FEDERAL TRADE COMMISSION

FIRST ANNUAL

FEDERAL TRADE COMMISSION & NORTHWESTERN UNIVERSITY

MICROECONOMICS CONFERENCE

Friday, November 7, 2008

9:00 a.m.

Federal Trade Commission

FTC Conference Center

601 New Jersey Avenue, N.W.

Washington, D.C.

FEDERAL TRADE COMMISSION

I N D E X

P R O C E E D I N G S

KEYNOTE ADDRESS BY SUSAN ATHEY

MR. ADAMS: Okay. I think we're going to get going here this morning, if -- if people like my -- my bosses sit down.

My name is Chris Adams. I'm a staff economist here at the FTC. And I wanted just before we get going just to thank a couple of people. The success of this conference is really due to the fact that five people answered in the affirmative when I e-mailed them. And those five people were Carl Shapiro, John List, Pat Barry, Scott Stern and Susan Athey. And I'm really grateful for you guys for both answering that e-mail in the affirmative and in helping me and helping put this conference together.

And then a group of colleagues of mine helped them. Matt Weinberg, Lauren Smith, Dan O'Brien and Rob Letzler really worked pretty hard to help sort through the papers. We had something like 90 submissions and 15 papers presented. So, that was a lot of work to go though those papers and, I think, get a pretty good agenda.

And finally I want to thank Marissa Crawford, who is out front there and really did all the work in putting the logistics of the conference together.

So, I'm going to introduce Susan Athey, who is one of the leaders in the field and is moving on to become one of the leaders in the field of online auction advertising.

MS. ATHEY: Thanks so much for having me here and giving me the opportunity to help organize this terrific conference. And I had a great time with the privacy panel yesterday, and I'm looking forward to the other sessions today as well.

So, today I want to talk to you about online advertising auctions. And I'm going to spend maybe half the time or a little more talking about sort of just general issues in the industry. I want to highlight some regulatory issues. And then for the last half of the talk, I'm going to give a little sort of sneak preview/synopsis of some work I've been doing with Glenn Ellison. And there's a paper on my website called Position Auctions with Consumer Search.

And I've actually -- I've been working on this problem really kind of full-time for at least a year now, and just in the interest of full disclosure, I've been collaborating a lot with Microsoft on this. Right now, I'm a visiting researcher at Microsoft research, which just opened up a new -- they have an academic style research organization like Bell Labs, and they've opened up a new branch on Memorial Drive next door to MIT. So, that's where I've been sitting for the last six months or so. And I've also been working with Microsoft to design their online advertising auctions. And then in the midst of that, I got thrown into some of the interesting regulatory issues which fortunately competition in search engines has lived to see another day as of this week. So, we're very excited about that.

So, that's just my full disclosure there. So, I've been spending -- as a result, I've been spending a lot of time talking to the regulatory community about this topic in the last couple of months. And I think, you know, it is a really important topic. And because of sort of the structure of the industry and all the various issues, this isn't going to be the last time that big teams of people at either the FTC or the DOJ invest a lot of time in these issues and other parts of government. And so, I think it's -- it is really important that we all sort of invest in this and learn about it so we can make rational policy.

So, online advertising is a really big business. Just, you know, Google as a company, that's one of their main sources of revenue, and they make more than $10 billion a year from auctioning sponsored link advertisements and search. And people say, well, does anybody even click on these ads? Well, I mean, if you look at, you know, Google's market value, you kind of have to believe that they do.

You know, Yahoo and Microsoft have similar businesses, and then content sites auction space via AdSense and related programs. So, the top three players are Google, Yahoo and Microsoft, and Google is the biggest by a substantial margin.

And another sort of interesting fact is that search earns, you know, depending on which display space, you're talking about four to 100 times more per impression than kind of the banner ads that you -- that you see. And that has a lot to do with the nature of what's going on with search. Just like, you know, you don't think about Yellow Pages, you don't spend a lot of time on Yellow Pages, but Yellow Pages are a big -- have a big market share of advertising dollars because people go to the Yellow Pages when they're ready to buy. And that's sort of a must buy for any kind of direct marketers.

Just some of the competition policy issues. So, in the last -- you know, in the last two years, this has become a topic that's absorbed a lot of time. So, the Google/DoubleClick acquisition, which was allowed by the FTC, and the Google/Yahoo agreement, which was blocked by the DOJ, and I think, you know, like I said, we're going to be back. Google's dominant position and then the relationship between search and other markets suggests there's many regulatory issues to come.

It's not just that we have a big important business that has a small number of players, but there is important relationships between those -- there are things that happen in that market and other markets like -- you can think about, you know, the information that is then input to ad platforms, which came up in Google/Double-Click. You know, Google has a check-out program, which, you know, gets -- which operates in the search market, which gets information which can then be used in other ways. And, of course, there's all the privacy issues as well.

So, one reason that we sort of expect that, you know, we will continue to have regulatory questions is just that we generally expect that there's going to be a small number of firms in these markets. So, you know, we have -- generally multi-sided platform markets, so if you have advertising networks, you've got indirect network effects. The more consumers you have or the more publishers you have, the more advertisers you get. And, you know, you can't get a publisher to sign on to an ad network unless you can promise them a certain -- a certain number of advertising dollars per page. And you can only get the advertising dollars per page if you have a lot of advertisers in your network.

So, we're expecting that there's going to be, you know, a relatively small number of players, although interestingly the display market is still fairly fragmented.

I -- the other thing that's really important in search is just the huge, huge, huge investments and the huge amount of time it takes to kind of build an algorithmic search engine or a search advertising platform. So, just when you think about algorithmic search, you have server farms, a statistic I haven't verified, but what I've heard is that, you know, Google, Yahoo and Microsoft are using 3 percent of U.S. energy consumption on their server farms.

You know, you're thinking about all over the world, you know, trying to place these football fields worth of computers near cheap energy. You have -- you have algorithms for parsing language and processing text. All the algorithms for page ranking, which basically means that you're running a big, applied R&D organization. And we know that it's not easy to run an R&D organization to attract star researchers, to get them functioning and doing productive work on a large scale.

That's something that takes, you know -- something Google has been very good at and that just it takes a lot of investment and long-term -- long-term planning.

You know, there's been -- you know, as you move between the algorithmic search and the advertising platform, there's algorithms for quick prediction, there's a whole experimentation platform, which needs, you know, to be built. It needs to have metrics. It needs to have scientists designing how you do your experiments, how do you evaluate experiments. When you do an experiment, how do you know that it works? You know, we've got all these measurers of what happened to consumers. You know, which metric is most predictive of short and long-term consumer engagement? Which one is most reliable statistically?

You know, so just think about any kind of research project that you've been a part of and then think about sort of starting it from scratch, you know, building up all of the intelligence and all of the approaches, the empirical approaches and so on.

The huge database architecture and storage issues. This is something I didn't really appreciate. The Department of Justice actually helped me appreciate that more when I -- when I sort of saw Microsoft trying to comply with civil investigative demands, and I really had to get inside of the databases of Microsoft. And you just -- just the project that they had to design to come up with a system that's going to be able to take tens of thousands of advertisers, each of them placing orders on thousands and thousands of keywords, the orders themselves are complex, there's broad match, there's exact match, there's targeting, and then you have to have a system that will allow you to query that database in real time and basically run, you know, thousands of auctions a minute, maybe, and then provide all the data back to the advertisers whenever they choose to log into the system. This is a system with terabytes and terabytes of data that has to serve many purposes.

And so, then there's -- and then finally you have to have an auction mechanism which has to be designed conceptually. It has to be tested. It has to work really fast and potentially be flexible to hold real time auctions. This is just a huge -- I mean, it's just amazing, really, that these things got built and deployed so quickly, but it's also very -- a very complicated problem. And there's tons of things that you say, oh, well, why can't we do this? And, you know, it's like, well, you know, we haven't been able to build it yet because there's so many things to be built. And, also, it's just highly innovative. You know, new innovation happening all the time in sort of econometrics and statistics and in just how the auctions work and are designed. And so, it's just changing constantly.

So, that's a -- so, it's a very -- so, it's just an important industry. We're going to be involved with it from a regulatory perspective, and it is important to get it right and to think about how what you do affects the future of innovation.

Let me talk a little bit now about targeted advertising. It came up somewhat on the privacy panel yesterday. Targeted advertising has wide-reaching implications as well. So, if you think about the fact that right now TV programs are designed to deliver demographics of consumers, which are easy to sell to advertisers, the whole industry structure of content provision in television and in video is sort of set up around a certain way that you sell that content.

And if we go to -- if we imagine sort of a world where in contrast, like, say on Youtube, if Google knows something about what you've been viewing in your searching and can show you Youtube videos with ads targeted to your search behavior, suddenly there's a whole bunch of content out there that can be monetized in ways that was never monetized before.

And so, you know, that changes the incentives for content provision and it changes the industry structure. And, again, I think there’s so many ways this industry could play out. I don’t pretend to have all the answers, but I think it’s going to be incredibly interesting and exciting to see how innovation happens, how -- who gets the rents from all the value that’s created from targeting. You know, it’s a huge amount of wasted advertising right now. Think about all the purple pills you see that could never be relevant to you.

You know, imagine if every ad you saw was something that was interesting to you. There’s huge amounts of value creation. And the question is, you know, is that value going to get created? Are we going -- is it going to be created sooner rather than later? Are there going to be -- are the confirms going to safeguard the data? Is the competition structure going to be such that those rents flow to consumers and firms, or are they going to be extracted mainly by a small set of advertising platforms or content providers? Where are the rents going to flow?

And, you know, related to that, the privacy issues are important in that neutrality is also going to be important. And one thing I’ve seen first-hand is that the regulatory uncertainty inhibits innovation. Do you think about, okay, well, if I’m going to -- if I want to do a certain merger, you know, I don't know what the regulators are going to say about it, and if you lose six months or a year in this business, you know, you can really end up behind.

If I'm going to think about certain kinds of alliances or investing in certain technologies, if regulation goes one way, that whole business model may not work.

And so, I think the investments that, you know, economists at the various regulatory agencies make in learning and understanding the industries, putting out white papers and just eliminating some of the uncertainty is really -- is really valuable for helping the industry move forward.

Let me throw out some interesting questions that I think are open in display advertising that could be interesting for research. And I'm going to spend the remainder of my talk talking about search advertising. I just want to -- not that -- there's not that much research. The guys at Yahoo research have been active in display advertising, but there hasn't really been a lot of research in the rest of the community on display advertising markets. And I think there's some really interesting questions there.

So, just as some background, you know, what is the current status of things, in a lot of -- a lot of content producers like MSN, a lot of those banner ads are hand sold. So, the salesperson who has advertising accounts and they just call up and negotiate prices, and there's various degrees of targeting that can be sold. So, you can be sort of sold a bundle -- you know, here are soccer moms, you know, how much do you want to pay for a certain number of impressions for these soccer moms and so on.

But it's really because -- when it's hand sold, there's limits to how refined that can be. And part of the reason it's done that way still is that -- is that, you know, you -- that's where you make the most money. There's a lot of automated networks for pricing display, but at the moment, you know, they don't tend to get full value, at least not for all -- for all publishers.

So, what's called remnant, those are things that sort of aren't sold directly, sells for much less. Even, like, you know, a New York Times page can end up selling for much less if it's an automated type of ad network. So, ad networks create spot markets and ad impressions. There's over 100 ad networks and there's many different business models for those ad networks. And so, there's some -- so, this is sort of an -- there are indirect network effects. You sort of think that eventually this might consolidate to a certain extent, but we don't -- it hasn't yet. And so, we don't -- we don't really know exactly how it's going to play out.

So, there's questions about what's the best market design and how the markets compete. You know, is it possible to have, say, some -- a lot of MEESH (phonetic) networks that serve certain industries. You get all the advertisers in that industry and that has enough scale to sort of -- to succeed as sort of a MEESH player. Are we eventually going to see consolidation?

Why is monetization still so low? Why haven't these ad networks been able to sort of close more of the gap between hand sold and what they get? And then another -- again, coming back to the regulatory theme, a crucial input for making, you know, an ad network, certainly like in five or 10 years out, work very effectively is the information for targeting. And so, there's just a lot of questions about how the -- how the -- how that information is going to be shared. So, how can you have kind of a -- is it possible to have a decentralized platform where people are sort of coming and going, but yet very -- very fine grained information is needed to figure out what the best match is between the advertiser and the publisher and to create the value.

So, there's lots of -- there's lots of things people are thinking about and trying to do here, but we haven't seen yet the answers. And then how will regulation, competition policy and technological innovation impact the ability of competing firms to access information? Are we going to see a lot of exclusive contracts, and will the regulators permit those exclusive contracts? Can dominant firms leverage their positions without regulatory oversight? And, again, depending -- you know, if competition ends up being healthy and there's lots of different sources of information, then this won't be a problem. But there's various ways the industry structure could play out where the information gets more concentrated.

Of course, we still don't know as an empirical matter what kinds of information are most valuable. Are the -- you know, can you come up with information from sources that's a substitute for the information you get from search engines? Or does that information end up being sort of hard to replicate?

So, now let me turn to the thing that I've spent personally more time on myself, and that's sort of searching contextual advertising. So, there the objection that's being auctioned is a position in a list or for a short text ad, and higher positions get more clicks.

So, one thing that, you know, might be a little counterintuitive at first is if you think about, say, Google offering eight positions and then realizing that typically they only have, you know, one, two, or three ads, you know, how is it that they're making any money at all because it seems like the supply of spaces is sort of less than the demand for the spaces. But there are sort of two reasons why they can end up making a whole lot of money even though there's empty spots on those screens.

The first reason is that there's more clicks at the top of the screen. And so, even number two competes to be number one to get more clicks. The second reason is that these -- these things are sold at auction, they're sold at second price auctions, but there's a very active role for reserved prices.

And so, you generally have to meet a minimum reserve, and a fairly large fraction of advertisements out there are actually paying a reserved price rather than an auction price. And so, you know, it can be sort of intuitively, do you think about, say, the third ad doesn't get a lot of clicks, then, you know, you can set a higher reserve price and the second ad pays a higher price, you lose the revenue from the third ad. But if the third ad isn't getting that many clicks anyways, then you'll bank more revenue by raising the reserve price.

So -- so, you know, there's -- so, it's possible -- so, as it turns out that, you know, you can -- you can make a fair bit of money with trying to control access in the sense where people bidding for access to the highest number of clicks.

So, then another thing about -- and as I said before, you know, people are looking for what you're selling on search. It's similar to Yellow Pages. And that's part of the reason that this is just such valuable advertising.

I also want to mention contextual ads because, you know, contextual ads are -- are also fairly important in terms of revenue. And I think they play a really special role in terms of providing incentives for content provision on the Internet. So, if you think about, you know, especially small -- small published sites, even, you know, your blog, your fishing afficionado blog, how can you profit from that?

And, of course, you know, lots of people like to put up free information on the Internet, but it takes a little bit of time to make a nice site that's easier for people to navigate, to take the time to continually update it. And there are a lot -- there is a lot of really great content out there on the Internet. And the main way that people can make money from smaller sites is through contextual advertising. And there's a couple of reasons that works well. One is that it's sold on a per-click basis. And so, the advertiser doesn't have to evaluate the quality of your site or the quality of your audience. And that's really important if you are a small site.

A second thing is that if it's -- especially for content-related site, like say a fishing afficionado website, in fact, direct response ads from people who sell fishing equipment will be the right thing to put up anyway. It's better to put that up than it is to put up a generic Coke ad. And so, it sort of can be more efficient to have contextual advertising for that kind of content.

So, that -- that revenue from contextual advertising, which is sold basically by Google, will read the content of your site, look for key words and show advertisements for people who have bid on those key words, those sites can be a real -- the advertising can be a really important incentive for the creation of content.

Now, on the Internet, content is being changed every day, and so, it is -- you know, it is important to have continued incentives to investment. It's not a one-time fixed cost in creating content, but rather it's dynamic.

So, the way that these incentives are provided, I mean, it's kind of interesting. You know, if -- there are sort of two types of -- two types of relationships at a broad level. You know, there's -- you're -- you can sign up your blog for AdSense and just show ads and you don't have any negotiation. For that, historically Google would just send you a check in the mail every month. But they wouldn't really tell you how it computed that check, or even sort of what revenue share you were getting. They just sent you a check, which is nice because you'd rather get a check than no check. But it also -- that lack of transparency is a little complicated for thinking about, you know, if your check falls, like why did it fall, is it just that people didn't like your site any more, or did they cut your revenue share?

Then for larger sites like the New York Times, you know, you'll have a search bar where you can search the web. And this, in the end, the aggregate of all these things drives a fair bit of search traffic. And so, for those types of negotiations, it's really -- it's money. You know, Google is going to pay you money. Yahoo will pay you money. Microsoft will pay you money. It's really a substitutable good. And so, you're going to end up getting sort of a second price auction. So, you know, say Yahoo and Google will compete against each other. At some point Yahoo drops out and Google pays the price that -- where Yahoo dropped out. And so, again, this competition sort of determines the payments.

So, that's -- so, that's an area where, again, the industry structure has an effect on the incentives for content provision.

Finally -- so, okay. So, let me now talk a little bit about the auction itself in search advertising. So, it's a really interesting market design thing. And the auctions have evolved over time. Just in the course of 10 years, we've seen a migration from auction systems that didn't work very well to some that work very effectively. So, there's a real time pay per click -- click and/or quality weighted, generalized second price auction. That's easy, right?

So, let me tell you a little about the different parts and why they're there. First of all, it's a real time pay per click auction. So, advertisers maintain lists of pay per click bids attached to key words. When a search engine -- search query is entered, the applicable per click bids are applied, and then bids are assigned an advertisement search query specific quality score.

So, you know, the way this was first rolled out is these were just click through rates. So, these were the -- the probability that an ad gets clicked, and over time the different search engines have evolved subjective scores that are assigned -- that are part of this quality score as well. And so -- and so, the bids are ranked according to the product of their per click bid and the quality score, and what they pay is the general -- you know, the rules aren't actually completed disclosed and aren't completely committed to, but at a sort of first approximation, what we think that Google is doing is that they are -- they have the bidder pay the minimum price that would keep them in the same position.

And so, your price that you pay per click depends on your score and the score of the person below you. And so, a change in your score would be just a proportional change in the amount you pay per click.

So, why this format? Well, a real time auction could be a rate card, it could be negotiated sales, it could be periodic auctions. But I think that this was partly -- I think that you could actually use periodic auctions in this market for auto insurance. You know about how many search for auto insurance. You know who the bidders are. You could hold an auction for the next six months impression of auto insurance. People would come and you would make some money.

But overall, it’s that you’ve got your millions and millions of products. They’re highly variable prices. The demands can change over time. You’ve got a lot of small advertisers who want to kind of experiment and learn about how their campaigns perform. And so, this real time auction tends to work pretty well.

You’ve got a lot of direct marketers who are interested basically in -- you know, they’re -- it’s not terrible for them to buy impression by impression as opposed to, like, a Coke campaign where you want to plan a whole campaign at once. But it does have costs. Firms have to monitor and fine tune, and there’s limited abilities -- there’s some limits to the abilities of firms to price discriminate, although the more, you know, kind of fuzzy the auction gets, the less that’s true.

The pay per click auction, you could also have people pay per impression. After all, that’s what’s being sold. Or you could have people give their whole shopping cart to Google and just pay the different price for each thing that the customers buy. Both of those are technologically possible. It turns out there’s some complications in getting the pay per conversion to really work well, and it requires a lot more data. Pay per click is sort of easiest to implement, and it manages to take some of the risk off the advertisers, which I think really got this market jump started.

However, because bids are weighted by their click through rates, there is a sense in which the pricing is on a per impression basis. You're ranked in part on the revenue that you will provide, the expected revenue, which is your per click bid times the click through rate.

The generalized second price, the early designs had pay your bid auctions, which led to cycles, and now the fact that you pay the minimum price that keeps you in your position allows for a more stable outcome. It means that small changes in your bid don't affect your outcome very much, and it allows -- and it removes the incentives of firms to kind of continually outbid each other by a penny.

Finally, the click through rating, again, it ranks firms by expected revenue for impression. The -- but it does require the estimation of click through rates. And that's actually a difficult problem on small -- on infrequently searched phrases.

It's also the case that an unweighted pay per click auctions and lead to much lower revenue. So, let you take an example. You search for Paris, you can have an ad for Paris, France, travel that gets 50 cents a click and a click through rate of 5 percent. Ads for

Paris Hilton sex videos could make a profit of $5 per click, and a click through rate of only a quarter of a percent. If you rank only by bids, Paris Hilton sex videos wins, but it generates less revenue. Okay?

