[SQUEAKING]

[RUSTLING]

[CLICKING]

GLENN ELLISON: Today, I am picking back up with a gap in between on advertising. And so the plan is today, I'm going to talk about empirical work on advertising. And then talking about empirical advertising, it's a big challenge for a class like this because marketing is a field unto itself. And I don't know if there are half as many marketing economists as there are economists of all-- you know, economists of all fields together.

You know, much of marketing is an overwhelmingly empirical field. If you look at what marketing faculty do, you know, there's multiple motivations.

One is there is tremendous demand from practitioners to understand the value of advertising. You know, this is what advertisers want to do is know how valuable the advertising they're buying is. There are many consulting firms who try to help them figure out how valuable the marketing is, and there are many academics who try to help the consultants try to help develop new cutting-edge techniques for how you would measure the effectiveness of advertising.

But then there's also a substantial marketing literature that is about basic economic questions. You know, as I said, we have these different ways in which advertising could be effective, whether it's changing tastes or whether it's providing information or whether it's signaling something.

And you know, I think we want to know why the-- sometimes those effects, the welfare effects were different depending on what advertising was doing, so I think there's also academic interest in, what is marketing-- what is advertising doing? And what are its general equilibrium effects on the markets involved?

So you see many different methodologies going on. Marketing faculty at places, top places like Sloan, a number of them have relationships with advertisers or with advertising firms that let them do sophisticated experimental work to evaluate advertising with that advantage. There's also-- you know, if you look at IO, it's not like IO is really ahead of marketing in structural demand models for understanding consumer preferences and how A, affects consumer utility. So there's a lot of complicated structural stuff.

What's actually sort of funny is the more recent trend in marketing is there's actually been this recent rise in just plain vanilla causal inference techniques and using plain vanilla causal inference techniques to try to evaluate advertising. And in some ways, when you're trying to talk about, what's the general equilibrium effects of advertising, like, if you allow advertising in a market, how does it affect market prices, you know, you can't do that experimentally because it's a market-wide question.

And so for those market wide questions, we're seeing more people do just simple difference-in-difference and other causal techniques to try to figure out, what are market-level effects of advertising? I don't know. Catherine Tucker in Sloan, she has a recent survey article discussing a number of recent papers that have done this.

So I'm going to start with, first, one of the earlier of these difference-in-difference papers. This is Milyo and Waldfogel. So you know, the paper title is "Effect of Price Advertising on Prices, Evidence in the Wake of 44 Liquormart." 44 Liquormart, I think, is the name of a Rhode Island liquor store.

And the backstory of this is from 1956 to 1996, Rhode Island banned all advertising of alcohol prices. So these kinds of circulars you'd get in the mail of liquor warehouse mix and match. Buy this many bottles, you know, Sam Adams, $7.99 a six-pack, this was all illegal in Rhode Island.

This was also illegal in Rhode Island. You could not in the state of Rhode Island have a sign in your liquor store window visible from the street showing how expensive the beer was. So it was a just complete, total ban on any advertising of alcohol prices.

And a liquor store, I think called 44 Liquormart, sued the state of Rhode Island, saying that this was a ban-- an infringement on their freedom of speech. And in 1996, the Supreme Court overturned the law as an unwarranted restriction on freedom of speech. And so starting on that one date, May 13, 1996, advertising alcohol was legal in the state of Rhode Island. And so what Milyo and Waldfogel were trying to do is do a difference-in-difference approach to see, how does that affect alcohol pricing in the state?

And I can say there's actually-- in some sense, they're building-- there's an earlier paper-- there's a classic paper from the 1970s by Benham on advertising eyeglasses. So some states allow you to advertise eyeglass prices, and some states do not. And Benham has this classic paper from 1970s, basically just a cross-sectional regression showing that glasses cost 20% less in states where you're allowed to advertise eyeglass prices than ones that you don't.

And so I think that was-- part of the background is, is there really a big-- going to be a big effect on alcohol prices of allowing advertising? Which, if your view of advertising is this is reducing consumer search costs and consumer search costs are big part of markups, you could think advertising would have a big effect there.

Anyway, I think the ambitious thing that Milyo and Waldfogel did on this paper was they gambled on the result of the Supreme Court case. And basically, they just-- you know, it was Rhode Island in the 1990s, it wasn't like every liquor store was big and had supermarket scanners and you could make contracts with a scanner company to get all the prices.

So what they started doing is just manually collecting prices for 33 different alcoholic beverages at 115 liquor stores in June 1995. So this was basically-- the Supreme Court announced that it would take the case in June. They had no idea when the case would get taken, what the decision would be. If the law was upheld, they would have had no paper whatsoever. But anyway, they just started gathering data on liquor store prices in anticipation of-- in anticipation of the law being overturned.

And to do a difference-in-difference thing, they did a control group of Massachusetts stores. They picked Rhode Island stores and then Massachusetts stores, some of which are near the Rhode Island border and some of which are not. And then they can do a difference-in-difference and compare, how do prices in Massachusetts and Rhode Island change when the liquor store-- when the advertising ban's overturned?

In addition to these-- and let me say, the price data collection is pretty impressive. It's like they just went to the liquor stores personally. And this is something I think also mentioned in Alan Sorensen's paper. You can't take a clipboard into a liquor store and start writing down the prices of everything that they sell and walk around for 15 minutes, writing down all the prices they sell. They kick you out of the store.

So what they actually did is they would go in there, and sometimes they would have a set of prices. They'd then try to memorize them, and go-- walk in, memorize them, and walk out. Joel also has these stories of he would bring his child with him. And you can spend longer if you've got a four-year-old wandering around the aisles than you can as a person and put a voice recorder in his pocket and just read out to himself, like, the Bacardi's $8.99 or whatever until he collected the prices.

Anyway, they did this for a year beforehand. They did it for a year afterwards. That's the main data set. But then they also have several other things they were able to find. They have Massachusetts and Rhode Island both have regulated wholesale prices. They got the wholesale prices in the stores. They gathered data on whether the Rhode Island stores used print advertising, whether they used window advertisements.

And they also got lottery ticket sales. And the idea was that they didn't have any way to gather quantity that the liquor stores were selling, but they could get data from the state on how many lottery tickets were sold at each liquor store. And they're going to use the lottery tickets as another prox-- a rough proxy for quantity, figuring that if more people come into the store to buy alcohol, more people are going to buy liquor-- buy lottery tickets at the stores. And that they could get for all stores that sold lottery tickets, which is a lot of them.

So here's a rough schematic of what they were doing. So their early data collections, they're visiting 20 stores in June. Then they do a bigger data collection in September and visit 30 or 40 stores. They get tired waiting around for the law to be overturned. They only visit 10 or 15 in February. And then once the law is overturned, then they ramp up the data collection. They're collecting from 40 or 50 stores in each wave.

Advertising. You can see that advertising didn't actually get all that big. So there are-- you know, in their sample stores, it's like three, one, five, six. So it's less than 10% of the stores are choosing to advertise once advertising is allowed. Anyway. So then they collected by hand whether the stores were advertising, but it was not-- you know, many stores which obviously had not been advertising for 40 years just went along happily not advertising once the law ended.

Here's an example of-- these are the products that were in their study. So they have a bunch of different standard liquor items. They have several different beers. They have a bunch of relatively-- a bunch of relatively cheap wine. They have a few champagnes and sparkling wines.

