders. So how do they do that? By virtue of having different portfolios, different shareholders might have slightly different interests. Their interests might not be fully aligned.

So in this model, the way this is handled is that we're saying the firm places different weights on what different shareholders want. So these are pareto weights. So you can write the firm's objective function, which is a function of its own prices and other prices in the market, as a weighted sum over what its different shareholders want.

Now, if you plug this in here using this formula from above, the shareholder objective function, the portfolio, you find that you're suddenly in a situation where this firm, F, might be caring about the profits of some firm, G, in the market through the incentives of its shareholders. So these are the pareto weights. These are these cash flow shares that the investor, S, is entitled to terms of firm G.

So you can rewrite this. You can do some algebra and normalize the share of firm F to 1, and can write the objective function of firm F as its own profits plus a weighted sum of other firms in the market. And so we call these profit weights here kappa. They're a function of these pareto weights and these betas.

So just to link this back to what I discussed at the very beginning of class today, this should remind you of something that we have seen on the first slide. So here, we have a firm that's maximizing its own profits plus some weighted sum of other firm's profits. This is very similar to this idea of a multi-product firm that owns different brands, and will take into account that if I lower my price on one of my brands, that might potentially cannibalize one of my other brands and divert some market share. So here, you get a very similar first order condition for firm J that places these profit rates that come, again, through that are a function of the shareholder holdings on other firms' profits. So I hope this connection is clear.

So what can we actually say about those profit weights? So this is what this companion paper by Backus, Conlon, and Sinkinson does. The first document, is it, indeed, true that, just as a pure matter of accounting, that firms are placing, or have these incentives to place more weight on what their competitors are doing, or any other random firm in the market?

So first, if you look at the time series of what fraction of the typical firm is actually held by one of these large institutional investors-- BlackRock, Vanguard, and State Street-- we find that this did, indeed, go up a lot. And, I mean, this is actually not just-- this trend towards diversified investment is not just driven by these institutional investors. This was another result in the paper was already going on even before that. So they use some data from SEC filings that large investment managers have to submit to document this.

So then, they compute these profit weights that I've just shown you that you can derive from the firm's profit maximization problem, taking into account the incentives of the shareholders. So the one issue that they have here is they don't actually see this pareto weight. So they make some assumption that these pareto weights are a function of these shareholder weights beta, and so they show how this time series look for different functional forms. But the general message is-- and you can also break this down by industry-- that if you take a firm in the US, it's a publicly traded firm in the US economy, and ask how much weight does that firm place on the profits according to these diversified investments on any other random firm, that this has gone up over time.

So this suggests that in principle, these kinds of incentives could be at play. Any questions here?

AUDIENCE: Is there any sense of how sensitive these measurements are to the pareto weights? Like were there different specifications tested, or [INAUDIBLE]?

PROFESSOR: This trend is not very sensitive to that. The levels are-- I mean, in the paper, they show you, I think, three different types of weights. But the general message is that this has been upwards trending regardless of-- I mean, you can then, of course, argue, is that an exhaustive search over the potential specification that you could use? Yeah.

So coming back to the cereals market. So again, what I've just shown you is from this descriptive paper is that they document that these incentives, the fact that firms may internalize profits of other firms that they're competing with could potentially be an issue based on the time series that we observe because of this trend towards more diversified investment. If you look at the breakfast cereal market and the different producers here, you'll find that, indeed, BlackRock also, in 2016, held a large share of General Mills, of Kellogg, of Quaker Oats, and Post brands-- so all the big producers. So in principle, this could be an issue in this market, as well.

OK, so what is it that they're doing? So I'm going to be-- you know, I'll give you the main economic idea here. This paper, big part of the contribution is develop a new and efficient testing procedure for this and allowing for flexible cost function. So I'm not going to talk about the econometric details, but this is very much in the same spirit as the other papers that we have been discussing-- Nevo's paper on conduct testing.

So the idea is that with my demand model in hand and conduct assumption, I can invert my first order conditions as a function of the observed market shares, prices, and the demand model. Again, the demand model, of course, has some demand parameters, and can back out my effective marginal cost. At the same time, I may be willing to also assume a functional form for the marginal cost directly. Again, because I know that the cost of producing cereals are driven by the cost of sugar, wheat, flour, and so on.

So what they do is they basically specify this equation here by allowing for a flexible supply function or marginal cost function. And based off this, get additional moment restrictions that they use for the construction of a GMM objective function under different conduct models, M. And of course, in particular, the conduct model that they are interested in is the common ownership hypothesis. And that is, again, very similar to this idea that I might be internalizing the profits not of other products that I hold, but of other firms through the incentives of my diversified investors.

So the idea here is to construct this equation for different conduct models. You get some GMM objective function out of that. And with this GMM objective function, you can construct a test statistic.

This here is a test that's-- it's a non-nested test by Rivers and Vuong. The idea of this test is that there are two different models of conduct, and the null hypothesis is that both of them are equally far from the true model. But you're always testing the two models against each other. So what, at the end of the day, you can say is, is this model better describing the data that I observed in some other model?

And so, for instance, here, based on, let's say, common ownership or own profit maximization, if this test statistic here becomes very large, then I'm rejecting model 1 in favor of model 2. Because the violation of the moment conditions is much larger under model 1. This is larger.

So it's shown that this is asymptotically normal. So you can use just standard criteria for the normal distribution to test with this test statistic. So just, again, to link this back to what we have already discussed today, so like Nevo, they estimated the demand model for cereals. Do this first order condition inversion. And like Nevo, they test for different models of conduct with a focus on common ownership.

But unlike Nevo, what they do is instead of comparing the markups that I get from this inversion here of the first order conditions directly to observed markups, they're seeing whether the marginal costs that they back out are consistent with the functional form that I imposed on marginal cost. And they basically now used demand side variation-- so things that move around the demand side but I exclude it from the supply side-- as overidentifying restrictions to test this model and construct this test statistic. And this allows them to compare, let's say, the common ownership model against other types of models.

So let me tell you what they find from this. Again, there's some details that I haven't touched on, which is how do you specify the supply function, and what are the best instruments? Their preferred instrument are these optimal demand side instruments.

Again, now you want to say something about this flexible supply side, so you use demand side variation which moves around the markup. And under the right model, these are excluded instruments-- instruments that are excluded from the supply side. These should be orthogonal to the variation in omega.

And so what you find here is that using these optimal instruments, they soundly reject your different models of common ownership against own profit maximization. And so there are different specifications whether or not you allow for firms responding to a lack of these profit rates. They also reject perfect competition using these optimal instruments as well as a full monopoly model. So it seems that, against what these time trends suggest, that firms do not seem to be internalizing these common ownership effects. And instead, the data is much more consistent with own profit maximization.

So this is, of course, just a case study. So this doesn't speak to the US economy in its entirety. So again, to wrap up, this is a very interesting hypothesis that has been very hotly debated in IO, and people feel very strongly about it.

You could plausibly conjecture that firms behave in this way based on these time trends that we have observed. It's not crazy to think that they somehow internalize this. This case study rejects this.

And more generally, people have been highly skeptical of this hypothesis, in part due to a lack of a plausible mechanism of how these firms would actually-- how these institutional investors would actually bring the firms to, let's say, soften competition, or engage in less price competition. So the mechanism is still a bit elusive, and, to the best of my knowledge, there has been no smoking gun-- like a story where it's clear that they have done this.

Some people argue that managerial incentives could be a potential explanation for this. I think this is an exciting topic that you should make up your own mind on. And that's all I have for today.