So, clearly weighting by click through rates is important. On the other hand, there is a -- there is a counter bailing effect which is that an advertiser doesn't necessarily care about writing accurate text when you weight by click through rates. And the basic thing is that if Paris Hilton sex videos disguises its topic and just says Paris Hilton on it, then more people click on that. That raises their estimated click through rate, which lowers the bid they have to make to stay in their position.

And so, in fact, getting unnecessary clicks doesn't cost you an expectation as an advertiser, because every extra click you get lowers the price per click you have to pay. And so, you get this unintended consequence of the click-through rating, which is that you can get imprecise ad text. And I would argue that, you know, you do see some of that on the web.

So, let me just -- I wasn't planning to go through all that anyway. Don't worry. So, that was what I would have done if I was going to advertise my paper with Glenn.

Let me just in closing kind of tell you a little bit about that research agenda, which kind of helps think about these search costs and it tries to build a model where consumer search costs are taken into account, which would help you do welfare analysis in terms of thinking about reserved price policy or thinking about the negative effects of having imprecise ad text.

So, the -- kind of what we do in our paper is we developed a model that kind of formalizes the idea that sponsor link auctions provide information, and we show that -- we show that the reserved price policy is somewhat different in a model where you're worried about consumer search costs.

Some small or moderate reserved prices can be socially optimal because they help eliminate bad ads and direct consumers towards more relevant ads. On the other hand, reserved prices also redistribute surplus between a search engine and advertisers. And so -- so search engines will typically have an incentive to set reserved prices that are higher than what's consumer optimal in order to -- in order to extract more revenue from advertisers.

So -- and then finally as I mentioned, the click through waiting can, you know, distort incentives. And we also show that click through waiting can lead to inefficient outcome, and, in fact, eliminate efficient equilibria altogether from the auction.

So, there's lots of interesting problems left to explore in this area. And, you know, I hope that -- one side benefit of all the regulatory intervention is that now over the last two years, between Google/ DoubleClick and Google/Yahoo, lots of economists have had a chance to learn about this industry and really get into the problem and even get access to data. And so, I'm really looking forward to the next year or two in the academic community of seeing the research move forward, and also the -- some of these regulatory issues get resolved. Thank you.

Questions?

MR. DANIEL: Beat you to it, Paul. Good morning. I'm Tim Daniel. I used to be at the FTC. I'm now with NERA. Your welfare considerations, talking about whether the reserved price is set at the right level, whether there's enough -- whether there's a problem with inappropriate or inaccurate ads, that sort of thing.

My competition background, you know, sort of leads me to think, well, those are the kinds of things that regulation isn't really good at. And so, perhaps we should let competitive markets play out. And you started your talk by saying you thought that this week’s result in the Google/Yahoo was the right outcome. Yet this is a market with lots of indirect and direct network effects, and economies of scale, you talked about all the costs to develop the search engines and do those effectively in the R&D.

Tell us how you think to balance those things. You know, the competition effects -- the competition in the market, but, I mean, you wouldn’t have to worry about reserved price policies so much. Yet you have these benefits from consolidation. Yahoo decided that Google really had a good product, they wanted to back -- piggy-back on it. How do you make -- how do you make those trade-offs?

MS. ATHEY: Sure. Well, I mean, obviously these are -- you know, people will come down on different sides on this. But I think that it is a very difficult industry to regulate, especially when it’s so innovative and just even -- you know, there are -- if somebody could come up with an idea next month about pricing that isn’t even in place right now, that might be extractive or might have trade-offs for consumers and advertisers and publishers and so on, we haven’t even thought of.

And so, I do think that in the end, it’s -- it’s better for everyone. In fact, probably even better for Google to be in a sort of competitive market where we don't have regulators calling the shots and you're able to sort of innovate in a more free-wheeling way. At the same time, having the knowledge that -- you know, say you try to extract all the advertiser's surplus, you go too far with price discrimination that they have another choice in terms of -- in terms of where they can go. So, I think in the end, when you think about as a search engine changing your reserve prices for example, you think about what's the benefit for consumers, you think about what that's going to do to advertiser engagement, and you think about what that's going to do to revenue. And so, you know, especially if you're a smaller search engine, you worry a lot that when you raise your reserve prices that advertisers will just shift their campaigns to other search engines.

And so, that kind of competitive pressure can help -- I think can help make sure that all the surplus kind of gets distributed among the different parties.

MR. PAUTLER: Paul Pautler from the FTC. You may have already answered the question, which was close to mine before. You mentioned at the start that regulatory uncertainty was a real problem for these firms, and if they lose six months or a year, they're really far behind.

Then for us, the question is, okay, I understand you want certainty, but does that mean it's better to get the wrong result than the right result? And that's sort of what you were just answering.

MS. ATHEY: Yeah. I mean -- and I guess I would just add to that that -- again, I -- I see the process of having engaged with all of the different regulators and having so many people become informed, makes it much easier to then have a conversation about, you know, other things that might happen and have informed people that can respond to that. So, I think that just the general process of education is a beneficial one.

MR. SHAPIRO: Inasmuch as Google and Yahoo and Microsoft basically have different sets of users at any point in time who are searching, I know at least Google has mounted the argument that they're not directly competing for advertisers just the way radio stations in two separate cities aren't competing for advertisers because they're reaching different users. How do you see defining the relevant markets and what do you make of that argument?

MS. ATHEY: That's a good question. I think it's a -- you know, it's partly an empirical question in the sense that, you know -- I mean, of course, you know, any time you make a change as a search engine, some people are going to respond to that. Microsoft, you know, we're very sensitive to the fact that, you know, people often will choose between -- you know, some advertisers will actually just quit the platform and just choose to only be on Yahoo. So, you're very -- you're very cognizant and you see empirically the fact that, you know, changes in policy can lead to that kind of a shift.

I think that overall that's a -- it's an empirical question as to how much -- how much that happens. So, you know, and it's important to understand that -- but I think generally, you know, you're going to see in a competitive environment that, you know, when you -- when you have competitors there and people have another place to take their campaigns, that's going to be a disciplining device.

MR. SHAPIRO: Great.

MR. ADAMS: Thank you very much, Susan. Let's give her a round of applause.

(Applause.)

PAPER SESSION FOUR: DEVELOPMENTS IN DEMAND ESTIMATION

MR. ADAMS: Next we’re going to have Pat Bajari with a paper session.

MR. BAJARI: This will be a session on demand estimation. Our first speaker is Matt Weinberg.

MR. WEINBERG: Okay. Thanks for giving me the opportunity to speak here. This is joint work. I’ve got a co-author named Daniel Hosken, who’s typically here at the FTC, but unfortunately couldn’t be here today. So, because we’re both working here, the usual disclaimer applies. These are our own views and don’t necessarily reflect those of the FTC.

So, first, just a few big general big picture things about horizontal merger enforcement in the United States. So, over the past decade, there was decrease since the late ‘90s. On average, the FTC and the DOJ conduct about 75 investigations of mergers per year. And antitrust policy towards mergers in the United States, as we talked about briefly yesterday, is largely prospective. So -- because it’s very expensive to break up firms that have already merged. The regulators have to make a forecast as to whether or not a merger would reduce competition, and then they have to sue to attempt to block such mergers.

So, for my purposes, I want to talk about two classes of empirical merger studies. So, the first I'm going to classify as retrospective. And by this I mean papers that have data before and after a merger or several mergers within an industry occurred. And the goal of these papers is to estimate what actually happened to prices in the past. That's not an easy thing to do.

And, typically, what people do is they compare the change in prices and markets that are affected by the merger to hopefully a change in prices in markets that are otherwise similar but were not affected by the merger. So, the change in the prices is the baseline as would have happened in the absence of the merger. It's not often easy, but the information inside of these papers is useful. In particular, it's useful for looking back at past anti-trust decisions and getting a sense of whether or not anti-trust policy was too loose. So, you can answer that question with those papers.

But, unfortunately, it's pretty difficult to figure out how to generalize from such studies and answer the question that the guys at the agencies have to try and answer. And that is, will this new merger cause prices to increase.

And that's where the second class of studies comes in, simulation studies. So, here, by simulation study, I mean the narrow definition that was briefly talked about in the introductory panel yesterday. I mean three things. I mean an assumption that firms compete in prices, in the static or tran game (phonetic). Second, that you know the functional form of demand and can't estimate that. And, finally, there's an assumption on the shape of the firm's marginal cost functions, typically if they're constant.

And so, if you knew all those primitives and it's relatively straightforward to simulate how a change in market structure, a change in the ownership structure of the firms, would affect prices, that's great. That's exactly the question that needs to be answered.

However, the results in this exercise depend upon a lot of strong assumptions. So, those three main assumptions that I talked about. And to the extent that any of those three things don't hold, the simulations may produce inaccurate results.

So, in this paper, what Dan and I have done is we were trying to use the former study to evaluate the latter type of study. So, here's what we do. So, we've got data before and after two different consumer product markets occurred. And -- and these mergers were -- the first one was a merger of motor oil companies that combined Pennzoil and Quaker State brand motor oils. The second was a breakfast syrup merger. So, they combined Ms. Butterworth's and Log Cabin brand breakfast syrups. And so, we're not just interested in breakfast foods. We're interested in these things for two reasons.

The first is that based on -- based on public documents beforehand, it looks like they were likely on the enforcement margin. So, the -- these are pretty concentrated markets. It's likely that the products are pretty close substitutes. And so, if mergers have passed that resulted in price increases, you might expect these to be that type of merger.

The second reason we're interested in these particular mergers is that it seems like they match up with the assumptions needed in the simulation exercise fairly well. So, there are relatively few products in these markets. That's going to allow us to estimate demand in a fairly flexible way. There's not been a lot of entry or exit or repositioning of the products. So, if you expected simulations to work well, anywhere you might expect them to work well in these particular instances.

So, while it is kind of a case study approach, and ideally you would have a whole lot of these things to be able to do it many times, these are two particular cases, in our opinion, or two particularly interesting cases in our opinion.

So, a preview of what we find, first the simulations. So, the syrup merger had relatively large simulated price changes. So, typically larger than 5 percent. On the other hand, the oil merger tended to have fairly small price changes; in many specifications less than 5 percent.

So, after we calculate that, we add the post-merger data in a couple different ways. We go back and directly estimate what happened to prices. We do this in a few different ways. And the main -- the main result in the paper is that the simulations reverse the rank order of price effects.

So, here's what I mean by that. So, we got large simulated price changes from the syrup merger, but our direct estimates of price effects using the before and after comparisons are -- are pretty small. Basically, we find that that merger didn't have much effect on prices at all.

On the other hand, the oil merger had a -- had a pretty small simulated price change, but moderate actual or directly estimated price effects. So, the next step is to figure out why -- or attempt to figure out why the simulations don't match up with the actual price changes. And so, remember the three assumptions that you need are the aesthetic for training competition, the particular functional form of demand, and the constant marginal cost assumption.

So, the extent that any of these things change before and after the merger occurred, that would be on reason why the simulations are off.

So, first, we explored changes in demand. We looked to see if demand shifted before and after the merger occurred. That could be because of some sort of product repositioning, or alternatively another explanation would be that it's difficult to identify demand in different product markets, and if -- think about like the very simple case of, like (inaudible) you don't get back demand, you get back the shared demand and supply. We know that supply changes as a result of the merger. There's got to be another reason why you might find that demand changed before and afterwards.

Second, we explored changes in marginal costs. Particularly, we calculate the necessary changes inside of the marginal costs that would be required to equate the simulated and the actual price changes. And, finally, we explore a few different assumptions on our demand system; specifically, how consumers would substitute to the private side of the market.

So, briefly -- this is probably familiar to a lot of people in the audience. I'm going to describe how the simulations work. So, using the pre-merger data, we estimate three different demand systems. These are all demand systems that can be -- we estimated, like, fairly quickly. So, that's pretty good. That's -- that's the benefit of these things.

First, we estimate the AIDS system. If that were given a more descriptive name, it'd be called the proportion log price demand system. A simple linear demand system, just levels of quantities and levels of prices and some other things. That has the benefit of being able to calculate the simulated price changes analytically. And then finally logit demand. We do this under a couple different identification assumptions that I'll describe later.

So, using AIDS as an example, here's how the simulations work. So, assuming the static Bertrand pricing game, you can write the pre-merger first order conditions as follows. So, this is just like a multi-product oligopoly extension of the normal markup equal to whatever the elasticity rule that (inaudible) most likely. And given that you knew the pre-merger equilibrium prices and revenue shares, and the demand parameters necessary to calculate the elasticities at those points, you can -- the only thing that you wouldn't know in that equation are the marginal cost curbs and -- or, sorry, the points on the marginal cost curbs, and you can -- you can back those out easily. It's a linear problem.

So, you do that and you calibrate the mileage of the pre-merger data. And then you just change the profit functions to account for the change in ownership, and you re-write the first order conditions like this. It's straightforward. It's the same thing as the first one, just different ownership structure. And the -- the post-merger equilibria will be the vector of prices that satisfies this first order conditions. But it's one for each agreement in the market. And we calculate the price effects as the percentage difference between the post and the pre-merger prices.

So, data. So, we've got data from IRI. It's scanner data. And for the motor oil merger, we got data from their mass retailer channel. So, this is data that's aggregated up to the region level. So, it covers 10 different regions of the United States. We don't have store-specific data. It's at the weekly frequency, and it covers a period from January '97 until December of 2000. The merger was consummated in December of '98.

The syrup merger is from the IRI's grocery channel, and it covers more regions. We got 49, but a little bit less of pre-merger data in terms of the time dimension. So, we observed -- you know, it's like a three-way panel. We've got observations that vary by brand, region and time.

So, here's how we calculate the direct pricing. There's a slight typo in the first equation here. So, we add to the sample the post-merger data, and the first thing we do is very simple. We just compare a change in the average prices, before and after. It's a simple time difference.

So, here we've got region specific fixed effects. That's the alpha. These are months, seasonal dummies. This should really be a subscript. I do this separately for each brand in the market. And then there's the post comparison -- or the post study variables. And what we do is we make the data symmetric around the merger date. We drop an interval of three months, centered at the merger because some strange things might be happening around then. We don't want to pick that up. And -- and, you know, 100 times the beta is the percentage change in the average price.

The second thing that we do is we follow a paper by Ashenfelter and Haskin (phonetic) that computes the -- that does this for three more different consumer product markets. They look at the actual price effects for just the merging brands in that paper, whereas we're going to do that for the merging brands and also for the non-merging brands as well. And our point is not just to compute the directness, but to use that as a benchmark to compare with the simulated price changes, just to differentiate the product briefly.

So, we've got -- here what we do is we compare the change in prices to the change in prices of private label products. So, we've got regions branded. So, alpha here is an interaction between branded and the region dummies. The multi-seasonal effects, again, the post-dummy, and then the interaction between the post and the branded dummy, the coefficient on that, you'll see the change in the prices of the brand name product relative to the change in the prices of the private label product.

So, if you believe that the change in the private label products is going to be as it was in the absence of the merger, then the difference estimator would have estimated the effect of the merger on prices. If you think that the private label products increase the prices, you're getting a lower amount.

So, here's our direct estimated price effects. I've got the merged firms brands in bold. Those are Pennzoil and Quaker State, just to refresh your memories.

And we've got almost an 8 percent price increase for Quaker State, and the difference in difference specification, and nearly a 4 percent price increase for Pennzoil. Private label products, actually -- their price actually dropped a little bit here in the simple before and after comparison. So, the difference estimates are going to be a little bit less. You know, almost 2 percent less.

And the -- the rival brands for the most part increased their prices as well, sometimes substantially. We get about an 8 percent price increase for Gastrol GTX. The only exception to that is Havoline, which their price dropped by about 4 percent. And I've got some stories based on marketing documents for why that was the case, if you're curious, later on.

So, just to walk through a simple -- a simple example slowly, here's some simulated price changes. So, this is -- we estimated an AIDS system, an AIDS demand system, on the pre-merger data, calibrated the marginal costs, and then simulated the price changes. And in parentheses, we've got 90 percent confidence intervals. And -- and this looks remarkably close to the directly estimated price effect. So, you see the merging brands. It even gets not only the magnitude right, but also the rank order. So, the simulation actually is picking up that Quaker State is going to simulate its -- or going to increase its price by more than Pennzoil.

On the other hand, the price effects, the simulated price effects aren’t as close for the non-merging brands. They tend to be smaller. That’s important because, you know, obviously consumer welfare depends on what everybody is doing, not just the merging brands. And so, the next thing that we do is that we -- instead of estimating by OLS, we try an IV strategy. And so, remember how I said the data was structured. It’s this three-way panel. This is like a pretty typical thing to do. A lot of people do this. And you’ve got -- so you’ve got prices of other regions. If you think that there’s going to be a common marginal cost component, then those prices in other regions are going to be correlated. And if you think the demand stats are independent, then those would be good instruments.

In our data, we didn’t get very plausible demand parameter estimates out of that exercise. Sometimes we get cross price elasticities that make the products look like compliments instead of substitutes. I think that motor oils and breakfast syrups are substitutes, that this happens. And while the model is predicated upon all those things being right, for completeness I went ahead and simulated what would happen if you used those things. And the results are a little wild.

So, next -- again, usually if somebody was doing this, they would look at the (inaudible) and probably wouldn't go forward with that part of the exercise. But for completeness, I put it there. Thanks. So, here are the other specifications. And the conclusions are roughly the same. So, linear demand gives a slightly smaller simulated price effects. And the logit model, we get really small cross price elasticities. If you look at that, that -- and that's going to give you very small price effects. The non-merging firms, their price effects are second order things, and so, when the merging firms are barely increasing their prices, you're just not going to get any movement in the non-merging firms.

Here are the results for the breakfast syrup merger. So, first, start on the left, the first column. We don't find much evidence that this merger caused prices to increase. And that doesn't really depend upon our -- our method for estimating the direct price changes, although you get slightly bigger price effects in the straight difference estimator.

On the other hand, the simulated price changes can be pretty big. So, the AIDS model, we're getting simulated price effects of about 20 to 24 percent. Now, this is pretty remarkable to me. This is a three to two merger. Again, you -- the products are likely pretty close substitutes. And it didn't affect prices. The simulations say that they would for most specifications. And -- and that gives me some pause.

So, if you move across specifications, we get smaller price effects for the linear demand system, and the logit demand system (inaudible) because of the linear one, and the specification.

So, the next thing that we do is we try and figure out what could explain the discrepancy between the simulated and the actual price changes. So, the first thing that we do is say, well, we need to assume again that demand is constant before and after the merger occurred. So, what I did is I took the post-merger data and I estimated demand on that. So, if it had shift, and we are identified, then using that should -- we should be right on, if everything else is okay.

So, here's what I find when I do that exercise. It does slightly better in some specifications, but overall the conclusions don't really change that much, particularly for the syrup merger.

So, the next thing I do is I calculate the percentage changes in marginal cost that would be necessary to equate these two things. So, focus on, for example, the AIDS system for the syrup merger. Those are pretty big. The first column. Don't pay much attention to the IV. We found that they need to be, like, between 22 and 24 percent. But that's pretty big given the technology of breakfast syrup. I mean, that stuff is like sugar water. It's like corn syrup and, like, something that smells like maple. That's the marginal cost of breakfast syrup. So, it's unlikely that that fell by that much.

I'm out of time. Okay. So, let me just get to the conclusions. So, again, the big finding here is that the simulations reverse the rank order of price changes. We had one merger, the direct estimates, they -- they seemed to imply -- they implied modest price increases, but the simulations gave small price increases. On the other hand, we got another one with no price effects. So, even though it was a three to two, that didn't go through with the right thing. It didn't reduce consumer surplus. But the simulations gave large price effects.

Just to -- just to compare this to the only other work that we know that's directly comparable to ours, Craig Peters has a paper that was mentioned briefly in the panel yesterday in which he does a similar exercise for five airline mergers. And our results are similar to his. So, he also finds that the simulations reverse the rank order of the price changes. So, I'm sure that you guys don't remember the slides from Mike's talk yesterday at the panel, but he found -- he found the same effects.

Some of the airline mergers had big price effects and they seemed to have the lower simulated price changes. So, thanks again for giving me the opportunity to talk here, and I look forward to your comments.

MR. BAJARI: Our discussant will be Matt Osborne from the Department of Justice.

MR. OSBORNE: Okay. So, as Matt discussed, what this paper does is it looks at how well merger simulation does in predicting the price effects of mergers. Now, the agencies would care about this because we have to predict what a merger is going to do before the merger actually happens. There's a lot of different tools that we use to do that. But one of the tools that we use is merger simulation.

So, as Matt discussed, the basic exercise here is you estimate demand and then you come up with a model of industry structure, which is often Bertrand, and then you feed the demand estimates into this model and then simulate it to try and figure out what the effect of the merger is going to be on prices.

So, they do this merger simulation on two different mergers. So, one of them is a merger of motor oil products. And what they find is they observe that the actual price of both of these products goes up, which suggests that maybe the merger should have been blocked before it happened. But the merger simulation predicts a really small pricing freeze. So, if we were just going by the simulation, we would have probably let the merger through.