And again, the goal of the data collection was to walk around the store and get all these prices. If they couldn't get all these prices, they just got some subset of these prices. And that's why you can see the number of observations differ a lot across product.

OK. So what does-- what does advertising do? So result 1, you know, they're going to estimate the effect of alcohol-- the end of the alcohol ban. Simple difference-in-difference estimate. You just do prices in Rhode Island after minus price in Rhode Island before minus price in Massachusetts after price in Massachusetts before. It's a regression with time fixed effects and state fixed effects and then a Rhode Island times post dummy, and that's the thing you're estimating.

And what they find is that, you know, the price was very low-- the difference was very, very low. So you put in the state product and the store fixed effect, and the prices on average in their sample are going down by, like, half a percent to 1% after the advertising ban comes into effect.

So it's just advertising leads to very, very slightly-- allowing advertising leads to very, very slightly lower prices. And this is-- you know, they're tracking it for one year afterwards. So you get one year afterwards, and within a year this effectively had very little impact on retail prices.

Breaking it down product by product, how do advertisements work? When stores choose to advertise, they advertise much lower prices. So the advertised items at an advertising store are roughly 20% lower than the prices of those items at the Massachusetts stores or the non-advertising Rhode Island stores. So stores that advertise do set very low prices.

And then your question might be, what's the aggregate effect of the advertising? You know, as I said, I did a loss leader model saying, like, you know, you're going to advertise very low prices for turkeys at Thanksgiving, and that doesn't mean that Thanksgiving isn't the single most expensive day of the year to shop because you can just use that as a loss leader, bring people in, and then advertise-- raise the price of everything else.

That's not what's happening. So the non-advertised prices at advertising stores are within a fraction of a percent of the prices elsewhere. And if anything, they're slightly lower. So it's not like the stores advertise and jack up the prices of everything else. What they do is they advertise low prices and basically leave the prices of everything else unchanged, or if anything, slightly decrease them.

And the non-advertising stores in Rhode Island, their prices are now going down by, like, 0.2%. So whatever the competitive effect is, the non-advertising stores are just not reacting-- they're not reacting to the advertising of their rivals other than this-- you know, just very, very tiny cut in average prices.

Other things in the paper. So it wasn't in the table. Rivals don't respond to advertised prices. So you advertise, I've got Bud Light, 20% off this week. Your rival stores don't have Bud Light at 20% off. They basically just don't react at all to the advertising.

On the quantity effects, advertising stores' share of lottery ticket sales went from 16.4% to 18.4% after the ban ends. So if you remember, it looked like the advertising levels were about 10%. So the advertising stores were bigger stores. They were selling more than 10% of the lottery tickets, but they go from 16% to 18%, so that suggests that their quantities are going up.

But again, if you think about-- so the fraction-- the half a percent or the 0.7%, that was the price on a just item-weighted basis. You could-- your question would be, well, OK, on a sale-weighted basis, could the drops in prices that consumers pay be much larger?

You know, this is suggesting that it's only 2% of the demand that may be moving. And so if only 2% of the demand may be moving, the aggregate effect on consumers is probably small. Now, if that 2% were all getting stuff that was 20% off, you know, still, 2% of the people buying stuff that's 20% off, that's, like-- what is that? 0.4%? So it's not a very big effect.

A second thing that they note is that the stores that eventually advertise were charging 7% less before the advertising ban was eliminated. So these larger stores were the larger, cheaper stores to begin with.

So again, if 2% of the people are moving and buying stuff that's on average 7% less, that's, again, adding a few tenths of a percent to what consumers are saving. So I think in aggregate, it looks like this shift-- consumer shift to lower priced stores presumably within the advertising stores, consumers buy the things that are the 20% off price levels. But on aggregate, average prices in an unweighted sense don't change. And on a-- whatever, consumers are better off sense, consumers are making a couple percent off this.

So I don't know. I don't know if this makes us think that all of these search costs for consumers, at least for buying beer and wine and liquor, are not such a big deal or such a big determinant of markups, or at least the advertising is limited effectiveness in changing those consumer search costs. Questions? Nope? OK.

So I thought next thing I would do is go to the other extreme in data and just give you a sense of, you know, what are the kinds of papers that people are writing with good access to-- good access to firms that allow them to do experimental work?

So this is a paper by Randall Lewis and David Reiley, "Online Ads and Offline Sales, Measuring the Effects of Retail Advertising via Controlled Experiment." So at the time that they started working on this paper, Randall and David were both employees of Yahoo or working at Yahoo research.

I think-- well, I don't know. This is-- I guess it's funny to be lecturing a class of people so young they probably don't really remember Yahoo. But in the year 2000, Yahoo had 25% of all page views on the internet. So if you think about, what do people do when they're on the internet? And the year 2000 was people were on Yahoo.

And it was like, you know, Yahoo was just like-- Yahoo was what Facebook and Google and Amazon-- whatever. Like, at the time in 2000, Yahoo was the dominant online firm and people are like-- Yahoo was their home page, and they go through Yahoo, and they would look through Yahoo shopping Yahoo mail, Yahoo whatever.

This is what the Yahoo front page looked like. Yahoo front page, it was kind of like a news site, your starting page. You can see over here, you had other things you could do on Yahoo like Yahoo Autos, Answers, Jobs, Groups, Mobile Web, Movies, TV, whatever. But it was mostly like a news page.

And then they would have these display ads like this one. This is what they call an LREC ad, which stands for Large Rectangle. You'd have this large rectangle on the front page with some ad. You then have other smaller ads here on the front page. And as you went through, if you read the news pages or whatever, there would be more ads like this appearing on subsequent pages. And Yahoo made its money by selling these.

How did those get sold? Yahoo sold those ads in multiple different ways. The front page ad there was actually sold by the day. So every day, there would be a new advertiser on the Yahoo front page. Some days it would be a 50/50 split. 50% of page views would be ad A. 50% of the pages would be ad B. Other days, it would be everybody. But basically, you could just run a one-day ad campaign on Yahoo and they'd sell you that LREC on the front page.

When you got to the inside pages on Yahoo, there were many, many, many different ways to bid for ads. Like, you could buy an ad that would-- if I'm Netflix or whatever, I could buy an ad that appears on pages that mention a movie title. And like, I want to buy movie title.

Or you could buy ads targeted to consumers. I want women 18 to 39. I want men 25 to 54, and I'm willing to bid this much if it's someone in that category.

You could bid re-targeting kinds of ads. Like, I don't know if you've ever noticed. You look at a pair of shoes on Zappos or something, and then those shoes follow you around the internet. You can bid, I'm paying $5 per 1,000. Anytime someone who's visited my site and puts something in their cart and not bought it, I'm going to buy an ad and target it there.

So you can target by the individual characteristics. You can target by the content of the pages. You can target on the person's past behavior using cookies, but they would sell all those things. This particular experiment they do, they're trying to assess the effectiveness of Yahoo ads-- ambitious project trying to assess the effectiveness on Yahoo ads on both online purchases and offline purchases, which is something that's usually much harder to get at.

So how did this experiment work? So they run a randomized controlled trial. So they have some large offline firm that's interested in running a big ad campaign on Yahoo. And they propose to the firm that they're going to evaluate the ad campaign as well as running the ad campaign for them.

So the retailer-- so think of this as a firm like Macy's or Bloomingdale's, some very large department store, probably. This large retailer had a data-- had a data set of 1.5 million customers that they wanted to advertise to. So they're going to display the advertising only to those people.

And then what they did is they matched it to the Yahoo database. So you know, like today, Google and Facebook and whomever know who you are when you're searching. People had Yahoo Mail accounts. They were permanently logged into Yahoo.

So they got someone to match this database that the firm had of people and their credit card numbers and their home addresses to Yahoo's membership database. And so they were able to jointly identify 1.5 million people who both existed in the retailer database and had Yahoo accounts that they may have been logged into when they were web browsing.

They used an 80/20 split of treatment and control. So 80% of the people, they tried to advertise to them when they saw them. 20% of the people they held out as controls and they never advertised to them.

One thing about the treatment control design like this is, again, the retailer had a bid that it was willing to pay to reach those people. And so I would guess, early 2000s advertising prices, it may have been, like, $1 per 1,000 or something like that for impression ads. So that every time-- every time one of these ads was displayed to a consumer, they owed Yahoo 1/10 of $0.01.

Whether those ads actually got displayed depended on what else that person-- whether that person was actually on the internet and what else they were doing. So if that person was on the internet, only rarely they might not see an ad. If that person was on the internet but they'd recently looked at shoes or something like that and had a re-targeting ad that wanted to follow them around, that re-targeting ad would outbid the Macy's ad and they would never see the Macy's ad.

Or it could be if that person was in a very valuable demographic group, somebody wanted young people, the young people price would be higher than Macy's bid also. So whether or not you see an ad, you know, it's like an intent to treat thing. Like, being in the treatment group means Macy's is bidding to get your their ad shown to you, but they may not actually get their ad shown to you.

And then the impressive part of this campaign was they're able to monitor the purchases by these people through the online-- so this store is a store that has an online presence, but they have much bigger offline sales. And they will monitor purchases both online and offline by those consumers.

So if somebody who saw one of these-- a Macy's ad, walks into the Macy's store, they buy something. If they use their credit card, Macy's is going to know who they are. They're going to record all those sales, and they match those sales back up to what you saw on the internet five days earlier.

So what if things-- campaigns looked through like this? You know, again, you're putting those relatively large ads on there. I don't know what the ads said. The ads could have said, Macy's sale, 15% off all purchases with this code, something like that. The ad got a 0.28% click-through rate.

So when people were displayed an ad, there was a probability-- you know, 28 per thousand-- 28 per 10,000 that they would click on it. 7.2% of those who saw at least one ad clicked on an ad during the course of the campaign. So they were advertising people many, many times. So you can get a sense of the scale here.

So there were two waves of the campaign. There was an early fall campaign. Then there was a late fall campaign. The early fall campaign was the main one. So they displayed 32 million ads to 814,000 people.

So if you remember, the intent to treat was 1.2 million. So 2/3 of the people-- 1/3 of the people either never got on the internet or got on the internet only rarely and Macy's never made it to the top of their bid queue to display an ad to them. But the people who were on the internet, they saw an average-- if you conditional on seeing at least one ad, you saw an average of 39.6 ads. And then conditional on seeing at least one ad, 7% chance that you clicked on an ad at some point during the campaign.

And then the second campaign was a bit smaller. You can see it's about 1/3 ads per viewer. 13%. Reach was only slightly less, 56 versus 63. And you'll notice the mean ad views per viewer isn't 39 plus 13. It's a little bit less because there are some people who only saw an ad in the first campaign and some people who only saw an ad in the second campaign.

OK, so what are my observations? The first one is actually-- sometimes you'd write a paper, and the paper kind of fails. And then in some sense they resurrect the paper. It's like, the failure of this paper is notable that, like, we get very imprecise estimates.

And part of their thing was that us going into this maybe as naive people, we thought, we have this massive experiment, and we're going to target a million people and send 30 million ads at them. We will be able to get very precise returns of the [INAUDIBLE] advertising.

And one of the first observations is that estimating return on advertising is really, really hard, and it's really hard even with advertising-- with campaigns of this scale. And so why is that? Here's a simple calculation.

So first thing about the sales data is that sales are very, very noisy. If you think about-- like, what does Macy's care about? I don't know that the store is Macy's, but just a store like Macy's. What does Macy's care about? It's how much people buy from them, and so they care about the dollar value of purchases.

What do purchases from Macy's look like in the data? It's an awful lot of zeros and then a few $100 and $200 and $500 and $1,000 things. So you get this enormous variance in how much purchases people make from Macy's in a given week.

So their sales and the treatment period average $1.89, although this is a fake currency to not reveal exactly how much money was involved. But they average $1.89 and they have a standard deviation of $19. So the standard deviation of sales to a customer in a week or whatever is-- actually, in a two-week period is 10 times-- standard deviation is 10 times the mean because you just have a lot of zeros and then this fat upper tail.

A percent of return on investment is a very small fraction of sales. So this campaign cost about 1% of sales to the targeted customers. So that says that they were paying-- if they were averaging-- if they were averaging $2 worth of sales per person, they were only averaging spending $0.02 on this campaign.

So if you think about the way-- like, if you're selling-- if you have a campaign, the campaign is costing you 1% of your sales, suppose your profit margins are 50% Then if your sales go up by 2%-- if your sales go up by 2%, then your profit goes up by 1%. That offsets the cost of the advertising campaign.

So if sales go up by 2%, they're break-even. If sales go up by 2.2%, that's a 10% return on the investment. If sales go up by 1.8%, that's a negative 10% return on the investment.

So the return on the investment, what you need to do is be able to estimate whether, did sales go up by-- you need a-- did sales-- you need a standard error of 0.1% on your estimate of the effect on sales to be able to tell-- if you're trying to tell 2% apart from 2.2% to 1.8%, you want a standard error of 0.1%. So you need a standard error of your estimate of 1/1000 of mean sales.

So then since the standard deviation is 10 times the mean, that's 1/10,000 of a standard deviation. If you think of variances going down like root n, that's like 1 over 100 million to get the standard error you want.

And so even though this campaign seems like this campaign had-- whether you think of it as a million observations of people or you think of it as 30 million observations of ads, orders of magnitude, you're going to need tens of millions or hundreds of millions to do something like this and get statistically significant estimates. So it's a cautionary tale that we should think of estimating returns to advertising for goods like this where there's an upper tail of sales as a very, very difficult thing to do.

OK. Second observation they make is that this is really another cautionary tale for many papers you see that are estimating advertising, that-- you know, viewing advertising is not an exogenous treatment. You can't just regress purchases on whether someone saw an ad.

Here, they have experimental variation. Let's suppose you didn't have the experimental variation. You try to regress purchases on whether people saw an ad. That's not going to be a very good estimate.

So there's this-- many firms like-- I don't know if you've ever thought it seems weird. Like, you do a Google search and you know you want to rent a car from Hertz. And you type in Hertz, and then you get-- the first thing is an ad for hertz.com, and the second thing is just hertz.com. And you know, Hertz is paying $0.50 per click for that top ad. And you click on that first click, and you're wondering, like, why does Hertz buy that top ad that says hertz.com?