And then if you look at the syrup merger, the actual effect of the merger seems to be ambiguous. So, the price of one of the products goes up and one of them goes down. But the merger simulation predicts large price increases. So, if we were going by the merger simulation, we would have stepped in and blocked the merger when that might not have been the right thing to do.

So, what the paper does, then, afterwards is it looks at some possible explanations for why we get these -- for why these merger simulations just don't seem to be working so well. So, they look and see, you know, could it be driven by demand specification. Some of the demand models don't give very reasonable-looking results, but some of them do. So, it may not be the demand specification.

They also look to see whether demand could have changed post-merger. So, they re-evaluate the exercise on post-merger data. That doesn't seem to do it. Another thing that they do is they look and see, well, what sort of marginal cost changes would rationalized the observed price increases? And the type of marginal cost changes that you need just don't look very plausible when you think about what these industries actually are.

And then the last thing they do is they look at the sensitivity of the result to market size assumptions. And that doesn't seem to be driving -- driving the results as well. So, the conclusion seems to be that if you're a practitioner of merger simulation, you're kind of in a conundrum. I mean, these merger simulations just don't seem to work very well.

So, let me talk about some -- let me give some of my comments. So, you know, I enjoyed reading this paper. I think this was a fun piece of paper -- or a fun paper to read. It was good work. I think the policy questions are important. It's -- it's clearly written. And I like the econometric work. I think it's nice and carefully done. So, as an example, the author is recognized, you know, for when you compute these counterfactual price increases, you can't use the delta method to calculate their standard errors. You're going to get the wrong results. So, they use a parametric bootstrap to do it.

I think there are some areas in which the paper could be improved even more, though. And so, let me talk about some of those. So, my main worry with the paper is that people may end up seeing this to be too -- as being too similar to some work that Craig Peters did, which Matt has cited. And what Craig does is a very similar exercise for a number of airline mergers.

So, let me suggest some ways that maybe the authors could broaden their conclusions a bit and build on what Craig has done and will differentiate a little bit more from what Craig has done.

So, one thing that would be interesting to see would be maybe a different demand specification used. So, it's like I felt that some of the demand specifications were a little bit too -- perhaps too restrictive. You know, for example, in the AIDS model, while potentially it's a very flexible model, there's a curse of dimensionality when you estimate the model and the number of parameters that you have. So, what the authors do is they aggregate out some product characteristics like the weights of motor oils.

If they were to say use a random coefficients logit specification, you know, like was done in (inaudible) they could include those as product characteristics. So, one suggestion might be to use -- to see what a random coefficients logit specification would do, because that's a baseline for a lot of work.

I think, though, an even more important point is that there doesn't seem to be much discussion on what sort of alternative competitive models would -- might explain these results. So, if you look at Craig Peters work, you know, his -- he finds that marginal cost changes don't do a very good job -- okay, I'm almost done -- don't do a very good job of explaining the results. And his conclusion is that, well, Bertrand is not a very good assumption.

So, I think it would be interesting to see some sort of other simple competitive models used, like maybe -- we know that there's retailers and manufacturers in these industries. Perhaps there's a Stockelberg (phonetic) game being played, or there's some sort of tacit collusion going on. And I had some other sort of smaller comments, but I'll give them to you after the -- I'll discuss them with you later, Matt. So, thanks.

MR. BAJARI: One or two quick questions for (inaudible).

AUDIENCE MEMBER: (Off microphone) (Inaudible) three months right around the merger, but I'm wondering if another possible explanation besides the (inaudible) demand is just that even though you removed the immediate (inaudible) insinuations are about a static, steady (inaudible) and what you’ve got is (inaudible).

MR. WEINBERG: Yes. There’s pretty limited empirical evidence on the long run results of horizontal mergers. The only one I know of is this paper by Focarelli and Panetta that look at banking mergers in Italy. And they find that immediately after the mergers occurred, prices went up, but then in the long run -- this is kind of a hard thing to figure out, but it seems like prices initially fell.

So, the story is that market power, you can exercise that like immediately, but marginal costs, it takes a while to garner these efficiencies. Again, I don’t think that they’re going to be -- I mean, this is me saying, but it’s not based on data, but I think that it’s unlikely that there are going to be big efficiency gains inside of either of these mergers to marginal costs. Right? That’s what you need? Yeah.

Yeah?

AUDIENCE MEMBER: (Off microphone) (Inaudible). First, I think scientifically there’s something a big flawed about (inaudible). (Inaudible) and we don’t want to give the research (inaudible).

MR. WEINBERG: Right.

AUDIENCE MEMBER: (Off microphone) So, scientifically the only way right to do this is to actually get the (inaudible) before anyone knows (inaudible) you know, put those in a (inaudible).

MR. WEINBERG: Right.

AUDIENCE MEMBER: Okay? And then see what happens.

MR. WEINBERG: Yeah, that would be -- that would be excellent.

AUDIENCE MEMBER: (Off microphone) (Inaudible). And the other sort of part of this is also, so you didn't (inaudible) talk about the mergers, but just in the one you showed us, I mean, I think that in some (inaudible) to express (inaudible) six players or five players.

MR. WEINBERG: Right.

AUDIENCE MEMBER: (Off microphone) (Inaudible). So, again, without knowing any of the (inaudible) very, very important. And without knowing any of them, we probably would (inaudible) larger in the syrup? Right. I think (inaudible). So, you know, the (inaudible) before going to the mergers, whatever (inaudible) it's really not about, you know -- I mean, one thing specific about the (inaudible).

MR. WEINBERG: Okay.

AUDIENCE MEMBER: (Off microphone) So, I think that's (inaudible). That's comment number one.

The other question -- the other comment on the -- you guys presented this, and I (inaudible) discussion went along (inaudible).

MR. WEINBERG: I agree. Some of the -- in particular the --

AUDIENCE MEMBER: (Off microphone) (Inaudible). One is, you know, hey, we got one of these (inaudible) I'll take those off any time. But the other thing is, you know, basically you showed us garbage in, garbage out, right? We (inaudible). So, I think it's still worthwhile to figure out what happened to the syrup case. But overall, you know, this is (inaudible). (Inaudible) and now we have to explain to (inaudible) figure out what it is that we're missing. (Inaudible).

MR. WEINBERG: All right. Yeah, thanks. So, first briefly, the -- the goal in the study was to do what I thought as a non-FTC employee at the time would -- what you guys in agencies would do on the pre-merger data. So, that's exactly it.

The actual things that the FTC and the DOJ are -- sorry, the FTC in this case, that would have handled these, the retail consumer product mergers, what they were thinking exactly, that's private information. That can't be discussed. It's all proprietary. The -- and personally, I don't even know it. So, the -- I mean, I can guess, but, like, nobody has told me anything.

So, the other thing is, if you look at the demand elasticities for some of the specifications that do lead to wild -- or not wild, but, like, inaccurate price effects, they look plausible. Like, if somebody handed you those demand elasticities for the syrup merger, estimated by the AIDS model, and you just saw the elasticities, that's it, you looked at those things, you would think, no, okay, they look reasonable to me. But they still give simulated price effects that are 23 percent bigger than what the direct estimates are.

On the other hand, yeah, the oil results are something that look pretty good. And so, I also view that as encouraging. And I think that this is -- the policy question here is just so huge that, like, it's -- this is a benchmark to guide future progress. And that's how we'd like the paper to be viewed. So, I look forward to things like the rest of the sessions. I should let the -- let it get on with. So, thanks.

MR. BAJARI: Our next speaker is Jeremy Fox from the University of Chicago.

MR. FOX: Okay. This is joint work with Che Lin Su, who's here at the conference, and Jean Pierre

Dube'. Their affiliations are now at the University of Chicago Booth School of Business, and I guess that's ane example of display advertising.

So, if you learned about econometrics from Art Goldberger or somebody, you probably heard about something called the Best Linear Predictor. Well, fortunately, that acronym has been stolen by some self-promoting (inaudible) economist, and it's now known as Berry, Levinson and Pakes, which is this very commonly used demand estimation method. And it's a pretty helpful technique because it allows us to talk about demand and differentiate our products industries where we have all these product characteristics. It's a fairly flexible specification. It doesn't impose as many restrictions on elasticities from functional form. We can use with commonly available aggregate data sets, and we can control for price endogenetic using instruments as we saw in Matt's favor.

What did BLP do? They kind of introduced in some sense a computational algorithm to compute the fiscal objective function, and allowing these -- this kind of complicated non-linear demand model. And then, you know, as we saw, it's very useful in a lot of applications for measuring market power, computing demand elasticities and so forth. And it's been used extensively by Aviv (phonetic) and others in applications.

So, I think I'm going to start from this point of view that demand estimation is a very useful technique for both the research and policy work. You know, I've gotten some attention from this from some European anti-trust agencies as well, and it seems like at least in some anti-trust agencies are entrusted in this type of technique.

And the down side is that, you know, this method is not easy to use for someone who has not been trained to use it. So, if I just gave a grad student a copy of BLP's econometrics article, told them to go code this up and produce estimates, you know, who knows what would come back? You know, probably not the correct estimates. And Aviv has been a leader in trying to give some advice to (inaudible) uses here.

So -- but the concern, I think, is potentially from, you know, people within the literature. And outside of the literature are these estimates coming back from this somewhat complicated method, the correct ones. And there's really no point in doing a complicated method if you're not going to do it correctly and produce the right estimates. And, you know, there's actually another paper out there in the literature by Chris Knittel and

Metoxaglou (phonetics) saying that, you know, this is -- you know, basically giving warnings that this might be not always producing the correct estimates.

And, furthermore, you know, Robin and others are doing work on BLP and models of (inaudible) consumers. So, the consumers are also solving a dynamic programming problem. And so, the research frontier in this demand estimation work is to go into more and more complicated papers. And then, you know, that's great in terms of research, but it also, you know, is a good time to kind of take a step back and make sure everything is going exactly right.

So, what we're going to do is document some potential computational concerns about BLP and maybe offer some solutions. So, for those of you who don't know what's going on with BLP, there is this computer program that's kind of embedded inside of BLP. So, you're both searching over parameters like you would in any non-linear econometric model. But there's also this kind of inner loop, which is a step where you're trying to solve a system of equations. And I'll go over that -- in detail what that is.

BLP developed a computer method called a contraction mapping to solve those systems of equations. And our basic point is that this computer loop -- inner loop is not always going to produce numerically the correct answers. And the researcher might have an incentive to make that computer loop a little -- inner loop a little inaccurate in order to speed up the results. So, if you have to go to a conference at the FTC in five days and, oh, no, my routine isn't working so well, so, let me just -- and it's taking too much time, let me just cheat a bit on this inner loop. Then that's going to produce numerical error and that might lead to wrong parameter estimates.

And this has nothing to do with the statistical properties of BLP. If it's coded correctly, it's purely a computational idea. Do you have a question or --

AUDIENCE MEMBER: (Inaudible).

MR. FOX: Excuse me?

AUDIENCE MEMBER: (Inaudible).

MR. FOX: So you'll have a computer program called this inner loop that's both -- and I'll explain what that is in detail. But the idea is that it's going to stop at some point, and it can stop when it's really, really accurate or just stop before then. And it's stopping before then, which saves time, but might introduce error.

Okay? And so, we're going to produce an alternative method to solve some of these issues called

MPEC, which stands for mathematical program with equilibrium constraints, and some other work. Che Lin has been investigating the properties of this and some other types of economic models.

Just to get up front, just to clarify what's going on, MPEC is not going to be a new statistical estimator. It will be a new computational approach to computing the same estimator that we've all been doing.

So, our contributions are going to be we're going to talk about BLP's approach, show that this can -- if you don't do it right, can lead to the wrong estimates; introduce MPEC as an alternative, and it's not going to have these numerical problems with this inner loop. It could work faster in some cases, which we'll be explicit about, and it could -- I won't talk about this at today's talk -- apply to models more generally, models where we don't have a contraction mapping property where they could be in some cases multiple solutions. And this might be important for some of these new dynamic demand applications.

And it's particularly -- and we're not going to talk about that today, but we're trying to push this in terms of these dynamic demand applications. That's like a new frontier where MPEC could be especially useful.

So, I'm going to go over the model pretty quickly in the interest of time because a lot of practitioners are already familiar with this. It's going to be micro-founded by a demand specification. We have a bunch of product characteristics for each product.

BLP studied cars. Think about cars having miles per gallon, fuel economy, speed, different measured characteristics. They'll have a price. They'll have a demand shock, which as you see this Greek squiggle letter here, that's allowed to be -- you know, that's going to be product in market specifics or product J and market T. And there's going to be some individual specific errors, which are logit. You pick the product at the individual level and it maximizes utility. We have aggregate data, individual data.

So, we're going to just aggregate up this demand specification to the market level by integrating out these error terms. There's two different types of error terms. Your different preferences for these different car characteristics, like some people care about speed, some people care about fuel economy, and that's -- there's going to be some distribution of that, Epha Beta (phonetic), and Epha Beta is indexed by some parameters data. And that's our goal of estimation, is to estimate these distribution of preferences.

The main point of this thing here is we have an aggregate data expression here. Inside of this aggregate data expression are these demand shocks. These squiggle marks exceed J and T, the demand shock for product J and market T. Because these things enter this equation non-linearly, it's going to be hard to back them out of the equation, which is kind of like an additive specification where the error term is just sort of sticking -- floating around there, and it's easy to back out once you guess at the parameters data. Here the error terms enter the model very non-linearly.

So, because of this complicated functional form, for every guess of data, we want to evaluate what are these error terms. We're going to have to compute the error terms numerically. And what BLP will do is, you know, they have a computer program called a contraction mapping that's going to solve this problem.

For each guess of these parameters, we're going to iterate on this inner loop and we're going to keep doing this. We're going to compare our guess of market shares due to actual data in market shares, and if we're within some error tolerance, which is, I guess, an answer to your question back there, we'll stop the inner loop at some pre-specified level when our changes and our guess of these demand shocks stop.

And then this is a nice approach because it’s guaranteed to find a solution from any starting values. Once we do that, we’re going to evaluate a condition that says our demand shocks are uncorrelated or they’re instruments, sort of a standard IV approach, and we’re going to plug in our demand shocks to this equation. And there’s going to be two approaches to then doing this.

So, what BLP will do is they minimize this objective function, which is just sort of a weighted product, or these demand moment conditions that says our demand shocks are not related to our instruments. But it requires sometimes they guess at new value parameters to back out these demand shocks using their model. We’re going to say another approach to doing this, which might be more common in a numerical methods literature, which is to do a constrained optimization problem where we’re going to maximize the objective function subject to the constraints that these -- at the solution that these market shares predicted by BLPs demand model (inaudible) data on market shares.

So, our alternative approach, we’re going to be minimizing over both structural parameters data and these preferences and these demand shocks.

So, I’m going to skip to -- I’m going to go through these slides relatively quickly for these theories. A contraction mapping has a certain rate of conversions, which is determined by something called a Lipshitz constant. Okay? The Lipshitz constant in the BLP demand model had something to do with the elasticities of demand in respect to these demand shocks, squiggle or exceed. That's a complicated expression, but that's what it is.

We have a bunch of theorems in the paper. In the interest of time, I won't really go over these. But the idea is that there is some error coming from these inner loop errors, propagating to our objective function. If we do what BLP suggests, use this so-called nested fixed point approach.

Furthermore, these problems will be especially severe when we use numerical derivatives to -- in part of our optimization approach. Here I'm kind of comparing the radiant based solvers, which are the techniques used in a lot of applied work, starting from (inaudible) work, have error in them if you approximate the rate of numerically in that error is going to be compounded a lot if you have an inner loop that's nested.

Well, I'll just demonstrate these parameter errors. As an aside, we're -- I think it's important to use a professional, high quality optimization package. So, we're going to use this commercial program called nitro. We're going to code it up in MATLAB.

So, here's the first example of some errors and mistakes one can make. So, there's going to be three algorithms here. These are all BLPs nested fixed point approach. There's going to be one approach where this is sort of the first column. It's sort of the impatient researcher who has to go to that conference in a couple days and sets the inner loop tolerance to be too loose. So, here tenant (phonetic) minus four, but keeps his outer loop setting, the tolerance for choosing our structural parameters, to be the default setting of tenant minus six.

Then the second -- what's going to happen for this researcher is his routine is never going to converge. I'm going to report solution found. We can see it on the first column, first row, where it says zero percent of runs, the routines had report conversions.

The second column refers to a -- reflects a reader who -- researcher who says, well, a solution to that problem not finding conversions is to set my outer loop tolerance to be low. So, now -- to be loose. So, no I'll just accept anything that looks like a -- vaguely like a solution and call that a solution. Well, that will solve the problem of what your routine is reporting, but that won't produce correct parameter estimates, either.

And the third column is kind of the correct researcher who's set the inner loop tolerance to be really tight. What we'll see in the first two columns that we're getting really different, so we have this one data set here, many starting values. The first two people who have the wrong settings are getting kind of crazy estimates that have nothing to do with the truth. We see that if BLP is done correctly, it does produce an estimate very close to the truth. But the first two columns people are just getting all sorts of crazy answers depending on your starting value reflecting these.

And then I didn't go into very much detail, but how these new -- these are the results that we predicted by the numerical theory that answers are crazy.

Now, because your answers are so crazy, a careful researcher in this example would have said these results don't make any sense, I must be doing something wrong. If the person really did try multiple starting values and got these crazy elasticity estimates that don't have -- that vary a lot by starting value.

Now, another example, we took -- and, by the way, the previous slide relied on using numerical derivatives in your solver. Here is an example using --

we’ve actually coded up BLPs derivatives analytically, doing some additional programming work, and we used these serial data. And here the two kind of wrong methods produced the wrong estimates. So, the true -- the correct estimate from this data set is for own price elasticities, negative 7.4. This is serial. But in using these bad implementations, we’re getting negative 3.7 or so as the elasticity. So, we’re off by an order of magnitude almost.

And -- but we see that the bad methods consistently get the wrong answer, and the same wrong answer. You get a standard deviation across starting values is really low, suggesting that even a careful researcher here who wasn’t aware of these issues could produce the wrong estimates, even if they were trying many different starting values, just because we’re always converting to the same wrong point.

As an aside, the problems we’re finding are quite different than those in this paper by Chris Knittel, which, you know, we took a look at their code and found that they were finding all sorts of crazy estimates but for different reasons. And what they’re using is sort of solvers that are not high quality, and they’re reporting solutions that the solver is saying are not true solutions.

So, the problems with BLP really aren't about multiple local optima, which is the message you would take away from that other paper. Okay. So, I think these are important issues. We need to code up on stuff correctly.

Just briefly, an alternative suggestion is to minimize the objective function over both demand shocks and structural parameters subject to the constraints that these hold, that there's going to be no inner loop here. So, there's going to be no error from one part of the computer program ending up in the other part of the computer program. So, you don't have to spend any time fooling with tolerances on your inner loop to make sure that's correct. It's nice we can prove mathematically that solutions to the original problem that BLP wrote down are equivalent to the solutions to our reformulated problem.

So, the main advantage of MPEC is not going to have these numerical errors. It can be faster for a variety of reasons because you're giving the solver more information about the problem. And there's various reasons it could be faster, you know, it could be a sparser problem and various things. And it can be applied to problems more generally where the contraction mapping doesn't exist, potentially including problems where there's not a unique solution.

We introduced earlier this thing called Lipshitz constant, which is a measure of kind of the speed of the nested fixed point inner loop. We can -- in here, this is varying the data generating process and seeing how close this thing gets to one, which is a measure when it's going to be slow. And we're just doing some speed benchmarks here. And we see that when the Lipshitz constant gets closer to one, the speed of the nested fixed point approach gets really slow.

And that's kind of the concern we might have about this frustrated researcher who in some data sets is going to have a really slow inner loop. Well, that's when the approach is getting really slow, is when the researcher might try to cheat. And MPEC is going to solve that problem.

So, there's some speed comparisons here. And we saw that in this speed comparison, and sort of the CPU times at the main column, MPEC was relatively invariant to these changes and the data generating process that made nested fixed points slow. Statistically, these are the same estimators as seen by having the same bias and root-mean-squared error across the two specifications.

And one concern you might have about MPEC is that, well, you know, it's not going to work if you have a lot of different products because you’re optimizing over these demand shocks, and that’s equal to the number of products you have. Here we’re increasing the number of markets. This is a very high dimensional problem when we’re seeing that MPEC is not slowing down disproportionate to NFP. In fact, here it’s kind of -- I don’t think this is a robust result, but NFP itself is getting very slow as we increase in our markets, not MPEC.

But also we ran it on Aviv’s data again, and we found that MPEC was slightly faster although the differences weren’t as huge on his data set. So, all we do in the rest of the paper is do dynamics. But I don’t have any time to talk about that. That’s kind of an advanced topic anyways.