You know, it does make-- it does bump Avis down to third on the page instead of second on the page, or something like that. But you know, again, if you do a naive study to Hertz and say, oh, you bought this ad campaign on Hertz, and of the people who clicked on your ad for hertz.com, you know, 30% of them went and rented a car from you. You did fabulously well on your $0.50 that you paid in the fee.

Well, obviously, those people were clicking on the Hertz ad because they'd already done a search and typed in the word hertz.com. They were going to buy from you anyway. And so the fact that someone sees your ad, in that case it's obvious, or clicks on your ad, that tells you an awful lot about the person's intent, and you can't view that ad as exogenous.

Here, they're arguing that this is true. You know, even in an environment like their experiment where-- you know, whether someone sees the ad for Macy's is different-- is determined by whether they're on the internet to begin with, whether by their browsing behavior. If you had just done a naive estimate of someone-- like, you think about it, you do an estimate. Even if you don't have this data set, you do an estimate of-- not the Hertz example.

Just when someone sees an ad for Hertz cars, are they more likely to rent a Hertz car than not? Yes, they're more likely to rent a Hertz car when they see a Hertz rectangle ad just pop up on their screen because they're on the internet. And if they're out hiking with their friends or if they're in a movie theater or they're in a bar, they're never going to rent a Hertz car because they're not on their computer or on their phone.

And so browsing isn't random. And you can't think of, I put individual fixed effects and that solves the problem. People are different people at different points in time. And there's you who's the browsing internet you and there's you who's doing something else you. Those are different people. They have different preferences and different likelihoods to purchase.

And then there's also these ad-serving engines have all these ads available. The untargeted ones, the content ones, the demographic ones, the re-targeting ones. All of those things can affect your ad views. So whether you see an ad or not depends on who you are, and it's going to be correlated with your likelihood of buying something from Macy's because people who buy from Macy's just are different than other people.

Usually, this makes us overestimate ad effectiveness. Here, what's funny is that this is an example where this was a campaign where the opposite effect was true. That effect would make you underestimate ad effectiveness.

And so that's why I use Macy's or Bloomingdale's as examples. My model for what's going on here is you have a store that sells to people who are in their 60s or 70s who are not on the internet much. And then if you learn that someone is on the internet for hours a day, you learn that they're not the typical customer of-- they're not the high-spending customers of that department store. They're people who are younger people who may have one time bought something at Macy's, but rarely shop at large, fancy department stores.

So anyway, the fact that you saw an ad-- like, seeing an ad in the-- because they have sales in the pre-period before the campaign started. And so what they find is that the people who saw ads were buying less in the pre-period than the people who didn't see ads. So here, there's a negative correlation.

So I guess the difference was, you know, $2 versus $1.81. And it's just-- the viewers of the ads are just a worse clientele than the nonviewers of the ads based on probably the demographics of who are viewers and who are nonviewers.

And this is a retailer where I think it was 85% of the sales in this campaign were in-store rather than online. So some people buy from macys.com, but probably most people-- store sales dominate their online sales.

OK. As I said, intent to treat estimates are inconclusive. So you do have this sort of very classic, high-quality IV thing where you just instrument for seeing an ad with being in the treatment group. And then you look at, how do the sales compare? That gives you a valid estimate that's not contaminated by any of these things.

And what you find, in the treatment group-- in the treatment group, sales are-- during the treatment period, sales are $1.89 in the treatment group, $1.84 in the control group. So the difference is 5.3 cents, standard error of 3.8 cents.

In terms of estimating the ROI of this campaign, they're estimating that this campaign has a return on investment of 40%. So Macy's should be very happy to run this campaign. However, it's got a return of 40% plus or minus 100%. So we get this very, very noisy estimate of the return to the campaign, suggesting that this campaign is effective, but we really don't know.

They then also do a differences-in-differences equation, taking into account that the pre-period treatment control group purchases were not exactly identical. You were randomizing 1.5 million people, but when you randomize 1.5 million people, they're not always exactly the same representative. And they didn't randomize based on pre-period purchases because they didn't have the data on pre-period purchase at the time they were doing the randomization.

So anyway, differences-in-differences account estimate that it's 8.3 cents plus or minus 5.9. So even bigger return on investment, even noisier estimate.

Difference-in-difference estimates also suggest that there may be long-run benefits after the campaign ends. The point estimates are you're making 5-point-something cents during-- 5-point-something cents during the treatment period. You're making another $0.06 in the week after and then $0.07 total for the three weeks after.

You can certainly imagine that if this is-- you're showing ads for a department store, some people see the ads for the department store. They decide-- that makes them think about going to the department store, and then they go the week after they've seen the ad online or they've seen 30 ads online, or something like that.

The paper referred to the sort of difference-in-difference estimates as their preferred estimates. I mean, my view is I think the difference-in-difference estimates are somewhat problematic. Again, if you take this view that people are different people on different days, do we really want to trust putting in individual fixed effects? Because maybe the-- maybe the reason that people spent less in the pre-period in one group than the other was because during the pre-period they were doing things other than shopping and postponing their shopping.

And then during the next period, they may be doing, like-- you know, there must be something different about them. And the fact that they were doing less shopping in one doesn't mean that they're going to be doing less shopping in another or they're spending more time online, one versus the other. I find that in some sense the paper is very skeptical that individual fixed effects are a way to control for things like this. Any questions on that?

And let me say, if you-- are there other ad effectiveness studies that give more significant estimates? Yes, there are. You know, one thing that's difficult here is this-- you get that factor of 10 that the mean is $1.89 and the standard deviation is $19.

If you're doing a zero-one treatment where you're trying to get people to sign up to buy cell phone service, you show them an ad, they can click on it and buy the cell phone service right there, at least that's one thing that-- you've eliminated at least that one source of the standard deviation being 10 times the mean. And so you can-- at least some part of it you can-- some part of it you can overcome when you have the low variance in the individual-level data rather than the high variance in the individual-level purchase data.

Yeah. Obviously, you can also estimate ROI better if the campaign is more expensive. If you have someone who's basically spending all of it-- you know, this advertising campaign is scaled to be a small fraction of the revenue of this store. If you have a campaign that's a large fraction of the revenue of some online business, that, again, makes it easier to estimate the return.

OK. So question? I'll go on-- I'm going to do another difference-in-difference paper. And I wanted to include Tobias's paper, both because it's been great to have him in the course, and because I think it's a really nice paper.

And I think actually, this is evidence on what I think of as a very, very timely and important question in online advertising. You know, in online advertising, targeting is just playing a tremendous role now. Most ads that you see on the internet are targeted somehow to you.

Targeted ads sell for much higher rates than untargeted ads. You know, it can be a-- you know, order of magnitude or two orders of magnitude difference in prices. Just if you have just a random web page and someone's using privacy blocking software, you know nothing about them, you're just trying to show them a display ad for Coca-Cola or whatever, those ads cost very, very little. Those ads following you around of the things that are in your shopping cart somewhere, those ads just cost, you know, two orders of magnitude more than the ad that's just a random banner ad displayed to anyone on the-- every single person on the internet.

You know, there's been increasing interest in privacy, I think driven by two things. One is there's a set of consumers who we know care very much about privacy, or at least say they care very much about privacy. It's always one of these puzzles that people say they care a tremendous amount about privacy, but then they do things to give away their privacy for almost nothing.