So, to conclude, I think we’re positive on BLP. It’s a useful technique. It can allow us to do a lot of cool stuff in demand estimation. Unfortunately, its implementation could lead to errors unless you’re informed about these types of problems and you plan around them in your implementation.

MPEC is an alternative that doesn’t have exactly the same sensitivity to these types of errors, can be faster and could apply more generally to more complicated models. Thank you.

MR. BAJARI: Our discussant is Denis Nekipelov from U.C. Berkeley.

MR. NEKIPELOV: So, in this paper, the authors are sort of -- the authors are -- were trying to infer the -- we're trying to infer the reference parameters in the standard differentiated demand model. And the utility is very standard. Jeremy has discussed that. And the nature of the inference problem that we're trying to solve is that we have panel scanner data. We observe some demographic information about the consumers.

In most modern data sets, we can also link individual demographics with individual purchase characteristics, and we can track consumers over time. However, for some markets, we only have aggregate market share data.

The assumption that we're making about the preferences here is that there are some unabsorbed characteristics. And we can find a set of instruments such that the unobserved characteristics means they were given these instruments.

What this implies in general from the econometric perspective is that if we're observing the data on individual purchases, then we have two moments. We have a system of two (inaudible) moments. The first moment will map the individual choice decision to the market shares that we are computing from the data, and the second moment is going to be the restriction on the unobserved characteristic.

So, the standard approach that we're going to follow is, first of all, we're going to put a lot more parametric structure on the system of moments. First of all, we're going to parametrize the shocks in the utility, and we're going to parametrize the distribution of random coefficients and the preferences of consumers.

The standard approach later on in the analysis, in the empirical analysis of differentiated product markets is that we're assuming that the first moment in the system isn't exactly quality. We're going to invert that and substitute the solution for the random coefficients into the second equation. And this is the way that's been used to solve that type of problems.

In the paper that I'm discussing, Jeremy and his co-authors are pointing -- pointing us to the fact that if we are using some of the iterations in order to do the inversion of this first equation in the system that will lead to -- that might lead to numerical errors in the -- in the estimation procedure and they provide a superior approach to standard contraction iterations, so they assume that we can actually define that first equation as an exact equality. Then if we define it as an exact equality, then the GMM problem becomes the optimization problem with the constraint. There's not going to be any need to control for the quality of approximation of the first equation, and it will lead in many cases to faster convergence.

So, in general, we can apply this framework in a lot of different settings, such as dynamic demand and dynamic models that Jeremy has mentioned.

Well, just kind of to explain the idea of the paper to myself, I'm looking at the following -- at the following very simple examples. So, suppose that we're trying to compute a numerical derivative of some function. Then usually if we take the find a difference approach, then we take the symmetric differences around the true value, and if we look at the (inaudible) expansion for the function that we're trying to numerically differentiate, you see that magically the second derivative is going to disappear. And the order of approximation is going to be the square of the step size of numerical differentiation.

When we have an error in evaluating the function that we're numerically differentiating, then unfortunately that error is going to propagate into the find a difference. So, what it means is that we're going to have one over the step size part in the numerical derivative, and that's going -- that might even lead to the loss of the first order approximation. So, that's why it actually is very important to control the quality of approximation of the function that we're trying to minimize or differentiate.

So, in general, I think the numerical properties here is very important. And we actually need to control very carefully the intermediate computational step, structural step, in the estimation exercise. And in general, the same arguments will apply to a lot of other quasi-likelihood and quasi-Bayesian type procedures.

And the authors give constructive advice for implementing these procedures in practice.

My comments are the following. First of all, I think that the way the paper focuses on numerical problems actually undermines the statistical aspect. And in a lot of cases, actually just the statistical noise in the objective function can lead to the similar results for the numerical -- for the numerical derivative and for the optimum.

Secondly, it seems that the constraint optimization procedure has obvious statistical problems. And, first of all, if we're looking at that as a GMM problem with the constraint, then the test statistic is not going to be squared as in a standard Houseman type of test. And what this means is that it will be very hard to use something like that for model selection or model testing, or specification testing. And so, I guess I'm just going to move directly to the end of my discussion.

So, first of all, I was -- I was going to say that the paper gives very important results about the relevance of numerical approximation. We can use it to improve computational performance of the differentiated demand estimator. And although this method is more interpretable, explicit inversion of fixed effects is not necessary for inferential purposes.

The real advantages of the method, when we're using the precise computations, is, first of all, we are producing more -- something which is more (inaudible) to the errors in large deviations, and that's going to be very important for counterfactuals. And, secondly, we can provide much higher precision for computing the welfare or the revenue measures in the models defined by differentiated demand. Thanks.

MR. BAJARI: In the interest of staying on time, I think we're going to postpone questions for speakers until the very end. So, let's hear from Katja Seim from Wharton.

MS. SEIM: All right. Well, thank you very much for having me. This is joint work with Michaela Draganska at the Stanford GSB and Mike Mazzeo at Kellog. And as the title suggests, what this paper is trying to do is look at how firms make product assortment decisions. And by that, what we're going to mean is how firms choose which subset of an existing portfolio of products to offer.

So, we're not going to be looking at how firms decided to position products and characteristics per se more generally, or how the decision to introduce a new product is made in terms of characteristics. Instead, what we'll be looking at is purely assortment choices. And the way the paper proceeds is to develop and estimate an empirical model of a firm's pricing and assortment decision.

We then look at a number of counterfactuals to try to look at how important consumer demand is in driving firms' choices, to what extent product assortment choices reflect back on the prices that we see in the market, and then lastly, which I'll spend time on at the end, to look at how market structure and changes in market structure affect the assortments that we see.

So, you know, why might you think that is interesting? I think, on the one hand, it complements existing work that looks at how market structure affects prices. So, if you thought about the effect of a merger, there oftentimes people look at what the likely price effect might be of that. We're going to also look at assortment choices. On the one hand, because the types of consumers that are affected by that are likely to be different, price effects in general tend to affect the marginal consumer who might choose not to buy any more as prices increase.

In our case, if you choose to fully recondition your assortment, you're actually going to affect the infer-marginal consumers as well.

On the other hand, we also think that these types of decisions are an important practice. And I've just put up a bunch of examples of settings where you think assortment choices are here, you know, product choices for a multi-product firm are relatively easy to adjust as a result of a merger in the short term and in the long run obviously as well.

So, store locations, closings, openings, adjustments to flight schedules, adjustments to the network, and then the last example that I have here is adjustments to radio formats. And there's actually some work there on what the affect of mergers might be on the variety of radio stations that we see by Barry and Goldberger (phonetic), and what they find for example is that in their sample of data, consolidation in value actually led to increases in the variety of radio stations that we see.

So, this would be a setting where consumers don't generally pay prices to consume the product, and so, these variety increases would actually make the merger beneficial to them from a welfare perspective.

That's likely not going to be true in a lot of other settings. And so, one thing that we're interested in is looking at isolating the effect on welfare of assortment choices, as well as the effect on welfare on prices.

In terms or where this paper fits in the literature, we're going to think of assortment choices as basically an entry game. There -- it's in the spirit of, you know, literature going back to Breshnahan and Reiss. This literature tends to think about at least initially relatively homogenous industries, and as a result capture demand in a reduced form by mostly characteristics of the market. That's not going to be very helpful for us in that the -- by default, thinking about firms that offer differentiated product, and so, it will be more useful to have a properly specified demand side to that.

So, we're going to merge into this literature then a more structural demand model. I guess here I'm benefitting from being the last in the session, so you've sort of seen how these approaches work. We'll use a very standard, discrete choice demand model that is very much like what Jeremy talked about. Sadly enough, even more simple, and I'll talk a little bit about that at the end.

And we're going to then take this model and as an application look at what kinds of estimates we get for the ice cream market. But I think it could be easily any kind of setting that you might be interested in.

So, I'm going to give you a quick overview of how the model works. It's going to be a two-stage game here that firms play. They're first going to choose which set of flavors, in our case here, or product more generally, to offer out of an existing portfolio of products that they have available. And then they're going to give them the assortment choices that they and their competitors may choose how to set prices.

As I said, our demand side is going to be a discrete choice model of demand at the flavor level, so the product level. We're going to use a random coefficient specification and have a logit demand shock. So, we'll get the usual logit demand estimates back from that.

In contrast to a lot of the other literature here has done, we're going to control for unobservable attributes of flavors, another demand shocks, primarily by including a host of market characteristics and time and flavor dummies rather than explicitly controlling -- (inaudible). And I talk a little bit at the end why we do that.

On the front side, we're going to look at the two-stage decision process. We'll have two types of costs. On the one hand, there will be a marginal cost to producing a product. In our empirical setting, the ice cream market, these are going to be primarily cost shifters of inputs, capital labor, et cetera. We'll assume -- which I think probably makes sense in our setting, that these are common knowledge. In contrast to what Carl talked about yesterday, our data on these marginal costs is actually very basic. And so, we will assume that there is unobserved component to marginal cost. And you'll see later how if you had better data, I think you could do much better on this front.

We'll also assume that firms pay fixed costs to offering a particular flavor. And so, what we have in mind here would be things like distributional costs of getting flavors to stores, the slotting fees that the brands contract over with the stores and having them on the shelves. We'll assume that these are flavor specific, that they're information to the firm only, but not observed by its competitors. And then in our empirical estimation about further distribution to them and assume that they're like normal.

So, the effect of these assumptions on the model are going to be, first of all, on the pricing side. So, the second stage, we'll assume that firms compete in Nash-Bertrand pricing. We are going to use that assumption to recover marginal costs from pricing first order condition. And based on these costs, together with our demand side, predict what a firm's profit would be under all possible assortments that they could offer in the market.

So, we can then look at the assortment choice where each firm will choose to offer that particular assortment that maximizes profits. This, the assumption that the fixed costs are private information here, will get an imperfect information equilibrium in the flavor strategies.

So, just to sort of point out some of the challenges in this particular setup, I've drawn here a little example of how this would work in a two-by-two case. We have firms from one's decision, they can offer up to two different flavors or offer no flavor at all. They are going to make the decision of which assortment to offer by comparing expected payoffs of the various choices. And the expected payoff of any given choice here is going to reflect what they are going to make in profit under each of the alternative assortments that firm two could offer, rated by the probability that firm one thinks firm two is going to offer that assortment.

And so, as this flow chart, I think, tells you, like the main difficulty in this literature is really the dimensionality of the problem. As you keep adding flavors here, computationally it’s going to be increasingly difficult. And so, in our empirical application, we’re also going to focus on a pretty small -- small scale example. This is more relevant for estimation because you keep solving the model over and over than it might be for the actual counterfactuals.

So, what we’ll do is we’ll do an estimation, start a demand side, calculated predicted market shares; use those together with the observed prices to figure out what the firm’s marginal cost would have been, and compute variable profits for all different assortments. Based on that, then derive what the equilibrium assortment offering strategies might be, and minimize the difference between what we observed in our data on prices, quantities chose and strategies to what our model predicts.

So, the data that we use is exactly the same data that Matt talked about in the beginning. So, it's IRI data at the market level. We have data from 2003 to 2005 for 64 markets. This is where they are. The data contained prices, quantities, information on sort of the flavors that are offered. And we're going to look at decisions at a monthly level, which is where we see some variation in -- in flavor offerings.

We'll focus on the vanilla subcategory here, which is about 25 percent of the ice cream market, and look at regular ice cream sold in three and a half to four pint packages. So, this sort of shows you a breakdown, we'll roughly capture 80 percent of the market that way.

The firms that operate in this market are really two types. We have Breyer's and Dryer's. They are national brands present in all of our markets. Then we also have a pretty large set of sizable regional firms that are listed here. They provide quite a lot of variation in the competitive environment in local markets. So, as you can see, they're not available in all of the markets over time.

The right-hand side here of the table just shows you differences in the number of flavors. Vanilla flavors that we see offered across markets. We're going to, in estimation, focus on the choices of the national brands. What you can't see here is that most of the variation in the numbers of flavors that are offered over time due to, one, in the case of Breyer's, and, two, in the case of Dryer's vanilla flavors that they offer selectively in different markets.

The regional and general not always have less variation in the numbers of different vanilla flavors that they offer across markets, and so, they're not model, they are strategic choice on that front.

So, we'll have demand for all models and for all brands in the market, and look at Bryer's and Dryer's decisions of offering the flavors that go in and out at the product choice stage. So, we'll call these optional flavors versus the staple flavors that they offer basically all the time in all markets.

I'll just go through our estimates here quickly. We have our random coefficients logit demand, which is on the right here. We observe very little about -- little measurable information anyway about attributes of the brands and flavors. So, we include brand and flavor dummies to capture heterogeneity and demand. Our price effect here implies an elasticity that's in line with the literature. So, that's maybe quite reasonable. And then we capture the demand for vanilla in general relative to the outside option of another ice cream by a host of demographic attributes of the markets.

On the cost site here, our marginal cost estimates, these are mostly, like I said, input price shifters. And if you look at the precision with which we're able to estimate these, in general not very pinned down. So, this would be an area that better data would really help.

We have one brand-specific cost shifter, which is the distance to the distribution center or transportation cost. Most everything else does not vary over brands, and just in general doesn't have that much variation.

And then lastly, the fixed cost estimates that we cover based on an assumed like normal distribution of the shop to offering a particular assortment, imply average and median flavor offering costs for a given month of, you know, on the magnitude of several thousand dollars, which is in line with the variable profits that we estimate for these flavors over time.

So, let me just turn to what we want to do with these results now that we're done. We're going to look at a bunch of counterfactuals. I'll only talk about the merger analyses that we conduct where we're going to contrast what happens if Breyer's and Dreyer's were to merger into a single firm, and offer the same assortment, which we'll call fixed product, to what happens if they're a duopoly and what happens if they adjust their assortments after the fact.

Now, as you can imagine in this kind of situation, the actual configuration and competitive environment in a market is going to matter a lot. This first example is one where we just basically took our empirical setting at face value and looked at what kinds of effects we get. And here the effects are very small, both of the merger in general and of androgenizing the assortment choices.

This is due to, first of all, vanilla being only a small share of the ice cream market; optional flavors being even smaller than that. And so, we're sort of looking at a merger here of products that are quite small relative to the big picture. In addition, the flavor offering costs are also relatively low.

And so, as an alternative, we looked at what would happen if we focused on the optional flavors only, so had Bryer's and Dryer's only, offer those, and assume that the market was smaller so that their overall share of the demand was significantly larger.

And then we're going to contrast our estimated fixed costs with a scenario where we jump up fixed costs of offering a flavor by a factor of one and a half. And so the main things to take away from this are the following: Both of these results, the settings give you pretty similar implications. And, first of all, you know, as we go from duopoly to any kind of a monopoly situation, prices increase. They tend to increase more with the settings that we've looked at so far for the case where we hold products fixed as opposed to the case where we allow firms to adjust their assortment.

In both of these situations, firms tend to decrease the number of flavors that they offer. And in terms of sort of how that's broken down between the three flavors that we look at, they tend to sort of decrease all of them as they go from duopoly to monopoly.

The effect of that on consumer surplus is going to be, you know, a reduction in surplus, both because prices increase relative to duopoly, but also because variety falls. And once we andogenize choices, the change in surplus also reflects that relative to the fixed products case, prices are actually not quite as high. And so, these two tend to offset each other.

So, let me just conclude here in terms of where we want to take this going forward. I think what this has shown you is that, you know, the results that we would expect to see from a merger on assortment is going to matter on the particular case study, which is not surprising. We are also able, for example, to come up with similar predictions to the Balferger (phonetic) Berry setting where variety actually increases as a result of a merger, which might actually mean that consumers are better off. And this provides you with a -- there's a setting that you can look at that.

There are a number of things that you might not like about the way we do this. I think there's things that we can do to improve on our demand side, sort of following on what Jeremy said. There's also things that we can do on how we estimate the product assortment game between firms, drawing on the recent literature.

What we're most interested in for now is actually looking at, you know, how the results of the predictions here will change instead of looking at a model where assortment is driven by fixed cost differences between firms. What would we get if instead we looked at a model where assortment is driven by selection in that there are unobserved things about demand and cost that firms might know that affect the selection that they make in a particular market.

This is more difficult in terms of solving it, which is why we started with this one. But I think having information on both of these would give you a nice picture of whether assortments matter in a particular case, and if so, how much. Thanks.

MR. BAJARI: Our discussant is Minjung Park from the University of Minnesota.

MS. PARK: Okay. Let me briefly summarize the paper. So, on the demand side, we have a discrete choice model for differential products. And the model allows random coefficiency. There's no site (inaudible) that represents an observed product quality.

On the supply side, we have an assortment decision in the first stage, and then firms engage in Bertrand-Nash pricing game in the second stage. And the fixed cost introduction, which is relevant for the first -- first stage decision, is assumed to be private information.

So, the author's applied a supply and demand model to the market for vanilla ice cream, and their paper shows that to get the count affecters (phonetic), it is important to first incorporate indulgence product choices, and also it is important to model demand and pricing decisions directly instead of using a reduced (inaudible) function.

So, this paper is very well motivated. I think most people in this room would agree that it is important to look at this issue. And the authors do a very good job of doing that. So, thank you, Katja, and thank you to the co-authors.

So, here are my comments. So, incorporating indulgence product choices in the demand model means that the typical (inaudible) instruments for prices are no longer valid.

Also, unless we are going to submit the demand and supply side at the same time, do simultaneously, we also need to start worrying about instrumenting for those indulgence characteristics.

So, when I first started reading this paper, I was getting all excited because I thought that the authors might have some ideas to share about, you know, how -- how to find instruments when we have indulgence characteristics. But it turns out that they didn't need to confront that issue because there is no (inaudible) in the demand model.

And I understand why they need to do that. You know, this issue about inferring the sites for those products that were not chosen, but I still think they sort of missed the opportunity. I mean, their setting is such a netrous place for them to discuss, you know, a lot of the alternative good instruments, or more generally how to make the demand side when we have indulgent product characteristics.

So, if the authors could offer some ideas on this, I think potential readers of this paper would appreciate that quite a bit.

My second comment is that product assortment decisions seem to be a dynamic decision, or at least it seems to be state dependent. So, for example, the fixed costs of introducing a product the second time around might be a lot more. Or if there's a serial correlation to fix costs, then a firm might be able to learn about its competitors fixed costs over time from the previous decisions. And the authors sort of assume away these issues and in their application they assume that the assortment decisions are made each month for each market separately in aesthetic fashion.

So, I think one simple way to check whether this concern is relevant for this particular market is to report the times where it's appropriate (inaudible) product offerings, so we see the products are offered for many months in a row and didn't get dropped, or do we see that they are offered on and off?

So, if you see the latter pattern, it might suggest that it's not such a big concern for this particular market.

So, what about dynamics on the consumer side? So, I don't really know much about this market, but the consumers have strong brand loyalty in this market. So, suppose the consumers have loyalty at the brand level, and they also seek variety of flavors. If that's the case, the firm might have an incentive to introduce a new flavor, even just so that they could lock in those consumers at the brand level, although the particular flavor itself might not be individually profitable.

Or it might be that it takes some time for consumers to get used to or try new products. And, again, if this is the case, a firm might have an incentive to introduce a new flavor, although doing so is not individually profitable for that particular period. And these conditions sort of make the optimality condition that you use for product offering to be incorrect, and in that regard it would be nice if you could provide some discussion about, you know, consumer behavior in this market.

So, for ice cream, we have a very simple form of differentiation. For many of the products, they like you to have multi-dimensional product differentiation. And we are likely to encounter the curse of dimensionality, as she mentioned in the discussion -- in the presentation.

So, just to get a sense of how serious this issue might be, and also just to get a feel for how feasible the proposed methodology will be for potential users of this approach, it would be nice if you could report, you know, how long it take to submit a model when you have one dimensional differentiation, two, three or four, those cases.

I sort of found it intriguing that these firms charge the same price for all of their flavors. And it also helps simplify the analysis in the paper. So, it would be interesting to know what are these features that sort of justify the practice of uniform pricing in this market. And also just, you know, in addition to that, in Monte Carlo, can you actually -- if you try -- can you actually show that the uniform pricing decision to lead to a lot -- much reduction in firms profits compared to unrestricted pricing, optimal pricing behavior. So, that would be sort of interesting to know on the side.

So, last two comments. So, they used to make these fixed costs from the optimality conditions for product offerings, and they find that the fixed costs differ greatly across flavors for a given firm. On the other hand, when they submit the supply side, they assume that the marginal cost is the same for all flavors in the same market, for a given firm.

So, it’s kind of -- it’s kind of strange to argue that the marginal cost is the same, but fixed costs are very different.