So there are many-- you probably know, on your phone there are many apps that track you and will report back to the mother ship everything about you. Like, you install-- you have a weather app on your phone that you look at once every month to look at the weather, and that weather thing is actually reporting your location back to the owner of the weather app every half hour when your phone is active, and sending that information to all kinds of advertisers who are targeting you based on knowing where you are.

And you gave up all that privacy just to install some app and then not bother to figure out in its privacy settings whether it's allowed to send your location to people or not when you're not using it. People are often also-- you'll get some click on here for enhanced displays, or something like that, and you get better service.

But anyway, people say they don't want to be tracked on the internet. And then also, there's a battle between different internet firms. So you know, Apple in particular probably prefers for Google to make less money rather than more, and Apple prefers for Facebook to make less money rather than more, and Apple would like Apple to make all the money.

Apple knows that these other firms that may-- like, Google may be able to sell Android phones really cheap because Google makes all the advertising revenue. If Apple can get Google to make less advertising revenue, Google may be less able to compete with Apple in the phone space.

So Apple's been a big proponent-- maybe for the goodness of their heart, they've been a big proponent of online privacy and of settings in the iPhone that make it harder for firms to track you using data that's on your iPhone. So there's both sort of Apple trying to be the protector of privacy and governments, especially in Europe, trying to be the protector of people's privacy and make it harder to track people.

And you know, it's a big question of, what is going to happen if it does get harder and harder to track consumers on the internet? One is you could have a real decline in ad-supported content.

It may be that websites-- you know, obviously, *The New York Times* does have a paywall. But it could be that many online newspapers that relied largely on advertising, if they can't track the consumers and their advertisements, instead of being able to show you the shoes that you put in your shopping cart, all they can do is show you an ad for Coca-Cola. It may be they don't make enough money to continue to survive, and we may have a big decline in ad-supported content relative to non-ad-supported content.

There could be twisting of what's available on the internet. Like, you don't-- if you know exactly-- like, if you know that someone is reading a, you know, vintage sneaker website well, you know what to advertise to them if they're browsing a vintage sneaker website. Whereas if they're just browsing a news page or boston.com, you have no idea what to advertise to them.

So it could be that we get this twisting of content away from places that need targeting ad based on exogenous information and the ones who can target based on the content of their page. You could also imagine this is going to increase the power of large versus small firms because people are often logged into Facebook or logged into Google all the time.

Therefore, Google can advertise to you in a targeted way, even if you're turning off your cookies for many other things because you're still logged into Google. And so it could be that blocking other people from tracking then just makes Google and Facebook all the more dominant because they're the ones that you're always logged into and they can target you when others can't.

Anyway, so most important regulation on this thing is EU's General Data Protection Regulation, GDPR. It was adopted in April 2016. It went into effect May 2018. And GDPR says that websites cannot track European consumers without explicit opt-in consent.

So I guess we don't have many Europeans here. So if you've-- I mean, for me, I know it was a shock going to Europe in 2018, and just every single website you went to-- and when you had a slow home internet connection, which I did, every website you come, you're just trying to look at a website.

And immediately pops up this sort of informed consent thing where you have to give your consent and click either yes, I consent, or click on the show me the privacy settings, and then click on-- if you show the privacy settings and then click on the minimal, automatically the default is you get the minimal-- only the strictly necessary cookies are enabled. But every European website has this pop-up thing that you have to answer the very first time you visit it.

Maybe you're used to this. Like, if you go to look at any World Cup website or the guardian.com, you always get this pop-up thing that you have to get rid of. Anyway, that's GDPR. It went into effect on one date, and not everybody complied on the first date. But you know, there were very large fines for not complying with GDPR, so many, many firms did comply pretty much right away.

And then what Aridor, Che, and Salz are trying to do is use a difference-in-difference design to explore a number of things about GDPR. The first basic question is, how are consumers going to react to these pop-up, opt-out things? How many consumers are going to opt out of data collection, and how many consumers are just going to get used to clicking yes, yes, yes, yes, yes? Because it is the single quickest way to get rid of the thing is an accept all cookies box.

Second question is then, how does GDPR change the composition of consumers who are observed? You know, the consumers are not opting out at random. It's not like you're losing x percent of consumers chosen equally likely, You're losing people who want to opt out, and so the people who opt out could be different from the firms that opt in.

And then in some sense, it's one and two together that are determining how GDPR is going to impact firms that do rely on consumer data for their business model. And they have this interesting idea that it's possible that GDPR could even help firms because you're getting rid of customers who don't want to be tracked. It may be that customers who don't want to be tracked just naturally buy less, or it could be that consumers who don't want to be tracked are going to be screwing with your data by eliminating their cookies or using pop-up blockers and using other things.

And once you get those people who are obfuscating their presence and screwing up your statistics, the statistics on the remaining people are cleaner and easier to analyze. And it could be that-- at least that could offset part of what you lose by not being able to track people anymore.

So they have this-- again, it's a simple difference-in-difference estimate, but it's a difference-in-difference estimate exploiting, again, a very detailed relationship with some firm that does-- that operates online. I had not been aware that such a firm existed, but here's my attempt to infer, like, what does this firm do?

So this is an intermediary that contracts, they say, with almost all top EU and US travel sites. So that means it's a very large behind-the-scenes firm that has contracts with pretty much everybody, which I would not have guessed existed.

And what does it do? When you enter a query on one of the sites-- like, I just went yesterday to kayak.com and typed in a search for a flight from Boston to Orange County.

And to try to look like somebody who wasn't actually going to buy the flight, I didn't change the default dates. I just clicked Boston to Orange County, click Search, because I was trying to signal to Kayak that I wasn't actually going to buy this plane ticket. And I figured that if you ask, how is someone going to not buy the plane ticket they just searched for, the fact that they left the default dates on there of January 3 to January 10 is probably a good sign that they're not buying.

Anyway, so what did Kayak do? It popped up as-- it did give me the flight prices, but it popped up as the very first ad here above all the flight prices. I don't know if you can read this. It says aa.com ad. And this says, you know, Your next adventure's a flight away, $371 American Airlines, quote, "View Deal."

And this is going to take me away from the Kayak website to the American Airlines website because the description in the paper of what this intermediary does is any time anyone visits a travel website, it has cookies on them. It analyzes their behavior. It does a machine learning algorithm to figure out whether they're going to buy from the website they're currently searching on.

If it thinks there's a low probability they're going to buy in the website they're currently searching on, it runs an auction and auctions off the ability to show a display ad to them, to all the other firms that it serves. And then those other firms can bid for this consumer who is probably not going to buy where they are now.

And so in this case-- so what I think it did was it said, Glenn is not actually going to purchase a ticket on Kayak, so let's show him an ad for something he might want to buy if he goes to the American Airlines website. American is the biggest carrier, I think, in Orange County airport. He may be trying to buy-- it may be that American Airlines is willing to pay $0.09 to have me visit their website or something like that, OK?

Anyway, so what they have is they have access to clickstream data. So they have, in some sense, access to every-- you know, Glenn Ellison typed in at 10:13 PM last night, search Boston-Santa Ana. We conducted this auction. American Airlines won the auction. They bid $0.09. He either clicked on the American website or he didn't. If he clicked on the American website, did he buy from American Airlines?