Last comment. So, in merger simulations, I think eventually we would like to allow firms to introduce new products that were not present in the market previously. And if that's what we want to do eventually, then we'd like to sort of map this production to the characteristics space so that we know how close they are. And, you know, then for that we need to know -- how the consumer substitute patterns among these products.

So, in that sense, it'd be nice to buy three gallons of ice cream and try to come up with some measures that can map these flavors into the characteristics space and see how close they are. And I'll be very happy to offer my help for that task. That's it. Thank you.

MR. BAJARI: Well, I'd like to thank our authors and discussants for three interesting papers. And let's go have a little bit of coffee.

(Paper Session Four concluded.)

PAPER SESSION FIVE:

ECONOMICS OF NETWORKS AND THE INTERNET

MS. ATHEY: All right. So, let's get settled. So, where is David? We're missing a speaker here. David? Man of the hour. And you're ready, so we're going to do about 25 and five, so that way when I say zero, you've got, you know, 30 seconds or something. So, let's get started so we have some chance to get everyone out on time.

So, our last session is on, again, the topic near and dear to my heart, economics of networks and the Internet, and our first speaker is going to tell us how that advertising works. So, take it away, David.

MR. REILEY: Thanks. This topic of how does advertising affect sales is something that has interested me since I was a graduate student. In fact, I had hoped to write my dissertation on that topic and I discovered that all the data that I had been collecting for the professor that I was working for were not actually going to be able to identify these effects in a way that I was going to believe. So, that's when I switched to studying online auctions and running experiments.

Since I'm now working at Yahoo! research, I have some really great opportunities to return to this question that interested me from the beginning. And, you know, a couple of people have said, gee, you know, there are ads on all of these pages that I browse, but I basically ignore all the ads, and so it's sort of -- it's an interesting question. Are there people who are actually looking at them or are these things affecting us subconsciously or do they have no effect and, you know, people don't -- advertisers are wasting their money on these things.

I know economists are always assuming that firms are behaving optimally, but having worked inside a firm now, I'm pretty critical of that assumption.

So, I'm really excited to be able to talk about the effects of advertising on sales. This is joint work with -- this is joint work with Randall Lewis, who is a Ph.D. student at MIT, and was a summer intern with me at Yahoo! this summer.

So, the outline is, why is it hard to measure the effects of ads on sales, what's the experiment look like, what's the data look like. Then I'm going to talk about basic treatment effects from the experiment that we ran. Then I want to talk about what happens, sort of what are the long-run effects of the advertising campaign that we did as an experiment. And then I'm going to talk about some more detailed results if I have time.

So, this is John Wanamaker, department store retailing pioneer, and he famously said, half the money I spent on advertising is wasted, I just don't know which half. And this has been -- this has been my experience that advertising is -- it's not easy to quantify the effects of advertising, and it's hard to know where the advertising you're spending is actually having an impact for you.

So, to substantiate my claim that advertisers do not have good measures of the effects of brand image advertising, I want to cite a Harvard Business Review article published this year by the founder and president of ComScore, and in this article, he talks about measurements of the effects of advertising on sales.

So, ComScore is the largest Internet data firm. They have a panel of over two million customers worldwide have who agreed to let ComScore track everything that they do in their web browser. And, so, Abraham describes in his article the methodology here is simple, we take those people who saw ads for a particular good and we compare them to the people who didn't see ads for the particular good, and then we survey them to see whether they bought it or not.

The potential problem with that methodology is that the two samples don't necessarily come from the same population. Let's, as an example, take -- oh, I should also say, in this article, he says, search advertising has the clearest effects on sales, right, because they've done hundreds of these studies, or dozens of these studies at least, sort of summarizing them, and search advertising you get, you know, three times as many sales as with the people who saw the ads as those who didn't.

So, let's think who's seeing search advertising, who's seeing the E*Trade ad. The people seeing the E*Trade ad are the ones who searched for online brokerage and the ones who are not seeing the E*Trade ad are the ones who did not search for online brokerage. Do we really want to attribute any difference in sales to the fact that they saw the ad or are they different people, much more interested in making a purchase independent of the ad.

So, as we all know, correlation is not the same as causality, but it's really hard for people to remember this sometimes, right, and it's actually a more -- it's a pretty subtle point, and is often hard to get right, this identification problem.

So, measuring the effects of advertising on sales has been difficult for economists as well as practitioners. The classic technique that was used was econometric regressions of aggregate sales versus advertising over time. Marketing professionals do this and call it marketing makes modeling. And it's literally a textbook example of the endogeneited problem in econometric, see Ernie Barron's book on chapter 8, which is uses advertising in sales to illustrate endogeneity. You know, what causes advertising to vary over time? Well, you know, sometimes you run the regression and you get a positive slope, and then you realize, oh, gee, firm was setting advertising as 10 percent of sales, right, and so which way does the causality actually go?

So, there's two ways for observational data to provide inaccurate results. Aggregate time series data, the advertising doesn't vary systematically over time. You have endogeneity, individual cross-section data, you have admitted variable bias if you compare people who saw ads to people who didn't see ads. And so, you know, my point of view has always been, when the existing data don't give us a valid answer to our question of interest, we should consider generating our own data. And I think our experiment is the best way to establish a causal relationship.

So, we're going to systematically vary

advertising, showing as to some consumers and not others, we're going to measure the difference in sales between the two groups of consumers, you know, and this is almost never done in advertising, either in online or traditional media. Some exceptions, direct mail marketers are really good at doing experiments, and in search advertising, there is some degree of experimentation going on.

I claim that our understanding of advertising resembles our understanding of physics in the 1500s, and Galileo's key insight was to use the experimental method. It's not sufficient to observe that a bowling ball falls faster than a feather. You want to try to control everything, take the same shape and sized items and have one be wood and one be brass and then see which one falls faster, right? So, we're going to try to do controlled experiments here.

Market is often measuring the effects of advertising using experiments, but not with actual transaction data. So, typical measurements done by marketers come from questionnaires like do you remember seeing this commercial, how positively do you feel about this brand, you know, what comes to mind first? What brand comes to mind first when you think about batteries?

So, it's maybe useful comparing two different creatives, two different, you know, kinds of ad copy, but do these measurements actually translate into effects of advertising on sales? To my mind, it has never been documented.

We're not the first to do advertising experiments, so I want to make sure people are aware, if they're not already, about the IRI behavior scan split cable TV experiments. This is a great idea. They managed to get the cable business, you know, kind of late eighties, early nineties, they managed to get local markets with cable TV customers a split signal so that some households saw an ad and other households didn't. They had a scanner data card -- it was a scanner card, frequent shopper cards for individuals in the panel.

So, they knew which ones had seen the TV ad and which ones hadn't, or at least which ones had been delivered the TV ad and which ones hadn't. And then they get measured effects on sales.

Unfortunately, they tended not to get statistically significant effects of anything. And it's not -- it's not clear whether it was the advertising didn't work, or whether it -- the power wasn't high enough. They had a sample size of about 3,000 households.

And if you read one of the review articles like Lotus' 1995 article in the Journal of Marketing research, you see that their summary, meta analysis of 300 different tests, is that 30 percent of the tests were significant at the 20 percent level of significance. So, there's only a very little bit there. They were being pretty generous using a 20 percent significance level and they still had a hard time finding anything significant.

Okay, I'm going to skip a couple of other things here. Well, I should also say some studies derived valid insights from nonexperimental observational panel data. Example being Dan Ackerberg's work on yogurt, where he had individual diaries of TV ads, sample of 2,000 households, and can't actually, you know, get some measurement of the effects of advertising on sales by seeing how individual purchase behavior changes over time. And, so, you can sort of have -- you can get rid of the individual heterogeneity problem by using a panel.

So, our study combines a large scale experiment with individual panel data. We matched the Yahoo! ID database with the database of a big nationwide department store, and by matching email and terrestrial addresses, we got 1.6 million people identified as matches. We then put 80 percent of the matched customers into a treatment group, who saw three ad campaigns on Yahoo! from this retailer. The remaining 20 percent got into a control group who didn't see any of the ads in any of the three campaigns. These ad campaigns are run of network ads on Yahoo! which means you might have seen them in Yahoo! mail or Yahoo! finance or Yahoo! autos or Yahoo! home page.

Following the online ad campaigns, we received both online and in-store sales data each week for each person. So, this retailer is pretty good at keeping track of each individual when they buy in the store, you know, at the cash register. They, you know, if you use your Visa card and they know your name in their database already, they'll record it under your name in their database.

To protect customer privacy, we had a third party do the matching of the data and then de-identify it. So, neither the retailer or Yahoo! knows all of the data attached to a particular user identity. And the retailer also disguised from us what the actual dollar sales amounts were by multiplying by a random number.

Randall and I have reverse engineered what that number is, so I will be able to tell you -- I will be able to make some cost benefit comparisons for you. I'm going to report everything in the retailer's fake retail dollars. So, I'm measuring everything in another currency, you know. I'll tell you it's kind of the same order of magnitude as real U.S. dollars, but it's not -- it is a transformation.

So, ads on Yahoo! look something like this NetFlix ad on the right-hand side of the page. I'll blow that up for you. Of course, we weren't experimenting with NetFlix, we were experimenting with a department store, and the content of the ads was what we -- people in marketing call retail image advertising. It wasn't particularly emphasizing any price or product. I mean, they were advertising each of the three campaigns advertising a different product line, but it was heavily emphasizing the store name and, you know, we want you to come into the store.

So, by the end of the three campaigns, we treated over 900,000 people. So, you can see campaign one, early fall, lasted 14 days, 30 million ad impressions, we delivered 30 million ads on pages. Campaign two was ten million, in the late fall, and campaign three was 20 million -- 17 million in January.

So, not everybody in the treatment group saw ads. So, my graph here shows the control group with sort of 20 percent, all the way through, the treatment group was -- is purple, and part of the treatment group never actually saw any ads. As you can see how it's kind of -- how we saturated the treatment group over time and didn't -- never got anywhere near 100 percent saturation with them.

I'm pretty happy with the randomization we did in the experiment, because we do have some demographic data and some other variables that we can check. And, so, the same percentage female in both control and treatment, of course the percent of people who saw ads is different, because I delivered zero to the control group, but percentage of people who saw any page views on Yahoo! at all during the 14-day campaign is exactly the same. Number of page views per person is about 360 in both cases.

So, down on the lower right, I'll give you a few more statistics that may be interesting. About 25 ad views per person in the treatment group, and that's averaging even across the 36 percent of people that don't see any ads. 0.56 clicks per person. The click-through rate, which is kind of the industry standard for measuring performance of an ad campaign was up 0.3 percent, which is kind of an average click-through rate for a display advertising campaign. But I'm able to compute what I think is a more interesting number, which is what percent of customers actually clicked on an ad, and that's 4.6 percent of the treatment group.

The number of ads delivered has a skewed distribution. That bump on the right-hand side is actually -- you know, I'm top coding some observations there, and so actually the maximum is way the heck, you know, across the street, with 6,000 ad views. It's hard for me to imagine that that was actually a few men seeing 6,000 views of this ad, because only about 15 percent of all pages shown on Yahoo! had this ad campaign on it. All right, I have to speed up.

In-store sales are big compared to online sales, blue versus purple here, and there's a lot of variance from one week to the next. I have a little hole in my data there, in December, where I wasn't able to get the sales data.

There were lots of individual outliers. You can see, you know, in the first week that I have data, the mean sales are 93 fake cents per person, and the min is minus $932, the max is plus $4,000, fake dollars. This is a retailer who's pretty generous in accepting returns, so I think I actually believe the minus numbers. You know, none of these data were hand coded at any point. These are all directly from computer records from the register.

So, not all the treatment group members browsed Yahoo! enough to see the retailer ads. 36 percent of them in the treatment group did not see ads. So, I can assume that in the control group, 36 percent of them behaved in such a way that they would not have seen ads if I had tried to show it to them. Unfortunately, I don't know how to cut out the red people and just examine the green people, you know?

So, I'm going to be able to first measure the treatment effect on the intent to treat, but that's not so interesting in this case. You know, it's not like it was a take-up rate decision where the individual said, oh, yes, I want to see ads, or no, I don't. It was, you know, did the person happen to browse in a way that resulted in their seeing an ad? So, this is going to result in dilution of my treatment effect measurements.

So, the descriptive statistics are $1.84 in the control group is mean sales, $1.89 in the treatment group. So, I got a five-cent increase due to ads. The effect is not significantly significant. Even with 1.6 million people. And I sort of think looking for the effect of advertising on sales is a bit like looking for a needle in a haystack. Right? There's huge variance of sales across individuals. I can't expect advertising to explain very much of it, particularly just this one ad campaign on Yahoo!.

It was a little disappointing, not to get statistical significance here, and so, one of the problems is I have complete noise for 36 percent of the data because I know these people didn't see ads. So, here's another thing that I want to look at. Suppose I hadn't done an experiment and suppose I just looked at the treatment group, right, I just ran an ad campaign to these people. Some of them ended up seeing ads, some of them didn't, just like in this ComScore study I talked about.

So, here, ads decreased sales by 23 cents per person. Big negative effect, if you do it this way. But it's not really a causal thing. It's got to be admitted variable bias. Not just because I don't believe that ads would have a negative effect on sales, but because I have pre-campaign data to show it.

Here it is. Before the campaign, those who are going to end up being exposed to the retailer's ads by $1.81, those who are going to not end up being exposed to the retailer's ads by $2.15. And that's a big statistically significant difference. And this is just totally admitted variable bias. I'm not comparing apples to apples, right? So, this shows what's going on.

And, in fact, if you look at the time series differences, the control group falls from $1.95 to $1.84, so a decline of about ten cents, with some rounding, looks like 11, but it's actually ten. And if you look at the bottom line, those who were not exposed to the retailer's ads, those sales fall by ten cents. But if you look at the people who were treated, their sales stay constant, right?

So, it looks like we had a period where people bought more and then followed by a period where people bought less, and the ads prevented sales from falling by as much as they would have if they hadn't seen ads.

So, I'm going to skip that.

Very interesting that the distribution of sales looks so similar across treatment and control. And there's some very small differences that I magnify here, but I have to move on.

So, we can compare sales directly between treatment and control and correct for this 36 percent dilution. But that doesn't help us with standard errors, because I just -- I have to scale up both the estimated coefficient and the estimated standard error.

Another thing I do is repeat that, but exclude those people that I know had zero Yahoo! page use during the campaign. I can't observe -- I wish the ad server behaved this way. I can't observe that somebody showed up, you know, I would have delivered them a retailer ad, and then the ad server says, oh, wait, they're in the control group, I can't show it. I wish I had been able to record that event, but I can't. And I can't observe that somebody didn't show up to Yahoo! at all.

So, out of the 36 percent who didn't see ads, two-thirds of them didn't show up to Yahoo! at all and I'm able to remove those entirely.

The third technique I can use is a difference in differences. I'm going to skip over this. So, the first two ways that I estimate, if I rescale my effects, the numbers that I really want you to look at are the purple numbers at the bottom. So, if I just look at treatment/control differences and then I rescale the control for dilution, I have an eight-cent treatment effect, but a six-cent standard error. And then if I exclude the people who had zero page views, I get a nine-cent treatment effect with a five-cent standard error.

So, I'm getting closer to being statistically significant, and the estimates are being consistent with each other, so that's nice. But let's look at the difference in differences estimator. So, basically, I'm going to assume that I have individual fixed effects alpha I, and if I take time series differences, I get rid of them. And, so, the estimated effect that I get is ten cents with a standard error of four cents, which is a estimated sales impact for the retailer of $83,000 plus or minus $70,000 at 90 percent conference interval. Compared with the cost of those ads being $20,000.

So, you know, it looks like we're getting a positive return to advertising, and it does seem to be statistically significant.

Let's see, what can I say in my one remaining minute? I have a specification test that makes me feel good about the difference in difference molds. And, so, then we ask about persistence. And we say, you know, gee, what happens after the two weeks are over? We get a -- we get a treatment effect for two weeks of ten cents, we get a three-week treatment effect of 16 cents, that is the single week -- the third week of after the campaign is over, has a treatment effect of six cents, so a standard error of 2.4 cents.

So, there are statistically significant effects of the campaign even in the third week. And then we thought, well, gee, if we have it in the third week, I wonder what happens in the fourth and higher weeks. So, we plot -- we plot our treatment effect estimates, and they're always positive. What we have is sort of cumulative effects of all three ad campaigns at any given point, or if you've been exposed to two campaigns, what I'm estimating is the effect of having seen two campaigns.

And I am out of time. So, I just want to present one more -- one more result, which is if we compare offline and online, the total effect is 16 cents, the offline sales are 15 cents, the online sales are one cent. So, seeing ads is causing people to buy in stores. And this is surprising to a lot of people who think online advertising only drives sales of online retail.

The other thing I want to point out on this table is those who view ads but don't click them, there's a big effect on them, too. Not as big as the effect on those who click. And, so, a 50-cent treatment effect on those who click compared to untreated people, 14-cent on those who view it but don't click, but 93 percent of the treated individuals are those who view but don't click.

And, so, it turns out that 78 percent of the total treatment effect is those folks.

Okay, I am going to stop.

MS. YIN: So, I think this paper is just really cool in that it's yet another example of how, you know, it's really important to take these lab experiments and put them in the market because I think you just get something that's richer than either just doing the lab experiment or just, you know, doing exercises in a market. So, yet another sort of testimonial for that.

And, in particular, I -- you know, so unfortunately, David couldn't get to a lot of -- David couldn't get to a lot of the data that he actually has in this. It's a rich paper full of all sorts of observations and, you know, I think maybe there might be multiple papers, because there's so many other, like once you see the data set that he has, which is matching these individuals at the score, online, and, you know, in their advertising exposure, there's just so many interesting things to explore.

So, what I think are particularly nice is that he clearly identifies the selection issue in here. Like this is something that I'm going use in my classroom, right, because I really want to train the MBAs, like this is a problem when you guys go out and you try to interpret these results. You should really be careful of this idea.

And, basically, the example I'm going to show them is what David showed to you in terms of he's got random buys who possibly get the treatment, but even in this exercise, selection occurs within that set, right, because some people who are going to potentially be exposed to these advertisements are just not going to surf on Yahoo! at that point.

So, selection occurs within the set, and those are interesting observations that I think are garnered with the detailed individual level data, panel data. In particular, this persistence effect of advertising, as well as the difference between offline and online results, and I think that's something that, in particular, hasn't been exposed to a detailed and credible level yet.

But let me just, you know, caution David and Randall when they're going forward in some of the robustness checks to still be careful of that selection effect that they still have in their model, right? So, they treat -- possibly people can see the selection effect, but if you just don't go on Yahoo! you don't see it. And what he's going to do in some of the later campaigns is actually say, look, I'm just going to control the treated group with the other people who never saw the ad. So, that includes the people in the control group, as well as those people who should have been exposed to ads, but weren't. And David explained that already.

And what I just want to point out is that you have to be cautious of doing that a priori in general, because what he's going to do, he's going to say, well, here's the treatment group, maybe there's a high type and low type, right? And here's the control group. And they're involved with a high type and low type. And let's say these types are correlated with your propensity to actually get online on Yahoo!.

Well, the problem of comparing this control group is that you're going to estimate in the differences, the first difference is that they go from, you know, spending $13 to seeing the ads and spending $19, so that's a difference of $6. But you're going to average out all of these differences, right? This control group with this little portion of the treatment group. And you're going to say that, oh, well, that effect, that difference was $3, the difference in these two differences, that second stage is $3. When in reality you know that this difference is actually $1, right?

If you could separate -- this is what David was talking about. If you could separate out and just compare this part of this control group with this part of this treatment group. That's the effect you want to have. So, there's the potential for this bias.

Now, one thing that's very comforting is that David actually checks in his data on this first campaign, and shows that, well, in fact, you know, all of these average out to about $2. So, it seems like this part of the treatment group is actually the same -- has the same behavior as this control group together.

So, in that case, we wouldn't worry so much. But what I would encourage him to do when he's looking at these persistence checks later on, is just to check that are we comparing -- is there a control group, whoever we put in this control group, including those untreated treatment effect people, and the actual treated people, are they reacting to external shock in the exact same manner over time, right? Because that's kind of what you need with that difference in differences. If you have these two groups reacting to shocks differently, then it's not clear that you can do that.

So, maybe one way to address, you know, sort of the distribution problems, take a look at Beatty (phonetic) and Nevins and see that they have a way to nest differences in difference. Also, they show this ad about how these net -- these advertising effects are kind of counter-cyclical with sales cycles and I want to suggest that you check other options. Maybe there are lag effects, because these advertising campaigns are not only happening online but offline in TV and newspaper.

So, maybe just off line people react to things a little bit differently. As well as the hot topic are all these social networks and they have information from Yahoo! on the social connection of these people, whether they're involved in Yahoo! groups or not. So you can say the persistence effect might actually be connected with that kind of behavior.

And then, you know, how many licks does it take to click. I think it would be interesting to see, you know, if every one is watching 25 ads on average, do they click on the first ad or do they click on the fifth time that they see the ad? So, again, you know, all of these questions can be addressed in their -- with their data sets. So, I think it will be exciting to see the next versions of these papers and even find out more interesting facts about how retail advertising works. Thanks.

MS. ATHEY: (Off microphone). So, we're running a little bit behind, so we can visit questions at the end since we took a little longer. Any questions? Hearing none.

(Laughter.)

MS. ATHEY: No questions? Do you have the data set? Okay, so up next, we have interactions between organic search and paid search.

MR. GHOSE: Hi, everyone, my name is Anindya Ghose, and I would like to thank the organizers for having our paper here today here. This is joint work with one of my colleagues who is also in marketing, Sha Yang.

What I would like to talk about today is some of the work that we've been doing over the past two years or so with different advertisers who advertise in a sponsored key words on search engines. And as I talked through some of this work, I talk about the fact that none of the data sets that we've had so far is perfect, because, you know, every data that we worked with has some limitations.

But one of the interesting things of working with the advertisers is that you get to see the conversions from those ads, which is -- and we have also tried to work with Microsoft, but it's really that much harder for the search engines to have data on conversions, unless the advertisers actually release that data and allow them to install, you know, our cookies and so on.

So, I talk about trade-offs in working with some of this data, as I move along. All right. So, what is this all about? Well, just to sort of set the stage up, imagine a consumer who is searching for betting on Google, right, and this consumer then gets to see a bunch of letter links on the search engine page and many of us might have seen something similar or even come across and played around with this.

So, the one set that you see on the top have three slots and the ones we see on the right-hand side, these are all sponsored links. These are all based on auctions, standard advertisers that you see over here have worked with, and, you know, place bids in combination with the bid price, click on it and they get ranked. And then we also get to see what we refer to as organic or free links. In other words, these are the links that come up for free, and I say that for free, because there are some investments in landing page optimization, but it's free in the sense that there's no explicit auction going on over here.

So, the question we're trying to address in this paper is, well, if I'm an advertiser, and my key word or my link shows up for free on the organic side, should I even bother to invest and place a bid on the paid side, okay? I mean, paid ad is costly, you have to pay on a per click basis, the organic links are free. So, does it help me if consumers get to see my link on both the paid and organic side, is that a good thing for me, do I institute need to do that? Does it hurt my conversions? So on and so forth.

So, what we are trying to do is look at the interaction between paid and organic and then try to get a sense of whether this interaction is positive or is it not positive at all and is it negative, is it a substitution effect or a complementary effect. Okay?

So, the agenda, sort of two full agendas I have today. One is I'm going to be talking about what kind of ads drive variation in consumer demand, drive variation from the point of your purchase, click-through conversion and so on, and in particular I will be talking about specific attributes of these ads. Whether these advertisements actually release the advertiser's information or whether it's just a brand information that's encrypted in the ad or whether it's a very generic ad that doesn't specify the brand or the company.

And we'll also be -- so, if you also looked at this from a simultaneous sense, have you looked at it not only from the consumer's side but from the advertiser's side and from the search engine's ranking side. And it's important because these decisions are happening at the same time. So, it's important to factor in all three entities in the same model.

And the second point of this paper is look at this, you know, this interrelationship between organic and paid listings, and think about, well, is this a positive relationship, does it not matter. And the second question is, is it symmetric or asymmetric? In other words, does paid -- having your name on the paid ad, does it actually help your clicks and conversions on the organic side more compared to vice versa.

So, that's where I'm headed to with this paper.

All right. So, to, you know, give you a preview of what we have done. I'll be talking briefly about the model. It's a hierarchical based framework where we're trying to model all the three entities simultaneously, consumers, advertisers and search engines.

And we -- I am going to talk about the kind of ads that drive this variation. In order to examine the relationship between organic and paid ads, we built an auto logistic model. The auto logistic model helps us look at the sign of the interaction, and we find that there's actually very strong and statistically significant positive interdependence between organic and paid click-throughs.

In other words, if I'm an advertiser and I am showing up for free on the organic side, it does pay for me to be on the sponsored links and have my links there as well. So, the probability of click-through on the page, on the organic or the paid side increases if my ad shows up simultaneously.

Now, one of the limitations of an auto logistic model is that we have to assume symmetric ties, so we also look at the discreet game entry model in which we look at asymmetric effects. In other words, what we found out we do as well, is the effect of organic links on paid search, stronger or vice versa, and so we find that the effect of organic links on paid searches, paid links, is stronger than the other way around. Almost three times higher.

We do some positive simulations to back out increase in profitability and we find that there is between a 4.5 to 6.7 percent increase in profits for the advertisers who when they show up on the sponsored side, given that they also are showing up on the organic side.

And then, finally, if I have the time, I talk about a very interesting field experiment we did. So, what we did with this advertiser is we asked them to stop sponsoring a certain set of ads for two weeks and then track the conversion on the organic side only, and then we asked them to resume sponsoring those ads for another two weeks and we tracked click-throughs, conversions, revenues, from both organic and paid.

And, so, we repeated this experiment four times over an eight-week period. And this helps us tease out a little better the effect of organic search on paid listings, okay? And, so, that explanation also corroborates that the net, the click-through rates, conversions, revenues and profits from having both organic and paid significantly compensates for not having paid.

All right. So, I'm not going to have too much time to talk about prior work, but just to give you a flavor of what's been going on. There's been a lot of work done in this area from economics, in mechanism designs and options, so we have been doing some of that. In computer science, people are looking at algorithms, you know, what kind of algorithms best optimize these mechanism designs.

So, our paper is sort of looking at more of a business perspective, information systems/marketing perspective, and what we're really interested in looking at is, as I said, looking at user firm, modeling firm and user behavior and then search and then performance.

So, this requires internal advertising level data, and that's, you know, what we've been trying to do over the last couple of years, work with different companies and get the conversion rates as well. And, so, that's sort of the framework for, you know, where I'm headed at with this study.

So, what -- let me first talk about, you know, what kind of attributes we are looking at. So, the first one that we're looking at is the presence or the absence of the retailer's information in the ad. Okay? So, suppose someone types in the key word Kmart bed sheets. That's an example of an ad where the retailer -- that advertiser's name has been exclusively shown, so that's one kind of an ad.

The second kind of an ad is when the consumer is sort of agnostic to the retailer, but he or she is interested in the specific brand. In this case for bed sheets, a consumer is interested in buying Nautica bed sheets and he doesn't care or they don't care if they are buying it from Kmart or Wal-Mart or target or L&T, but they want to buy Nautica bed sheets. So, that's what they search for.

And then you've got the extent of specificity. So, some consumers in our data, you've seen them actually searching for cotton bed sheets, whereas others are very specific and they type in a query as long as 300 cotton Egyptian count bed sheets.

So, we've seen all kinds of, you know, specificity in these kind of advertisements. And, so, those are the three attributes we will be looking at in the study. So, whether the ads have retailer information, whether the ads have brand information, and whether the ads are longer or shorter. And then we can always look at how these ads are associated with click-throughs, conversions, ranking and cost per click.

So, to give you a flavor, I'm not going to have too much time to walk through all the details of the model, but this is a hierarchical based model, and we resolve it using Markov Chain Monte Carlo method, Metropolis-Hastings algorithm.

So, one of the models of decision making over here -- well, from the consumer point of view, there are three decisions. As you go on in a sequential manner, first the consumer searches. Based on a search, they get to see an ad. When they see the ad, they makes clicks or not. Based on the click-through, they decide whether they want to purchase or not. So, you have three decisions on the consumer point of view, the advertiser's cost per click decision and the search and rank decision.

So, we go in to solve these five models in a simultaneous conversion framework. So, let me first address the data. So, just from the purpose of NDA, I can't disclose the name of the company, but you can guess based on some of the ads that I was talking about. This is a Fortune 500 firm, you know, which has 700 stores in the U.S. and internationally. We've got a -- we've got a one-year data set recently, but the work that I'm talking about is sort of a three-month data set from Google. The longer the data actually has the same data from Google, Microsoft and Yahoo! so that's -- that helps us tease out a few more things, eventually. So, I'm going to talk about the Google data today.

They have -- we've got about 1900 unique key word advertisements, and these are all the ads that this retailer sponsored. So, you know, we didn't have to worry about selection issues here.

Here the broad category is bedding, bath, kitchen, home decoration, dining and so on and so forth. It had then 40 unique product categories and about more than 200 unique events. So, we got data from both the paid search and organic search. So, we have number of searches, impressions, clicks, cost per click, the rank, conversions, revenues from the conversion, whether the ad was retail or brand, generic, and the last similar data set from the organic set. So, conversions, revenues, from the organic site, at the same time when the ad was being shown on the screen on the paid and the organic side.

So, here's basically the sense of the auto logistic model. I'm going to be just briefly showing you a slide that has the joint level of distributions. So, there are four kinds of possibilities here, right? A consumer may click only on the paid listings. A consumer may click only on the organic listings. A consumer may click on both paid and organic, or a consumer may not click at all. Right?

So, we are going to be using the auto logistic model and, in particular, Besag's theorem to formulate the joint distribution functions, and the idea is to tease out the nature of the interaction effect. Is it a complementary effect between paid and organic, is the interaction negative or is it independent?

Let me -- so, here's the joint level of distribution. So, the first equation, essentially the probability that a consumer clicks on both the paid and the organic. And the teed up parameter is interdependence parameter. So, teed up parameter maps whether paid and organic, if the simultaneous is positive, that suggests that paid and organic have a positive complementary relationship. If it's negative, that means they have a substitution effect.

Pi is intrinsic to the function. So, what we're saying is users, when clicking on a paid ad, they have a market share utility from clicking on an organic ad and they also have some market share utility of clicking on both the organic and paid. And if they don't click at all, that's the last possible option.

So, as I said, we look at consumers, first with the number of searches, the first three questions look at the consumers, the number of searches as a function of the ad attributes, whether the ads have retailer brand or length, then the second equation is about click-throughs and click-through rank. We are looking at rank, also, because your ad click-through will be based on rank. And then the third equation we are looking at conversion rates, also.

From the advertiser's point of view, we have -- we are trying to model how the cost per click and the bidding behavior varies, based on their rank in both the paid and the organic. So, this particular company that we are talking to, they use a relatively simple intrinsic in which based on the ranks in the previous period, on both the paid and organic side, and the profits from both the paid and organic side, they decide on the billing period in the current period. And then the search engine certainly it's important that we factor in both the current bid and the prior year click-through rate.

So, as Susan talked about, they factor in, you know, the rate of bids, they pay their bids based on prior click-through rates. So we look at lag click-through rates on that. We have some information on competitive price and that's one of the limitations of the data. It's not, you know, we know the average competitor price for this advertiser in that time period. We don't know the individual bid price of those competitors. So, it's a little noisy, but it's relatively decent proxy for figuring out what's going on.

So, these are the five simultaneous equations and we are solving this and the data is all correlated with the rate of normal distribution. So, that's sort of the framework of the model.

I only sort of show the clicks on conversion in the interest of time. So, you will see that rank of the ad does play an important effect, in both clicks and on the paid and the organic side. And this is consistent with prior work that has looked at privacy effects, Mike's work and Shabbat (phonetic) has shown similar reserves that your click-through rates keep going down.

And, so, the interesting thing here it seems like click-through rates matter, the rank matters relatively more for paid versus organic. So, you know, on the paid side, it does play an important role relative to organic, but that's the interesting parameter from our point of view, the interdependence parameter is very positive that you would see, so that's basically suggesting that paid and organic click-throughs have a positive complementary relationship. So, it does make sense for advertisers to show up on the paid side.

Well, I'm going to talk about some robustness. Now, one of the -- you know, in the basic model, we don't factor in independence of the teed up parameter. So, we rate the robustness. We actually extended the model incorporating both independence and interdependence in the interaction parameter. So, it's also a mixture model, the estimate and the point mass of 9.72 on the interdependence model would suggest that the interdependence model is actually the right one to go with. We did some out of sample relations and found out the proposed simultaneous regression model predicts a lot better than the same model estimated aggression/regression. Same with when we looked at a naive, very naive non-model-based forecasting approach, the current model does a lot better.

We did some policy simulations, and I think the interesting result over here is that they're getting a 4.3 to 6.5 increase in advertiser profits if -- given this positive interdependence between paid and organic searches.

So, that was interesting and various things happen because of the kind of key words. For certain key words, retailer key words, the positive effect is more. For comparative key words, the positive effect is less. And, so, that's sort of another interesting paper. Or interesting result.

Then as I said, in one of the limitations of the auto logistics model is you have to assume symmetric interdependence. So, we also model a discrete game entry framework where we're looking at possibility of say symmetric interdependence. In other words, you could argue that as a consumer, the ad showing up on the paid side might have a probability of me clicking on the organic side than vice versa.

So, we're looking at the fact that there's say symmetric interdependence. So, we find that the effect of having organic listings on the paid search is much stronger. On average, about three times stronger than vice versa.

So, here's a few experiments that I talked about. We worked with this particular company for about an eight-week period in which we asked them to stop sponsoring a set of key words for two weeks. We tracked their conversions, clicks and revenues for the organic side for two weeks. For the next two weeks we said, why don't you start resuming those sponsored ads and we're going to track your click-throughs, conversions, revenues from both paid and organic, and we repeated this over an eight-week period.

And, so, this kind of, you know, turning the ads on, turning the ads off, experiment helped us to solve the different effects of paid and organic on click-throughs, conversions and revenues a little better than the data we had. And, so, that's why you see that a combined conversion rate and the combined click-throughs, when the paid ads were on, were significantly higher. Compared to when the paid ads were turned off.

Now this is a smaller sample, you know, because the advertisers wouldn't let us take all their 1900 key words, this was only less than 100 keywords. But at least this releases the fact that having your ads show up on the paid search side is definitely a good thing for you. So, the combined effects and for both click-throughs and conversions is a lot better. And, also, for profit -- if you looked at revenues and profits, so that went up, too.

So, basically to conclude, you know, we have a hierarchical basin model, when we estimated this model to figure out how ads impact consumer search, click-throughs and purchases. We also examined dispositive interdependence, which suggests that, yes, even if you're showing up on this free organic site, it does make sense and does pay for you to show up on the paid side. It is asymmetric. So, showing up on the paid side, also showing you on the organic side has an asymmetrical relationship. There is a 4.3 to 6.5 percent increase in your profits, based on some of the counterfactuals and policy simulations we ran.

So, that's sort of a -- you know, in the field of experiments validated that, yes, your combined conversion and click-through rates and combined revenues are much higher when you have both paid and organic, compared to when you only have organic.

So, part of this, as I said, the last couple of years, part of our work has also extended to working with, you know, this -- these results have some indications on whether advertisers should invest more on search engine optimization, like improving their landing page qualities, versus improving their -- you know, having higher bids, maybe, on search engine options.

So, one of the -- some of our current work actually involves working with the advertisers in manipulating the landing pages and trying to see, again, running some field experiments with them and trying to see if they shift the content in a certain way, does that lead to higher conversions and so on and so forth.

The one question that I always get, and from advertisers, across, like we worked with, you know, financial services, travel, IT, retail, that if this is true, if having your ad on the paid side does always lead to higher probability of organic and vice versa, would search engines have an incentive to play around with the organic ranks. And I remember Susan and I talked briefly about this and she had some interesting insights to share. So, we are sort of trying to -- so, possible future work is trying to look at this by working with SEOs who have data from multiple advertisers, and the conversions and click-throughs.

So, that's sort of where we're headed to, to trying to decide if there is an intent for searches to play around with the organic rankings. They get paid on a per click basis, so you could argue that maybe there is some incentive there.

But that's basically what we have so far.

MR. SMITH: So, my name is Loren Smith, I work here at the FTC. I have to give the usual disclaimer that these are my views and not the views of any Commissioner or the Commission.

I thought this was a really neat paper, and it was very well done. It's very complex, the estimation technique is quite involved, and I must admit that without the help of Matt Weinberg and Wikipedia, I wouldn't have known what was going on. But, eventually, I kind of got a basic idea of what he was doing in the estimation.

And, so, the primary question, or the one important question in the paper is, do paid and unpaid search advertisements, how do they interact? Are they complements, are they substitutes, and his simulations indicate that they are complements, and they are supported by his -- some field experiment results.

He also finds that retailer-specific key words, which are less competitive and more specific to this particular retailer, have a larger interaction effect than do generic or brand-specific key words.

Things I liked about the paper, he empirically qualifies something, a complementarity between organic and paid search listings that, you know, really without doing this exercise, we wouldn't have known what the sign of it was. If you have a high level, a high rank in an organic search listing, is it worth paying for a paid search listing? And I think that's an important question, and I think he answered it well in the paper.

The estimation routine allows for an arbitrary correlation pattern across the errors of the model, because of the hierarchy of decisions made in the model, you don't know what the error correlation across the errors of those decisions might be, and his estimation technique allows for that correlation to be arbitrary.

The estimate seems sensible. The structure allows for him to run some counterfactuals that inform bid strategy and key word selection in paid search advertising, and the results predicted by the model are supported by a really cool field experiment. I mean, it's very rare in IO that you have the opportunity to compare your results to what might actually happen in the real world, and he has that opportunity, and he took advantage of it. And we all wish that we had that opportunity, and I think it reflects well on his simulation results that it's at least indirectly supported by what he sees in the field.

So, the highlights of the model, it's a very detailed model, demand, consumer click-throughs and conversion rates, with some other equations modeled within there. On the supply side, he has cost per click. And then the estimation, he has the observed number. So, the data -- he has the observed number of each possible search outcome, and he estimates a likelihood function of parameters and consumer market share utility function from clicks, the observed number of purchases, the likelihood function of the conversion propensity, and then a cost per click, he parameterizes the linear cost progression equation, and within that, he estimates an equation for the rank of key word search using a similar set of covariate.

The very complex estimation routine, you draw a set of parameters from a proposal distribution, you accept the draw, if it meets some criteria, that depends on the proposal -- the relationship between the proposal distribution and a target distribution, the likelihood, and then you do this for a while, and you figure that you're getting close to the proper distribution, you throw all of those initial simulation or iterations through the algorithm away, and then you use the last end iterations of that algorithm and the M accepted draws and treat that as posterior distribution that you can draw the mean and the standard error of your parameters from.

This sample -- that's what it says in the last bullet there.

Applications, I always think about what could I do with such techniques. So, other advertising, we might want to know if they're complements or substitutes. Direct consumer advertising, detailing in drug markets, this is a question we might be interested in here at the FTC. Are online stores and traditional stores substitutes or complements for one another? That's beyond the scope of advertising.

Estimation, anywhere that we want to build a model of consumer decisions where we're uncertain about the correlation in the error structure and we're uncomfortable establishing a nest. We might want to use a method like this where you estimate a set of simultaneous equations, which allows for an arbitrary correlation in the error structure.

Some questions and comments that I have. One concern that -- the major concern that I have about what he's doing here is that neither the demand for clicks or the conversion depend on the characteristics of the actual product that ends up being purchased. So, for example, price. Is this information available? Could it be used as a covariate? I think that it's likely to be correlated with both your other explanatory variables and your errors, so it could cause some problems in your estimate.

The model fit, you talked just briefly about that managers don't appear to be behaving optimally. I would like to know more about is that behavior systematic. Are they doing something that's related to actually the complementarity, are they missing the complementarity or are they overestimating the complementarity in their decisions?

Can you pair the results of the field experiment that you see directly to what your model would have predicted in that situation?

And then another counterfactual, you might have to actually re-estimate the model with some interactions for rank, but how does the complementarity between organic and paid advertising vary with the rank order of where the ad shows up in the search listing? Something like that might be interesting to see as well.

And that is all I have. Thanks.

MS. ATHEY: Any questions?

AUDIENCE MEMBER: (Off microphone) relating to Loren's question, it's not clear to me that you have a control for the page where the organic listing appears. Is it just the list is also if it's on page 85 of an organic listing, then it just has a really high rank or is there an explicit control for that?

MR. GHOSE: No, that's right, we know the rank of the organic listing. So, if it appears on the fifth page and ranked sixth, it will show up as rank 36, or rank 56. So, it shows up as well.

AUDIENCE MEMBER: (Off microphone) there's not a dummy, though, of whether it's on the first page or not?

MR. GHOSE: No.

AUDIENCE MEMBER: So, not surprisingly, I was more excited about the field results than the observation results. Can you give me any intuition at all about where identification comes from in your observational study? Is it -- is it basically comparing one key word to another that has different ranks in the different positions and that's how you're doing it? Because I just --

MR. GHOSE: In the experiment you mean?

AUDIENCE MEMBER: No, no -- well, in the observation Al study. So, it's a very complicated system of equations, and you have a higher arch Cal basis.