So they have the whole clickstream of everything that I was doing. They mostly aggregated to the website country week level, so it's mostly going to be comparing customers in France on the first week of GDPR on kayak.com to customers, week before GDPR, kayak.com, first customers.

Difference-in-difference is it's a treatment control design. They have five treatment countries. UK, France, Germany, Italy, Spain. And then they have three control countries. US, Canada, Russia. And depending on which website it is, some of the websites are active in different subsets of these countries, but enough of them have activities both in treatment countries and either Russia or North America that they can then do regressions that have website fixed effects and country fixed effects and time fixed effects and try to analyze the effects of GDPR.

They have some nice simple pictures here, showing you that GDPR is having an effect. So this is an example graphing the number of-- and the way GDPR works, if the GDPR pop-up shows up and then you click Don't Track Me, this company never saw you because it's like Kayak-- someone did the search on Kayak. Kayak popped up the track me or not box. They click Do Not Track Me.

And then once they click Do Not Track Me, they can't send the query to whatever this behind-the-scenes intermediary is. The behind [INAUDIBLE] never get any message whatsoever. And they specify that if the behind-the-scenes intermediary never gets a message, it's basically this box isn't here, and just the lowest price flight would have been in this box instead.

Anyway, so what they track is they recognize that GDPR is in effect by the number of cookies that they-- the number of requests that they get from Kayak drops. And what they show here this is like the raw data behind the differences-in-differences.

This is the-- this is the searches that this intermediary's being told about in the US and Canada, and this is the number of searches that they're being told about in each of the EU countries. And you can see that immediately or perhaps even starting the week before, there's this big drop-off in searches that they're being told about in all the EU countries, and then there's not a corresponding drop-off in the US and Canada. So that indicates that GDPR is having an effect.

They also note in the data that there are other firms who you suspect-- and in this case, they were able to verify-- did not actually comply with GDPR right away. And so in this case, the searches in one EU country and the one non-EU country that this website served pretty much just continued on the same path across the line and didn't have a big drop-off.

So somehow this website, they were able to then go back and verify that website was not GDPR-compliant right away. And it took later to come into compliance, whereas this one did.

But when they do, it's immediately apparent that the regulations are having an effect, and they're seeing fewer customers from those websites. And obviously, it could be that consumers go to Kayak and do get annoyed by the pop-up boxes and never actually do their search, but this is-- fewer people are doing a search and then having the intermediary being told about their search, whether it's they're never doing the search or they do the search, but they do the search and they're not being tracked.

So what are the magnitudes? First one is that while people talk a lot about privacy, it seems like most consumers just go with the-- just get that thing off my screen by just saying, yeah, fine, I accept all cookies. And I know I accept all cookies all the time on European websites.

And what they find is that the log of unique cookies goes down by 12%. Log of recorded searches goes down by 11%. So there is this big drop-off right at the same time that people-- appears to be people not giving consent.

Also, interesting here is the-- they sort of-- every consumer who they are allowed to track, every consumer they are allowed to track, they know how many times they've seen that consumer search in the past and how many searches are recorded on the cookie that they're putting on that machine. And so what they show is a change in the-- so there's this distribution of, how many times have you seen a consumer?

And it's a histogram that's like a lot of the consumers you've seen exactly once, and then some consumers you've seen twice, and then many consumers you've seen three, four, or five times, whatever. So you have this histogram of how many times you have seen this consumer before.

What they note is that immediately after GDPR goes into effect, the number of consumers that you've seen exactly one time, this bar, goes way down. And the number of consumers you've seen two, three, four, five, six, seven, eight times seem to just increase proportionally and be very, very slightly.

So what they note is that it seems that when these consumers now opt out, you're seeing many fewer consumers who you're seeing exactly once, which would be consumers who may have something already installed to delete their cookies or to block the cookies from identifying them and make them look like a new person every time they show up, even though they're actually someone who you're seeing over and over and over again. And so what they note is that it seems like this drop in the number of people seen exactly once makes you think that of those 13% of the people who dropped out or whatever it is, many of them also were cookie deleters or had some blocking technology they were already using.

So advertising revenue. There's an immediate drop in the number of clicks that are recorded on the compliant websites. Number of clicks drop-- I guess it's 13%. Revenue goes down 16%, like, on the day the thing takes into effect. So this is certainly-- this is bad news for this advertising intermediary who we're getting the data from. It's like they just-- 13% fewer-- 13% fewer clicks referred to them to try to find an extra potential to assess.

You know, you presume they get paid based on a per click basis for how much revenue they can produce for Kayak. 13% of the searches disappear for them. 16% of the revenue disappears.

But then this interesting effect they find is that, you know, the bids drop off or appear to be slightly low on first week. And then over the next seven weeks, there's just this steady climb in what people are bidding for the ads. So again, the bidding for the ads is like bidding to what American Airlines paid to show that ad to me because I'd done the Orange County search.

And what they show is that there's nothing going on in the first two weeks, but then it's as if bids go up and up and up. So why would that happen? Two potential explanations.

The people who are left behind are just more valuable because they're different people. And those more valuable people, they didn't know that the people left behind would be more valuable people, but they're learning that those people left behind are more valuable people, and so the bids go up.

Or two, the intermediary is better able to figure out what those people want to buy. So the intermediary may also be providing American Airlines and hotels.com predictions of the probability that I will buy an American flight or that I will rent a hotel through hotels.com. The fact that they can tell-- provide better information about me may get their bids to go up in these auctions.

And so if the intermediary is better able to decide who is going to purchase from Kayak versus who would purchase from another firm, if they can do a better job of data analysis, that also makes prices go up. So either people are better or you're better figuring out which of-- who are the valuable people? And that causes the average prices to go up, OK?

And so this effect, it's not-- it's not like it's good news for all the firms involved. They can't track consumers. But at least what's happened here in the first seven weeks, that's offsetting about half of the decline that you would have gotten. So if you've gotten an initial 16% decline, it may be actually, you're only going to get an 8% decline in revenue because the bids are going to go up over time once people realize who the consumers are who are left over and how well you can track them.

OK. And then the third thing is they did have access-- this firm is doing a machine learning algorithm to predict whether you're going to buy from kayak.com when you search on kayak.com. And so they have access both to the predictions that the intermediary made at the time that the click-- at the time that the search occurred and whether the person actually then bought from kayak.com at the end of their session.

And what they find is that true purchase frequencies increase. And obviously, they can't-- they can't do-- this is not a difference-in-difference estimate because they don't know whose purchases would have been observed-- they're observing this selected sample afterwards. But the selected sample afterwards is purchasing more often than the unselected sample was in the pre-period. So these people are more frequent purchasers.

And you know, when they look at, how well is the intermediary predicting purchases, the mean squared error goes up after the-- after the regulation goes into effect, so they're predicting worse. But you know, mean squared error is sort of a problematic measure of prediction because the probability of purchase is a very small number. Like, it may be 10% or whatever.

And so if the mean goes up, like if you're trying to predict a zero-one random variable, zero-one random variables are hardest to predict when their mean is 0.5. If their mean is 0.001, they're very easy to predict.

And so in mean squared error sense, when the average purchase frequency goes up, their mean squared error goes up. What they show is that there's a graph-- I don't think I have it here-- that the mean squared error looks like this. The mean squared error goes up, and then the mean squared error comes back down. So by the end of the sample, even though the mean is higher, the mean squared error is not higher.

And so it seems that when the firm is updating its machine learning algorithms using the new, better quality data, it actually looks like, if anything, it's predicting better what consumers are doing. So it initially suffers because it's predicting in the wrong environment or for a selected sample of people using its unselected data.

But then once it does it, it looks like-- the intermediary looks like it's improving its predictive ability in this new cleaner-- in some sense, the GDPR has cleaned the data for them a little bit and made their job easier. OK? At least it's not so significantly easier, but at least it's clear that it didn't make their job harder on the remaining people. OK. OK. I think that's what I had on Tobias's paper. No.

OK. So I wanted to go to one other-- one other causal paper. So this is an example of-- as I said, there's been a big increase just in the last five years or whatever of people interested in just using same kinds of quasi-experimental methods you see in many other applied micro fields to do advertising studies and think about, how do we estimate the effect of advertising using quasi-experimental designs?

There are lots of neat effects here. There are lots of things that do just sort of, again, straight difference-in-differences using some kind of regulation goes into effect banning advertising or allowing advertising, or someone runs an advertising campaign that lasts for four months. And then it stops on one day, and you use it before and after.

Spillovers from political advertising are a really neat general purpose idea. I guess-- maybe you guys don't watch TV, but if you're watching TV in Massachusetts right before an election, you may have noticed tons and tons of campaigns for Maggie Hassan versus Don Bolduc in New Hampshire, you know, buying up all the ads on Massachusetts TV.

If you live in New Hampshire, your entire TV for a month before the election is nothing but advertising-- political campaign ads. The political campaign ads shift other ads off the TV. And you can do these designs taking into effect that New Hampshire versus Vermont, no one advertises on Vermont TV because the democrats always win.

New Hampshire has a ton of political ads because it's a toss-up state. You can do new Hampshire versus Vermont before and after an election and see all these other advertisers were getting blocked out of New Hampshire, but still advertising in Vermont and Maine, and use that as an instrument for advertising.

There are papers using sporting events. If your team gets to the Super Bowl, you therefore see all-- everyone watches the Super Bowl. They see many more Budweiser ads than they would see otherwise. How does that affect purchasing?

In some ways, I talked about the Bronnenberg, Dubé, Gentzkow paper about people moving from one city to another. If you move from one city to another, you're seeing different ads when you've moved there that could be an exogenous difference in the ads that you've seen. We can use ideas like that.

Brad Shapiro's paper is using what we call a border design in media markets. So the application here is to antidepressant drugs. So there was a class of antidepressant drugs called SSRIs. I forget what that stands for. Introduced in the late '80s and '90s. These drugs were under patent protection and earned billions and billions of dollars a year in revenues in the period he's studying, which is late '90s to early 2000s.

And you know, these are-- this is just a picture of-- that's what a Zoloft ad on TV would look like. You'd see this ad. This is the guy who's depressed and he's under a cloud, and symptoms persist every day for at least two weeks. And then later in the commercial, somehow he takes Zoloft. And the clouds part, and the sun comes out, and he's happy.

And then you have this woman talking about how, I couldn't do anything. My life now-- my life is so much better. And then obviously, then that 10-second comment from her, followed by this long litany of, may cause suicidal thoughts, may cause gastrointestinal symptoms, may cause loss of appetite, may cause cancer of various kinds.

Anyway, it was sort of-- [SIGH]. It was-- I don't know. Again, prior to 1997, TV commercials were not-- were not-- there were not TV commercials for prescription drugs in the United States. Prior to 1997, there was no TV ad-- the United States was like every other country in the world. There was no advertising for drugs.

Rules were changed for how many of those warnings you had to put in an ad around 1997. And so 1999, it became feasible to have only 20 seconds of your ad with all the side effects and to still have an ad that survived with all the side effect stuff. So anyway, first depression ads appeared in 1999.

What does Brad have in this paper? He's got prescription data. So the prescription data is a 5% random sample of physicians who prescribe antidepressants. The data set also includes physician characteristics, including the exact office address. Using the exact office address, he can then map each physician to a county.

Now obviously, not all patients visit a physician in the county in which they live, but these are relatively rural counties, so we hope that's not such a bad assumption. And then he takes the physician-level data, aggregates that to the county month level. And most the analysis is going to be done at the county month level.

And I guess, continuing the theme on many of the papers we look at, it is using an impressive data set. And you need detailed data sets to write good papers these days. So what does he have on advertising? He has TV advertising expenditures by brand in each of 101 markets in the US from September 1999 to December 2003. So the basic advertising data set is at the DMA, which is like-- you can think of as roughly as a city, the DMA times month level.

Now, much of the advertising of pharmaceuticals is national, but there is substantial local advertising also. So what he does is he creates a measure of local advertising in each market by basically apportioning the national spending according to each DMA's share of the national population, and then adding on top of that the advertising that's purely local to that DMA. Shapiro's strategy for identification is to limit the analysis to DMA borders and to treat monthly advertising as exogenous once you include the product-border-quarter and product-border-DMA fixed effects.

So why does that make sense? So imagine that you didn't have just the borders here, and you were going to try to compare advertising and sales in the Columbus area with advertising and sales in the Cleveland area. You know, there, you'd really worry that inherent demand is different in Columbus than it is from Cleveland and that the firms who are advertising know that demand is different from Columbus and Cleveland. And therefore, the advertising in Columbus is correlated with the demand in Columbus.

The idea of the border effect is to say that it's kind of like the same thing you would do in a regression discontinuity design, but in a spatial sense. So you know, if you look at Columbus versus Cleveland DMA, you see one, two, three, four, five counties along the Columbus side of the border. And then you see one, two, three, four counties along the Cleveland side of the border.

You know, these counties are going to be relatively rural counties. They're going to be fairly far from the center-- the cities that are the center of these DMAs. So the idea is supposed to be that the people on the Columbus side, the people at the Cleveland side are very similar.

The advertising that the Columbus people see is not going to be driven by the characteristics of the people on the Columbus side of the border because the people on the Columbus side of the border buy perhaps, you know, one or 5% of the prescriptions of antidepressants in the Columbus DMA prescriptions are going to be-- you know, the advertising will be driven by something-- a factor affecting the Columbus DMA or the Cleveland DMA. And so you can treat the advertising choices as exogenous once you put in these product-border-quarter and product-border by DMA fixed effects.

And again, the product-border-DMA fixed effects, those are going to capture-- if the people in the Columbus side of the border are slightly more depressed than people on the Cleveland side of the border, that would get captured by that nontime-varying fixed effect for the Columbus side of the border, OK?

So the basic exercise here is to just put in all those fixed effects and regress the log of sales in each-- on each side of the border in a month on the advertising on each side of the border in a month, plus these fixed effects. And what do you find?

So first thing that he finds is that advertising has large effects. So look at the first number that I put in the box there. The DTC has a-- direct to consumer advertising in this market has a coefficient of 0.024. I believe the right-hand side variable's in the scale of pennies per person. So this is if you conduct an advertising campaign that costs $0.01 per person in that area, in the DMA you would get direct to consumer-- you would get-- sales go up by 2.4%.