MR. GHOSE: Right.

AUDIENCE MEMBER: You basically have five equations with multinormal error term, and I just was having a hard time imagining what is varying in rank that's allowing you to identify the change in clicks due to the rank?

MR. GHOSE: Right. Sure. So, it's a fully recursive triangular system. So, this is, you know, this is based on some -- an identification strategy proposed by Lahiri and Schmidt many years back where what they are saying is if you have a triangular system and if it's fully recursive, then you don't -- you can identify without having any further restrictions on the dynamics and on the area and so on.

So, in other words, you start with one variation on the cost per click, which is the advertiser's decision, they base it on the previous period rank and the previous period profit. Then you look at the search engine decision, and the search engine then decides the rank based on the current bid and the prior click-through and they look at the consumer's decision on click-throughs and conversions.

So, what is happening is that we have for each of these cursive iterations, there are certain variables missing from the previous one, the previous egression, which is not there in the next egression and so on.

AUDIENCE MEMBER: (Off microphone) so, you're actually identifying the over time rate for the particular key word?

MR. GHOSE: Yes, identifying over time for a particular key word.

AUDIENCE MEMBER: So the rank changes for a particular key word?

MR. GHOSE: That's right, yeah.

MS. ATHEY: I would just kind of reiterate that question that there's the molding and sometimes it's so complicated that neither your presentation nor the discussion actually was able to articulate that particular issue because the cross-sectional variation could be a little bit problematic, just because if you're searching for -- if somebody is searching for bed sheets at Kmart and then your landing page is going to be very relevant for that. So, you are likely to be high in the organic listings and you're more likely to get a click-through on a Kmart ad. That's a -- that's a correlation and not causality. So, the -- clarifying that issue would seem to me to be the fundamental economic issue.

MR. GHOSE: Right.

MS. ATHEY: From interpreting your results.

MR. GHOSE: Right. So, we also did some simulations, you know, like in order to make sure that the parameters identified, we tried to back out of those from a similar data set, if you could back out the same parameter estimates. So, we did a little bit of that, but you're basically relying on the fact that, you know, the key word rank for a given key word varies across time for that key word.

MS. ATHEY: That's right. That's a more sort of compelling sort of variation. So, the comment I actually was going to make about the -- when you go to the field experiment, you have cleaner identification, but then you never get to see what happens when you take the organic link away.

MR. GHOSE: Yeah. That's -- yeah. Maybe that is up to the search engines to help us out a little bit.

MS. ATHEY: Exactly. So, a future collaborative project. And then the incentive for the search engine in the end, the claim that has been made was that the search engines don't want to put up -- they don't put up paid links on the left side very much. And, so, there's a claim that maybe they've -- they're trying to extract more revenue. You've found that the -- that being high on the organic will increase the click-throughs for a particular advertiser, but what you haven't been able to show is that having those firms on the organic side won't cannibalize clicks away from the ads as a whole, because putting an ad high on the organic side could shift clicks from one advertiser to another.

MR. GHOSE: Right.

MS. ATHEY: And simultaneously divert clicks away from the paid side all together. So, the search engine's incentive still is not clear.

MR. GHOSE: Right.

MS. ATHEY: One more question. Two.

AUDIENCE MEMBER: I'm going to bring this back to the antitrust world a little bit, rather than the identification world. One of the issues in the Google DoubleClick investigation, a key issue I think was whether or not search advertising and display advertising competed with each other. And your results would suggest to me that search advertising really is potentially quite differentiated from display advertising, especially this synergy between the organic and the paid. I wonder if that is a correct inference in your view or not.

MR. GHOSE: I mean, I haven't worked on something myself, but I remember actually someone from Yahoo! I spoke to someone some time back, and some folks in Yahoo! had looked at this possible synergy between display and search and they did find that there is, again, a positive synergy between display and search. So, yeah, that's what we know so far.

AUDIENCE MEMBER: Is it as big? Is it this four to six? Is it as big?

MR. GHOSE: I don't remember. I couldn't get that information from him, but he did mention that.

AUDIENCE MEMBER: (Off microphone) they've done several studies now that showed, you know, if you run display ad campaigns that your number of searches for your -- (inaudible).

MR. GHOSE: And the only other data point I have is from a company called I-crossing. I-crossing is the largest digital ad company in the U.S. and they also work with companies to look at these kind of synergies, and they also corroborated that they found something similar. So --

AUDIENCE MEMBER: (Off microphone) I apologize, since I haven't read the paper, so I'm maybe asking two very simple questions, but I was wondering, first of all, if you had separately looked at placement of paid search advertising in the top versus the right-hand side of the page? Meaning that I might imagine that actually if you got your ad on top, it's more of a substitute than a complement to organic. So, I was wondering if there's a differential there.

And my second question is that although it seems like you have a rank variable, it would seem intuitively to me like what would matter would be the relative position of the organic and the paid search ad. So, if the organic listing was very high, and the paid search ad was very slow, I might be more likely to click on the organic one, and if the results were opposite, I might be more likely to click on the paid one.

MR. GHOSE: No, good question. So, the first one we have looked at, so we have actually looked at the -- you know, it's a small sample, but if you sort of look at the only sub-sample where the ads on the page were on the top three slots, and then looked at what happened, we then find the positive complementary effect.

The second one we have, and this is going back to Mike's point where he said that you might want to add a dummy if the organic listing was on the six page. Even though we had the rank, but you might argue that the consumers can think of the 64 listing would be the sixth page and the fourth listing. So, you might have --

AUDIENCE MEMBER: It's linear when you get to the second page. So, on the paid search, it's a paid ad --

MR. GHOSE: Yeah. So, in another study, we actually looked at how conversion rates fall with the square of the rank term. So, you're right, so we actually find a nonlinear effect.

AUDIENCE MEMBER: (Off microphone)

MR. GHOSE: Well, from what we know, about 80 percent or so stick to the first page, and then the second page about ten more and then just keep going. It's a long, very long table.

AUDIENCE MEMBER: And that would then complement or substitute, right, because you're substituting the fact that the paid ad from the first page were back with the organic ad on page 85, right? So, I think the tease -- I mean, the experiment provides for perhaps in terms of the econometric model that the dummy would allow for that.

MR. GHOSE: Right.

MS. ATHEY: Great. So, let's move on to Gunter and talk about markets with indirect network effects.

MR. HITSCH: I hope so. Thank you. All right, well, first, thank you for the opportunity to present my research here. This is joint with J-P Dube and Pradeep Chintagunta at Chicago GSP, which as of today is known as the Booth School of Business, $300 million and no word yet on my raise next year.

This paper is about markets such as BlueRay versus HDTV, standard war of markets in indirect network effects now decided in favor of BlueRay. Economic theory that says markets with indirect market effects tend to become concentrated, and the goal of this paper is first to clarify how you could measure the market concentration in use by indirect network effects and then provide empirical or maybe I should call it semi empirical illustration for a specific case of a standards for the first generation of the Sony PlayStation versus Nintendo 64, about 12, 13 years ago.

So, very, very briefly, this might be too obvious for the audience, but what gives rise to indirect network effects? So, think about consumer adoption, consumer adoption for, say, a video game console depends on the hardware and the price of the hardware, it also depends, of course, on the software, which is a complementary good, in particular, quality and variety and price of software.

So, assume, which is, I think, true in many markets, that there are economies of scale, meaning if a larger number of people adopt the standard, there can be more software forthcoming. And that gives rise to indirect network effects, because then the adoption decision of consumers indirectly depends on the size of the network. Indirectly because consumers care about the software, not, per se, about how many other people have it up.

So, to illustrate this, just to lay the

groundwork here, think about two competing standards, is an empirical application, the crucial variable I want to focus on is the install base, which is the cumulative adoption for either of these standards.

So, we can visualize this adoption as a point here in this triangle, this triangle is distinct space of potential adoption patterns, and we can visualize the evolution of a standards war, it's just a sequence of points. What's the relative fraction of adopters for each standard in each period?

So, to talk about concentration, just introduce the main concentration measure we focus on, it's a cumulative one firm concentration ratio. Concentration ratio not defined in terms of current market shares, but in terms of cumulative adoption rates.

So, why can indirect network effects lead to markets to become concentrated? Think of some initial advantage for standard one, initial advantage in terms of adoption means it's more self forthcoming for the standard that reinforces the advantage and the market can become concentrated, quote unquote, "tip in favor of standard one," same thing could happen if just for some reason standard two adopted an initial advantage.

And as the literature -- by the way, I realize that I don't have any slides on the literature, all I'm telling you right here, on this slide, you could get out of a Katz and Shapiro, their seminal 9285 paper, which is not a dynamic model, but the information can be provided and stagnant.

So, here in this example, there's no initial advantage, but if consumers have expectations, nonetheless the market might tip in favor of one of the standards. Why? Well, suppose consumers expect that standard one will win; therefore, in a certain equilibrium, standard one will win, consumers will believe that standard one will win, will adopt standard one, and then more software will be forthcoming for standard one and expectations are self-fulfilling. And I could construct under certain parameter values and equilibrium that the market is going to tip in favor of standard two.

So, to sum this up, why can markets with indirect network effects become concentrated? First, there's the positive feedback effect. You have an initial advantage and just initial advantage tends to propagate, and this process can be exacerbated by self-fulfilling expectations, and then you can actually have multiple equilibrium. And all these mechanisms together, that's what the literature on indirect network and indirect network effects refers to as tipping.

Okay?

There's no -- the one I have up on the slide is, I think, the most concise definition I could find even -- and actually, I realized in Farrell and Klemperer's survey chapter in the current -- in the latest handbook of industrial organization, there's actually no concise definition of tipping. But I think all these things together, that's what it's really referred to as tipping in the literature.

So, now the main point of the paper is, well, how can you measure tipping? I already introduced a concentration measure, one firm concentration ratio. Now, think about what does this concentration ratio depend on? Well, it depends on all the model parameters that define demand and cost, and it depends on a certain equilibrium that's being played out.

So, if I know these parameters and if I know the equilibrium, I can, in principle at least, calculate the expected one from concentration ratio, say 25 periods after the initial launch of the -- of both standards.

So, now in the even measure of tipping that follows, some are calculated expected concentration ratio, and I compared it to 50 percent. You have two standards and they're completely symmetric, this might make sense. Any deviation from 50 percent might tell us, don't think about how indirect market effects lead to concentration.

The obvious problem with that is in empirical, general markets are not symmetric, there's demand and cost differences. So, really, what we would like to do is to make counterfactual predictions about markets if the parameters that in using our network effects were removed or were made smaller in size.

So, ideally, I would like to study a model variation where I change some of the parameters that lead to indirect network effects, then I find in corresponding equilibrium, and I compare the market concentration, expected market concentration under our actual market versus the counterfactual hypothetical market where indirect network effects were removed.

So, how do I come -- how do I calculate this measure here, this proposed measure of tipping?

So, ideally, ideally, I would study a couple of hundred independent markets, which are identical, and I have identical and initial conditions, then ideally I would like to experimentally manipulate whatever causes indirect network effects. And while everyone understands in this particular case, our standard wars, that's very, very hard to do. All right?

So, what we do in this paper is we've built a model, a model of standards competition under indirect network effects, and we try to calibrate this model from demand estimates and cost side data, and then use the model for an equilibrium and predict the market evolution, okay?

So, essentially in our computer we run this experiment where we study counterfactual markets.

This is the model in one slide, I believe that it wouldn't make any sense to bring up any equations here, given my time constraints. So, the model has three sides, and two are the really interesting ones. First, consumers sold some dynamic, durable good at option decision. The problem here, they choose between adopting one of -- well, in our application two standards, and they can delay the option until tomorrow, and whether they adopt or delay depends on their -- well, depends on current prices and software availability, but also on expectations about future software ability on prices.

Then we have another party, the hardware firms, who price their products dynamically taking into account how their current price affects current adoption with all of the future of the market. And we have software firms, and the main part is that software supply is increasing, and the cumulative installed base decides to network.

This model is close to using an equilibrium concept. Let me now explain why this is a phase in equilibrium, while there's some private information in the model. So, what does this equilibrium capture? It captures some interaction. It captures that consumers make adoption decisions thinking about how other consumers will adopt the standard. They need to know that because it tells them how much software they can buy in future. And consumer adoption decisions also depends on the expectations on how hardware firms will behave that will tell them at what price they're going to buy a product in the future, and firm decisions depend on the expectation of how consumers and their competitor will behave in future.

In our model here, there's no strategic role for software firms, okay? So, essentially all we do is with a reduced firm software side, where we estimate to what extent more software is forthcoming, if the cumulative adoption for a given standard is higher. And there is -- you know, an important, implicit assumption here, which is that you don't have any superstar games like Halo in the market, okay? That's something we assume away, and I think it's okay for our generation of video game consoles that we study, but probably certainly wouldn't be okay for other generations.

So, where do we get our parameters from? As I already said, we focus on this standard war between the first generation, the first Sony PlayStation versus Nintendo 64, we estimate demand based on sales price and software data, at the monthly level.

I have no time to talk about the exact details here. By the way, the empirical side of our paper is not really contribution. It mostly follows the lines of existing work on durable good adoption. Where do we get cost-side data from? We get cost parameters from industry records. So, that's the approach, that's the -- that's our data.

Now, as I said before, a couple of minutes ago, the goal is to calculate this measure of the increase in market concentration due to indirect network effects. And we do that by first getting our parameters, estimating them and getting them from industry records, and it's sold for equilibrium, and we simulate this equilibrium where initially nobody has adopted any of the standards and then simulate the model 5,000 times and record the adoption rates 25 months after the beginning of the adoption process.

And when we calculate counterfactuals, we manipulate two parameters: First, the market share utility of software, and secondly, the consumers discount factor. Well, the first part is, I think, obvious why do we manipulate the consumer discount factor. Well, it manipulates the importance that consumers give due to their expectations about the future evolution of the market, the future software availability. Which I think is an important part of -- we know is an important part of indirect network effects and the effects of indirect network effects on market concentration. So, I think it's easiest to give to you an idea of the flavor of our results, but first -- four? Five? Four. I'll further negotiate.

Let me show you outcomes with symmetric competitors, because it's easier to understand what's going on. So, I take the parameters for Sony, and assume there's two identical Sonys competing against each other. Okay? So, hopefully, this works when I move around with this cursor here. What do we see here? This is the state space. These arrows show you how in expectation the adoption rates move between periods.

This here is the 45 degree line, essentially this says, if the -- if the current adoption rates are symmetric, they're expected to stay symmetric. But you see here from the direction of these arrows that, you know, we see these positive feedback effects. If one of these guys gets an advantage, his advantage tends to propagate.

Nonetheless, across 5,000 simulations, here, this is the distribution of the share of the installed base after 25 months, there is exactly identical outcomes.

Now, these two graphs here are for our estimate of software where the market share utility is scaled down by 75 percent. So, it's 25 percent of the estimated value. Now, here you have the outcomes for 100 percent of the estimated values. Well, estimated values.

Observation number one, we are unable to sell for symmetric equilibrium, even if I start the symmetric equilibrium, it converts away. How is that possible? It's possible because the computer you have round-up error. So, that's how we saw the first time that, well, there are some strong indications for asymmetric equilibrium here. There's more than one equilibrium. I'm showing you a particular one. Equilibrium tends to favor standard one.

And here, this is distribution of installed share, installed base shares, after 25 months, typically a standard of one gains a very large share of the market in more than 25 percent of simulations, more than above about 90 percent of the market.

Sometimes, however, standard two gets a very large share of the market. How is that possible? It's due to random demand shocks, which are in our model -- I'm sorry, the estimated standard deviation of these from the data. Which might randomly move the state here under this part of the state space and then consumer expectations essentially flip. And due to the change in consumer expectations, standard two wins against standard one, so to speak.

Let me skip that. Pricing patterns, these are the penetration pricing patterns here. The same pattern if you look at the consumer discount factor of 0.8 versus 0.9. There's different ways of doing this comparative static, which I might be able to clarify later, but I'll skip over.

Now, we do the same thing with the -- with the asymmetric competitors, the actual estimates for both Sony and Nintendo. We see that if you scale down the software market share utility, Nintendo typically gets a higher market share. Why? It's because of -- mostly because of the cost advantage that Nintendo has over Sony. However, once we turn up the software, the market share utility of software to 100 percent of its estimated value, we see that Sony typically wins. Why is that? It's because typically in our data, Sony gets more -- there's a larger supply of software targets for Sony at any given state. Why is that? It's because Sony made it cheaper for software developers to develop games.

Similar pattern if I move around to consumers discount effect, okay? So, I guess I'll be very, very brief here, but I have two more minutes. All right.

So, this is our -- this is the promised measure of concentration, where I -- where we compare concentration under the estimated parameter values versus a counterfactual model, it's a couple of counterfactuals down here, especially along the lines I'm showing you. It seems that in our market, at least for the version of our market that we calibrate and simulate, indirect network effects lead to an increase in concentration, more than 23 percentage points, which is a very, very large economic significance.

So, summary, the main goal of this paper is to clarify and explain and show how you can measure tipping. Now, let me relate this a little bit to the literature. I think that the most closely related paper is a paper by Jenkins, Leo, Matzkin and McFadden, they do something very, very similar for the case of browser war, Internet Explorer versus Netscape. The main difference here is that we actually incorporate forward-looking consumers into the model and I think our results show that this is something very, very important. Quantitatively important.

Results show potential large increase in concentration. Our results also show that what we predict is very, very sensitive to a couple of things. In particular, the market share value of software, but more what concerns us more is the consumers discount factor. Why does that concern us? It's because that it is virtually impossible to estimate consumer discount factors if you just -- if you have consumer adoption data in the way we have used this data in the recent literature on durable goods estimation.

So, well the other thing that I think our results are sensitive to is, of course, the -- I think that it's something you want to discuss here, the assumption of rationality that is enshrined in this equilibrium concept that we have. So, here it is consumers taking the results of this concentration data, all right?

So, where we are moving forward is on the consumers discount factor. So, we are currently in the process of designing conjoined experiments. That conjoined analysis is a standard survey-based marketing research technique, and we try to learn more about discount factors and patience from these conjoined designs. And I think we're going to go back and say, well, how would that affect some of the predictions we make here for such a market of indirect network effects? Thanks.

MS. ATHEY: Thank you. Robin?

MR. LEE: The last, I guess, discussion, I'll try to be brief. First, thanks for your time and the colleagues formerly known as the Chicago GSB. Before I begin, I thought it would be nice to talk about tipping and network effects in general. I found this useful when talking about these type of industries.

But most of the early literature looked at these types of industries, focusing on standards battles or more or less one-sided networks, where agents just joined this platform standard network if you will, and this platform is nonstrategic, it has no stand-alone value. Consumers just join something if they think other consumers will also join it. So, it's really like a location choice or a bazaar or meeting place. And again, the value is strictly increasing the number of adopters.

And in these types of models, it's not surprising that we're going to see tipping. We should see complete market tipping toward one standard or another, and the only times we don't are typically as a result of coordination failure.

The issue is that in many real-world network settings, it isn't so simple, these networks are multisided, that is there may be consumers and firms using a joint applied forum. These platforms may be strategic and horizontally differentiated. These platforms may engage in different pricing strategies, different consumers or firms. Consumers or firms can join multiple platforms or multiple standards, and there may be even same-sided congestion effect. So, in auction markets, buyers may prefer auctions with fewer buyers, sellers may prefer auctions with fewer sellers.

So, in this regards, markets need not completely tip, even though there are strong indirect network effects. Even though I care there are a lot of games, because of these other factors, we can still see market splitting equilibrium. And this raises the question, do network effects still matter?

And I think this is one of the great strengths of the paper is it gives us a real way of measuring the impact of network effects by defining an appropriate counterfactual to compare the difference between industry with network effects, and industry without network effects. And notice that even in symmetric market, because of demand shocks, different prices, different marginal costs, we can actually get very different outcomes than complete 50/50 market shares. And, so, I think that's nice.

But my interpretation of what we can do with this may be a little bit different. As opposed to asking how much closer do we get to complete market tipping as a result of network effects, we can also ask the parallel question, how far away do we get from complete network tipping because of these other factors? Because of the fact that maybe consumers can multihome.

The counterfactual raises an interesting thought question, too, it's more food for thought. What does it mean to reduce network effects? I mean, because just to some extent the fact that consumers really care about the number of firms or there's this dynamic feedback loop, these two aspects are really fundamental to the nature of these two industries and by reducing them or removing them, what are we now looking at? Are we looking at something fundamentally different now?

So, maybe we can hold fixed network effects and maybe add the ability to multihome or, you know, change congestion effects or add strategic platforms. And see how that shifts where we go.

But let me move to the application, and the question is, you know, are video games tipping? Should we think about them as standards? In my opinion, maybe not so much, because these problems are indeed horizontally differentiated. Consumers may prefer one or the other, for some exogenous reason. Software, also may do so, and they may also adopt multiple platforms. You know, some people really like video games so they buy everything available.