OK. Second significant effect you see here is the spillover of rival advertising. So spillover rival advertising is a coefficient of 0.016, which is roughly 2/3 of the size of the own advertising effect. So that is, you know, it's better to advertise yourself, but having a rival advertise gives you almost 2/3 of the sales boost. And obviously, you don't have to pay for the rival advertising.

Second result here, if you look at the quadratic terms, this one and this one, quadratic terms have negative coefficients on them, which means that there is decreasing returns to advertising. If you look at the shape of those quadratic things, functions, it would be that advertising maxes out at about five and a half on the own advertising, nine on the rival advertising. Those levels are well above the levels of actual advertising the data, so it's just capturing the curvature that there are decreasing returns as you advertise more.

And then the third important effect in this regression is the DTC times DTT rival coefficient. What that says is that the marginal return to my advertising goes down when my rival is advertising more. You know, that fits with the general. If there's decreasing returns to aggregate advertising, that would be true. As my rival advertises more, we're already getting further out on the decreasing return side and that's going to cause a reduced benefit from advertising.

And it's fairly simple estimation strategy, but I think-- you know, fairly simple estimation strategy, you just have to buy these regression discontinuity-type assumptions that the people along the two sides of the border are similar other than a fixed effect that's not time varying and that the advertising is not being driven by something that's border-specific. It would be something that if it's anything-- if it's not just purely random, it's driven by something related to the cities, and therefore wouldn't have different effects on the different sides of the border.

The next thing Shapiro does, though, this is-- you know, like many papers, it both combines a standard, modern causal inference approach with a structural estimation. So his structural estimation is basically driven off the same levels of-- the same exogenous variation that's driving the reduced form estimates.

But here, he does this nested logit model of demand. The nested logit model has three levels. At the top level, there's a choice between an inside good and the outside good. That choice is driven in part by total advertising in the market.

The second level is a choice between categories. I think I mentioned there are six different categories of antidepressant drugs being sold at the time. One of them is "other." And so that choice is driven by his right-hand side variables, which is advertising within that particular category of antidepressant.

And then the bottom level, which is going to be the most purely business dealing level, is competition between-- people make a choice of which drug within a class of antidepressants they buy. Again, that's going to have advertising on the right side. That's advertising of the particular molecule. At least early in his data, there are actually only two of the firms doing advertising. There's GlaxoSmithKline advertising Paxil and there's Eli Lilly advertising Prozac, OK?

What do the advertising variables look like? You know, the advertising variable that's on the right side of this-- so advertising for Product J, Order Level B, Time T is going to be a sum of delta to the s times advertising by that firm in the same region at time t minus s. So he estimates these discount factors, delta to the s, to estimate, how quickly does advertising die out?

Again, there are three separate deltas being estimated. There's a delta being estimated for the aggregate demand for the inside good versus the outside good. There's a delta being estimated for the within category effect. There's a delta being estimated for the product effect.

Biggest results he has there on those deltas is that the deltas are fairly large, but the delta is-- or sorry, fairly small. Advertising dries out fairly quickly, and it dries out at different-- it dies out at different speeds at the different levels. So it dies out most quickly at the within category level.

There, he finds that within three months, 90% of the benefit of advertising has died out in terms of getting share to you versus other firms selling similar drugs. And then it dries out more slowly but still fairly quickly at the aggregate level. There, it takes six months for 90% of the demand to die out.

So in the graph I put on here at the right, I graphed these effects. So this is looking at what the marginal effect would be if Zoloft started an advertising campaign of some size at the beginning of 2002. The red line at the bottom of that graph shows how many extra prescriptions per month Zoloft gets in thousands. And so what they're getting is 20,000 extra prescriptions in the first month.

But then as I said, much of that dies out. And so you can see by three months later, many of those prescriptions are lost, and then the effect continues to die out.

The black line-- or is it dark blue?-- in contrast, is this is the gains in total prescriptions. So while Zoloft prescriptions go up by 20,000, total prescriptions go up by 70,000. So that's saying 50,000 extra prescriptions are being gained by other firms, not the manufacturer of Zoloft. Those also die out. And here, we see those dying out more slowly.

So at the beginning, it's 20,000 and 50,000. The other firms are getting two and a half times what Zoloft is getting. The dashed line here is showing that by the end of the sample, the other firms are getting five and a half times what Zoloft manufacturers are getting. So this is showing that there is really substantial spillovers to the rival firms.

Next topic he does in the structural model is asking how much advertising the firms do relative to what the model says would be the optimal level of advertising. You know, many, many structural papers impose the supply side first order conditions, imposing that firms are behaving optimally.

Shapiro does not want to do that. His argument for why he does not want to do that is advertising is brand new. It's being done for the first time in the United States. Doesn't really happen elsewhere in the world. Why would the firms know what the marginal return to advertising is until they've been doing advertising for some period of time?

What he finds in his simulation is that it looks like the advertising-- given the advertising is effective, firms should be doing much more advertising than they do. In particular, in the early part of the sample when it is just Paxil and Prozac being advertised, the firm should be advertising four times more than they do if they did the equilibrium level of advertising, even given the free rider problem.

It's different across the firms. You know, Paxil, which doesn't have a big market share, or doesn't have a very big market share, is doing about as much advertising as it should. And it's Eli Lilly who's doing much less advertising than it should, although Eli Lilly does come in notably and do a lot of advertising right in the very final month of that sample. So what he finds is that firms should be doing four times as much advertising as they could if they were playing the equilibrium of this game.

Second counterfactual said, well, what if they had an advertising cooperative? As he said, advertising can be much less than is optimal for the firms, given that there are so many spillovers and that advertising by one firm decreases the incentives of the other firms to advertise.

And what he finds is that an advertising cooperative would then do yet two and a half times more advertising than would happen in equilibrium. So that's 10 times the current level of advertising is what an advertising cooperative would do.

What would be the profit impact of that? Raising advertising by a factor of 10-- you know, the advertising does cost money-- ends up raising profits by about 15% for the firms.

So I guess that's roughly a summary of the paper is that he has these two parts. He's got the nice, simple causal analysis part using this DMA border variation. DMA border variation tells us that advertising is effective and advertising crowds out-rival advertising.

He then builds that structural model. The structural model lets us put more detail on what these curves look like. Obviously, more detail on what they look like, given assumptions on what they look like. But the advantage of putting that second part in is that the second part, we can't talk about optimality.

And the optimality, what he finds is that in equilibrium, once the firms learn how effective advertising is, you would expect them to advertise much more than they do. That seems like that could explain what's happened in American TV where we do see a lot of advertising.

And then he talks about the magnitude of the free rider effects. And again, the free rider effects are very large here, and firms are greatly advertising relative to the firms' collective interest. Whether you think it's social interest or not obviously depends on whether you think depression is being undertreated or overtreated with medications. And what are the benefits of that to people?

OK. So guess that's about all I had for today. So I'm going to then just finish up saying next class, we're going to be back in the theory empirical rotation. I'll go back to talking about theory work.

And I'll do three theoretical papers that are on search engines. The Edelman, Ostrovsky, Schwartz, my paper with Susan Athey, and Anderson and Renault's paper. And then I will do one paper by Armstrong and Zhou, which is about retail platforms and thinking about design of a retail platform, how much information to give consumers.