So, because adoption decisions are driven by more than software availability, we again need a relevant counterfactual, which is great, the paper stresses that we need to define a relative counterfactual to compare against when we measure network effects, but it might be a weakness insofar as if you don't capture all these dimensions, we might be able to, say, overestimating the impact of these effects.

And to this step, you know, I add some suggestions for modeling points. So, right now there's no heterogeneity multihoming. I read in the previous slides they're working on incorporating that, which is great. Right now, consumers only care about the number of software products. And I understand it's very difficult for them to care about the individual identity of software products, but maybe we can make some progress.

For example, there's this coefficient called gamma in their paper which refers to how much consumers care about software. And let's say now we allow it to be software-specific. That gamma on Nintendo is different than gamma in Sony. This means that maybe Nintendo games on average can be higher quality because they employed a quality versus quantity approach, whereas Sony employed a quality versus quantity approach. And I think you can identify this because you have a -- the panel metric nature of the data.

We can allow consumers to care about other things. Maybe they care about the number of consumers on board. But these are all easy to incorporate in the model as it's specified. The demand shocks are size or ID right now, because it allowed the variance in the market shares driven by these demand shocks. Maybe we can try persistent demand shocks, we try, you know, to test the robustness of the predictions.

And, finally, my understanding, I mean, this is a nice, clean application of a two-step estimator. It's really nice and it shows how we can estimate these complicated models, but still make it, you know, complicationally feasible. But the issue is that the first stage policy estimates are they still consistent in this counterfactual because if we reduce network effects, aren't we now changing something fundamental about the industry and maybe we should be a little bit more careful about thinking how we should interpret that.

But, in conclusion, it does provide us a nice framework to measure the importance of network effects, and insofar as it allows us to get away from the idea that just because there are indirect network effects, we should expect complete market tipping, I think that it makes a great point. It contributes to the literature on dynamic demand in pricing. And, interestingly, it actually endogenizes penetration pricing, which I thought very nice.

Gunter didn't have time to discuss it, but in the paper he discusses that the model predicts that these platforms actually priced below marginal cost early on, which is something we observed in industry. And if we can extend this and maybe allow for, let's say, estimating what marginal costs must have been to rationalize the observed price path, would be very nice. Or maybe we can endogenize the royalties that platforms charge to software providers.

So, can we allow the platform now to charge both sides and maybe test the two-sided market literature and the theory indicated that literature and see if they're actually behaving optimally, given what your cross-priced elasticities would predict.

And, finally, it allows us to maybe consider really what we'll talk about when they say network effects could be a potential barrier to entry. So if you look at an entrant, how much would an entrant have to invest in vertical or horizontal platform differentiation or investing in new software to actually gain some advantage against the incumbent and is this a result of network effects or is this a result of the other random things that go on?

But I very much enjoyed the paper, I liked it, and thank you.

MS. ATHEY: Questions? Very good. So -- (inaudible).

AUDIENCE MEMBER: (Off microphone) (Inaudible)

(Applause.)

MR. ADAMS: So, lastly we're going to have John List from University of Chicago talking. I ran an experiment, a third experiment where I put John on the experimental committee to see if we got more experimental papers. We seem to be concerned about identification. So, you're now invited next to see what happens. So, please welcome John List.

(Paper Session Five concluded.)

KEYNOTE ADDRESS BY JOHN LIST

MR. LIST: So, thanks a lot for that introduction. While I'm getting my slides up, I want to thank Chris for his hard work on this conference and also thank the FTC and Northwestern for the financial support for this conference. I think it's been quite nice and I've learned a lot. So, thanks a lot, Chris, and thanks a lot to the other organizers.

So, today, I'm going to have to go fast because I have a flight at 2:30 and I see a lot of you are jumping around. So, if you have to get up and leave, that's fine. I don't have a problem with that.

For the next half-hour I'm going to talk, first, about lab and field experiments, and secondly, I will go through an example of a field experiment that I've been working on for the better part of three years now and it's going to an actual marketplace and looking at different aspects associated with collusion.

But let me start by showing you a figure that Charlie Holt has put together, and this shows the last 50 years of -- '48 to '98, of the number of journal publications in our area of experimental economics that uses lab experiments. So, I think it's fair by any measure -- and this is an important measure here, number of actual publications, you see this significant growth over time, and in particular, you have this big jump up from ‘93 or so to 2000. And since 2000, we’ve still had substantial growth, as well as in running lab experiments.

Now, that’s not to say that lab experiments are criticism proof. Now, when I was working at the CEA, I argued that we should be taking account of the willingness to accept willingness to pay disparities that we have found in the lab when we’re revising our benefit cost guidelines.

A White House official commented, even though these results appear prevalent, they are drawn by methods similar to scientific numerology because students are not real people.

(Laughter.)

MR. LIST: This is exactly the criticism that you get when you present results from the lab.

Now, the next line of skepticism has been brought up in various areas and I think Cross summarizes it well in his 1980 book chapter. It seems to be extraordinarily optimistic to assume that behavior in an artificially constructed market game would provide direct insight into actual market behavior. Now, what Cross is talking about is the early work of Vernon Smith on markets.

So, that type of very vague statement makes you think about, well, what is different between the lab and the field? We can think about selection rules in the markets. We can think about the commodity, we can think about the scrutiny. There are actually a lot of differences between the lab and the field.

Now, Smith responds to Cross in basically saying, is there empirical evidence to support these criticisms? At the time, there was no empirical evidence. There still is very little empirical evidence. So, Vernon basically says, if not, then the criticism is pure speculation. So, that sort of reasoning induced Glenn Harrison and I to think about different classifications of field experiments. And that's where this paper in the JEL essentially came from.

So, we all know a lot about conventional lab experiments. Cary gave us a very good example of a conventional lab experiment yesterday. I'm thinking about a lab experiment in the sense that you bring students into some sort of experimental lab and you give them rules and real pay-outs and then you look at their behaviors.

What Harrison and I call an artifactual field experiment is essentially meant to go after the criticism from the White House official. So, if the only problem is that we don't have real people in the lab, that's not a real problem for lab experiments, because all you need to do is go out and get the real market players that you're interested in. That's what we call an artifactual or synthetic field experiment because it's not really going out to the field, but it's making an important step in looking at the population itself.

Now, to start thinking about going after criticisms of Cross, you need to start thinking about adding naturalness to the environment, maybe adding naturalness in the task, the commodity, the stakes, the time frame, et cetera. And there are a lot of different types of field experiments under this specific classification that we call framed field experiments.

But the important part of a framed field experiment is that people still know that they're taking part in an experiment. That might matter sometimes. It might not matter in other cases.

Now, the final frontier, so to speak, is what we've seen a lot of today and some of yesterday. David gave us a good example. Dean Karlan gave an example of this. So, it's what we call a natural field experiment where you're in charge of the randomization yourself and it's occurring in the real market. So, you have realism and you have randomization. So, there should not be a criticism of this is not real and it’s hard to criticize the identification strategy because the only assumption you need here is proper randomization.

So, the underlying idea is to think about this idea that there are a lot of ways to generate your own data, the lab and a lot of different types of field experiments, and you can also use naturally-occurring data. So, we should think about taking advantage of all of these particular areas, not only field experiments, but also lab and using naturally-occurring data.

Now, much of my work has gone after what I would consider important economic phenomena in small-scale markets. That’s not because I have some affinity for small-scale markets. It’s that it’s not possible to do large-scale experiments in larger, more important markets. So, the idea is to go to the small-scale markets, manipulate them, test economic theory or provide policy advice in that market, and then think about what are the important features of other markets that we want to generalize these results to.

So, now, I want to go to my example. What is my small-scale market in this particular example? It’s an open-air market. All of us have probably frequented these open-air markets at some time in our lives. You walk in, you negotiate bilaterally for the good or service in question.

So, now, the rest of my talk, I will be giving various examples of field experiments from various open-air markets in a region that -- I can't tell you where exactly the open-air markets are, but you can walk a few miles and probably see some of them.

So, let's think about two facts from open-air markets. One is that we know very little about the economics of open-air markets. I think that's because we've never really taken seriously the data-generating process and going in to open-air markets themselves and manipulating them. We see open-air market data, but we tend not to believe it because there are a lot of reasons to mis-state what's happening in open-air markets. So, that's fact number one.

Fact number two, which I'm going to focus on for the rest of the talk, is that there are some very basic questions in the collusion literature, such as are large coalitions more fragile than small coalitions, that are very difficult to address empirically with field data. So, what I'm going to argue is that we can make advance on both one and two if we take the data-generating process into our own hands.

So, what's the strategy here? Through various interactions and open-air markets, I learned that there were certain collusive arrangements that existed. And I'll talk a little bit about those.

So, what I'm going to do is I'm going to run lab experiments to begin my analysis and I'm going to make sure these lab experiments are very similar to the experiments that experimentalists have used to test models of collusion. Then I will slowly move form the lab to an artifactual field experiment to a framed field experiment and then look at results from a natural field experiment. And in this way, this is what I'm talking about, what I'm saying that there's a bridge then between the lab and the naturally-occurring market.

Now, it would be important to recognize that in the natural field experiments, there are things that have arisen endogenously that will not be able to randomize, such as how many collusive arrangements are you in or how large is your coalition? That will then induce me to go back and run frame field experiments whereby I can randomize group size and group composition.

So, in following this strategy, a few things that I'm arguing I can learn about in this paper, the actual economic underpinnings of open-air markets. We know very little about that question. I'm exploring bilateral negotiations with or without seller communication, provide some insights on a few comparative statics of interest and then compare behavior across the lab and the field.

So, here are the details. So, I originally went into this market to think about when we want to test economic theory or do prices and quantities approach the intersection of supply and demand, if we allow sellers to communicate, does that frustrate that equilibrium outcome? And what I began to realize is that in those experiments, some of the sellers were acting in a very curious way.

And then I stepped back and thought about, well, there are small numbers of sellers that are providing homogeneous goods, the ones that were acting in a curious way. And when I'm talking about homogeneous goods, I'm primarily talking about selling of DVDs and CDs in this case. There are certain barriers to entry and seller communication is continual.

So, it seems like these are some of the conditions that would make it, perhaps, transparent that there's some sort of collusion going on. So, I'm actually a seller at these markets and I'm talking to who is now my mole next to me, and after a sufficient amount of badgering, my mole informed me of various collusive arrangements that existed in this market.

Here's how much of them are set up. We will be selling this particular CD. The marginal cost is $7, so let's all agree not to price lower than $14. All of the collusive arrangements were based on mark-ups from marginal cost, not on the features of the demand curve.

I learned of 27 distinct sellers across eight different markets being part of some type of explicit collusive agreement. These are in groups of two to four and across goods. Some of these sellers have multiple collusive arrangements across markets.

So, what am I going to do? I'm going to have my confederate approach various sellers within the collusive rings and outside of the collusive rings and I will negotiate to buy these DVDs and CDs one by one, and then I will explore whether any of these other features are correlated with how often people cheat or what sorts of pricing deals will they give my confederates.

Now, some of these sellers will also be in other lab experiments. Some of them will also be in some of the framed field experiments. So, that will give me some leverage to compare behavior across these various domains.

So, let's talk about what I find. The two-person arrangements have less cheating than four-person groups. People cheat less when they have collusive arrangements with a partner in more than one market.

People cheat more on high-volume, busy days. But you have to step back and say, is this because of treatment of selection? It could be the case that trustworthy people just have more collusive arrangements just because they're more trustworthy, and that's leading to this effect of people cheat less when they have collusive arrangements with a partner in multiple markets.

So, that then induces me to step back and say, let's do a series of framed field experiments, and I apologize, I won't have time today to talk a lot about the details of the framed field experiments, but you can get that from the papers. But what I will do is I will randomize the ground size, I will randomize group composition, I will vary cheating profits, I will vary the time frame. Most lab experiments tend to be 30 minutes or 60 minutes or an hour and a half. And an interesting question is, if we want to take that short-run elasticity and go up to a week or months or years, does behavior stay the same? That's an open empirical question.

So, all in all, I will have some 19 treatments. These are acronyms that you won't know right now until you look at the paper. But I wanted to give you a sense of this. I have students and I have flea marketers in the lab and artifactual field experiments. Then I have a series of framed field experiments where I'm randomizing some important features from economic theory or in this market on people, and then I have the natural field experiment that I've just summarized.

And then what I'm going to argue is that it's important to have a draw from each of these, lab, framed and natural field experiments, each of these classifications. Because, together, we can learn a lot more from these than we could with any one in isolation.

So, summary comparative stats, this is what I just mentioned about two-person versus four-person. So, framed field treatments, of course, can help, and here's the results for the framed field experiment. What I have on the Y axis is a proportion that cheat and on the X axis, I just have just consider table two and table four, table two is two sellers, table four is four sellers. So, what you have here is cheating rates of about 16 percent in the two-seller treatments and about double that in the four-seller treatments. So, cheating rates roughly double when you go from four sellers to two.

What about the idea that people cheat less when they have multiple collusive arrangements with a partner? Now, the framed results add some power here because what you find is that the number of outside agreements is not correlated with cheating rates in the framed field experiments. And, also, cheating rates are much higher when groups do not have collusive ties outside the experiment. So, both of these results are consistent with this idea that people cheat less when they have multiple agreements across markets.

What about this idea of high-volume, busy days, what does this mean? So, the framed results, again, I can look at this because I can vary the rewards or the benefits from cheating. And what you find here, this is roughly a change of about five times in the stakes. So, in the table two seller, these sellers are earning about $40 for this experiment; in the high stakes, they're earning about $200 on average.

And what you have, just from the change in stakes, you have cheating rates going from 16 percent to close to 50 percent. So, very large effective stakes here.

Now, when I look across the entire bridge, I apologize, again, I haven't talked about all of these particular experiments, but when you look at the proportion that cheat across the lab, the various framed field experiments and the natural field experiments, what you get is a very interesting result in that these very neutral sterile lab experiments, in aggregate, do the best at predicting cheating rates of the natural field experiment. And I'll come back to this in a few moments.

Now, I've been talking mainly about the probability of cheating, but you can also think about, well, how intense is that cheating? And here I want you to focus on the middle column. This gives you a sense of the price deviation from the agreement transaction. So, as you move across these settings -- and I just want you to focus on the bottom, the natural field experiment, these 70 or so percent of cheats are cheating at a rate of about 20 percent. So, if the agreement was for $10, they're selling for $8 instead of $10, when the marginal cost is $5.

So, as you see, now all of the comparative statics that I've talked about on proportion that are cheating, those hold on the intensity of how much you cheat as well.

So, a summary, it's going fast, but I think you see that there is no possible way that I could have made a strong inference from the natural field experiment, and the reason why is because I was not able to randomize the important features of my theory. So, that induced me to have to go back and run some framed field experiments and then, together, both of the domains told the same story. Now, I can be much more confident in making inferences from my particular data set.

Now, the interesting part about those sterile lab experiments is that they could do pretty well predicting an aggregate. But when you look at the individual cheating rates, the individuals that cheat in the particular lab setting are not necessarily those that cheat in the natural field experiment. But it is interesting that the best predictor of the cheating in the natural field experiment is cheating in the lab or framed field experiments.

So, let's conclude. I've talked about some very specific field experiments and you've heard about some very specific field experiments during the conference. But, of course, field experiments, there are many ways, shapes and forms of field experiments, and I've created this website that you're welcome to go to. There are now about 300 or 400 different field experiments on there that also have PDFs attached to them. So, if you're interested in downloading some of those, please go ahead.

And I receive nothing for this and this is not an experiment, even though it's www.fieldexperiments.com, it's just something that I thought was appropriate to set up for people who were interested in field experiments.

Now, I want to end on a methodological note. In experimental economics and empirical economics, more generally, many times people argue that -- about representativeness of the population. Many times, people say, well, I don't believe your results because your population is not representative. That's exactly what the White House official was telling me when I was arguing that we should be accounting for WTA, WTP disparities when thinking about the benefit cost guidelines.

But what always receives short shrift is representativeness of the situation or properties of the situation. We, oftentimes, generalize across situations without even realizing it, but we oftentimes want to stop ourselves or stop others from generalizing across populations.

Now, my last example will be another government example. I apologize to the EPA for this. This actually also occurred when I was at the CEA. The EPA came to me and they were interested in whether male or female surveyors raised more money in these contingent valuation surveys. Contingent valuation is a very important tool for benefit cost analysis. Why? Because it's the only tool we have right now that can estimate the total benefits of the non-marketed good or service, not just the market benefits. It can estimate both the use and non-use values.

So, what do you think they did? Well, they spent a whole bunch of money, which they should have, to carefully draw a representative sample of respondents. No doubt that’s important. But then they had one man and one woman do the surveying. Now, it’s clear that if you don’t sample the stimuli, you would come up with very different inferences. Right?

On the one hand, you have John and Angelina and Angelina’s going to do much better than John, but there’s no possible way you want to generalize that to Brad, of course, and Miss Piggy. It’s clear that you see that now, but we always, always, always forget about generalizing across situations and realizing the importance of the properties of the situation.

I think one advantage of field experiments is that you are able to vary that from the lab to the naturally-occurring data and, of course, when you change each element, you can explore whether that change induced people to act differently, and then we can think about theory or other empirical exercises to learn more about that particular economic behavior.

So, thanks for your attention. I’ll take any questions if anyone would like.

AUDIENCE MEMBER: Very interesting work. I just had a quick question. When I teach the MBAs, when I teach the lecture on collusion, I basically go back to Stigler and talk about all the things that Stigler talked about in his famous early sixties paper. It looked like most of what he was talking about you're finding as correct. I was wondering if there was anything he said that you're finding was incorrect.

MR. LIST: No, I think that's right. And that's not because I'm at Chicago. But you're exactly right. In particular, I draw from his '64 paper. Yeah.

AUDIENCE MEMBER: (Off microphone) You haven't found anything that was wrong?

MR. LIST: Not so far.

AUDIENCE MEMBER: (Off microphone) (Inaudible).

MR. LIST: And then also use my paper as empirical support.

(Laughter.)

AUDIENCE MEMBER: (Off microphone) There's a paragraph in my notes.

MR. LIST: All right, very good, very good.

MR. ADAMS: Any other questions?

AUDIENCE MEMBER: (Off microphone) Do you have anything where you vary the information environment, you know, what the sellers know about each other and prices?

MR. LIST: So, what I do vary is the group composition. So, in some cases, there are CD and DVD sellers in the same group, and in other cases, some of those same CD or DVD sellers are in a group with people who are selling cigarettes or let's say a fruit or whatever. But I don't have anything explicit where I vary the information. I always allow them to discuss between themselves various elements of pricing and how they want to price and then I record what they're saying and what they've agreed on.

But I have not systematically varied the level of information disclosure in these treatments.

AUDIENCE MEMBER: And the nature of the cheating is in the bargaining process as opposed to the posted price?

MR. LIST: In these particular markets, there is very little, if any, posted prices. All of the prices are via bilateral negotiations. So, what will typically happen is the people who are in a collusive arrangement will, just about every time, start the initial price out higher than what they've agreed on. So, people who aren't sometimes even start lower. So, that's some -- I have three or four robustness tests about is this information accurate. Is my mole telling me the truth? So, that's, of course, important.

And then what -- of course, what happens is they're cheating by going below $14 or $16 or $20, whatever the agreed-upon price is. And that's happening through the negotiation process.

Now, it's important, at this point, that I tell you that my confederates are blind to the actual sellers who are part of a collusive arrangement. So, of course, that's important. Otherwise, people always want to bring you back the results that you want and I think there's just a human tendency to want to do that. But my confederates are actually blind to that.

AUDIENCE MEMBER: Okay, thank you.

MR. LIST: Thanks.

MR. ADAMS: Great job. Thank you, everybody, for coming to this conference. Any thoughts and comments you have about it, you can send to Mike Baye, who's not here at the moment.

(Laughter.)

MR. ADAMS: Good comments come to me, bad comments go to Mike. Thank you.

(At 1:23 p.m., the conference was concluded.)

C E R T I F I C A T I O N   O F   R E P O R T E R

MATTER NUMBER: P085800

CASE TITLE: ANNUAL BE INDUSTRIAL ORGANIZATION CONF

DATE: NOVEMBER 7, 2008

I HEREBY CERTIFY that the transcript contained herein is a full and accurate transcript of the notes taken by me at the hearing on the above cause before the FEDERAL TRADE COMMISSION to the best of my knowledge and belief.

DATED: NOVEMBER 25, 2008

____

ROBIN BOGGESS

C E R T I F I C A T I O N   O F   P R O O F R E A D E R

I HEREBY CERTIFY that I proofread the transcript for accuracy in spelling, hyphenation, punctuation and format.

____

ELIZABETH M. FARRELL