Second Federal Trade Commission Conference on Marketing and Public Policy | October 18, 2024

Ginger Jin:

My name is Ginger Jin. I'm a professor of economics at the University of Maryland. On behalf of the scientific committee, which include me, Professor K. Sudhir from Yale University, Professor Yesim Orhun from Michigan, Dr. Devesh Raval from FTC, and Dr. Nellie Lew from FTC, I want to thank all of you for being willing to devote the full day to discuss academic research in the public policy about consumer protection. Oh, next slide. Okay. Okay, and thank you, Katherine.

80 years ago, when Steer and I talk about this, we thought it would be nice to have a way for an economist and marketing researchers to talk with FTC economists about consumer protection. Eight years later, we feel that demand is even stronger. The scientific committee actually receive over 130 high quality submissions for this conference. We're even more excited to see so many people coming from near and far to attend this conference in person. So hopefully, you'll all enjoy the conference.

Just a few housekeeping items. Please silent your mobile phones and other electronic devices if you must use them during the conference. Please be respectful of the speaker and your fellow audience members. And second, just be aware that this building is under security. So if you leave the constitutional building for any reason during the workshop, you will have to go back through the security screening process. Just bear in mind of this issue and plan ahead, especially if you're participating in a panel, you need to be in and out of the building. Please build in extra time. Make sure that you'll have time to go through the security. More importantly, I think most of you will receive this FTC event badge. The FTC will need this badge back at the end of the conference, so just make sure you return the badge when you leave for the day.

If some emergency happens that will require you to leave the conference center but remain in the building. We'll hear some public announcement. Just follow the instruction. If the emergency occurs and require us to get out of the building, then alarm would sound and everybody should just leave the building in orderly manners through the main Seventh Street exit. After leaving the building, turn left and proceed down the Seventh Street and cross E Street. There will be a FTC Emergency assembly area. Just remain in that area until being instructed to return to the building. If you notice any suspicious activity, just alert building security. Please be advised that this event will be photographed, webcast and recorded. By participating in this event, you're agreeing that your image and anything you say or submit may be posted indefinitely at the FTC.gov website, or on one of the commission's publicly available social media websites.

The restrooms are located in the hallway just outside of the conference room. The conference will provide lunch for everyone, but you're welcome to go to the cafeteria inside this building. The lunch hour will be open from 11:30 to 1:00 P.M. And most importantly, food and drinks are not allowed in this conference room, so only water is allowed. If you want to drink and eat, please use the outside area. If you're speaking for a session, we have reserved some seats at the front row. If you can kindly sit at the front row for your session, that will smooth the process. Thank you so much. Without further ado, we'll welcome Aviv Nevo, the FTC director of the Bureau of Economics for open remarks. Thank you, Aviv.

Aviv Nevo:

There we go. Okay, I hope you paid attention. There's going to be an exam at the end about all the instructions of where to go. Okay, well thank you, Ginger. Thanks for having me here. My name is Aviv Nevo. I'm the director of the Bureau of Economics here at the FTC. I'd like to welcome you to the second FTC Marketing Conference, or should I say the second annual FTC Marketing Conference, since our plan is to now turn this into an annual conference. With the next one scheduled, we want to be on a spring schedule rather than a fall schedule. So the next one is scheduled for March and 20th of 2026. Okay, so a year from this coming spring. So please put on your calendars, make sure to coordinate your schedule so you're here on that day. But it'd be great to see you. So I think we have a great program.

I think I can honestly say that because I had nothing to do with it, and I'd like to thank the people who did, made it happen. So first from BE, Jules, Marilyn, Aidan and Dina, and especially Katherine Chang who's running there very nervously. Katherine, it's going to be okay. It's going to happen. Thank you very much for making it happen. Also, the FTC events and media team that really allow us to set up and recording it. Kendall Milles from Yale and the Scientific Committee, Sudhir, Ginger, Yesim, Nellie and Devesh, who've really put down a truly fantastic program. And finally, I'd like to thank know SOM and [inaudible 00:42:02] for the support to making this happen, and all of you for participating. So just a quick hand for everyone.

For those of you who are outside the FTC, I want to say a few words about our agency and the Bureau of Economics, just so you have a little bit of a background sense. I think you'll learn more throughout the day. So as you probably know, the FTC is an independent agency and it has two primary enforcement missions, consumer protection and competition. The Bureau of Economics supports both these missions. We have roughly 120 FTEs with roughly 95 of them PhD economists. That makes us, well, one of the larger groups of microeconomic economists in the federal government. I haven't scientifically checked this, but I think we're probably the largest group of microeconomists, and we do a lot for the agencies. We support directly both the competition and consumer enforcement missions. We provide economic analysis and support of investigations and litigations, and we apply in many cases, cutting edge economic analysis, both theoretical and empirical, to these cases.

You'll get a sample of both what we do in enforcement work, but also the research that we do. Later on, one of our panels is going to specifically highlight what we do here and the work that our economists do. Now, this is a particularly exciting time at the FTC. We have a lot of things going on and BE is right in the middle of all the action. And indeed, to meet these challenges, we've been growing quite significantly. Over the last three years, we've hired 27 new PhD economists and we've grown, and I've seen some of the faces of our new hires here, and hopefully you'll get to meet them during the day. The agency is strongly committed beyond the enforcement aspect to research and really a strong connection to the academic community. We really think that's a very important part of what we do. Historically, this connection was stronger with the economics research community.

Specifically, we have this microeconomic conference that we're going to ... The one in about a month is the 17th one. So it's been going on for a while. But I think the marketing research community has a strong role to play as well. Several years ago, I wrote a handbook chapter for I think ... I forget what handbook it was, but something the JP ran, and wrote a chapter on literally called Marketing and Public Policy, just like the name of this conference. And in it, basically what we did is we urged marketers to get more involved in policy. Therefore, I'm really delighted to see this happening and I'm really delighted to see that ... Hoping this will really become a usual event. At the time when I wrote the handbook, I focused only on competition. There was two parts of it, but the part that I wrote really focused more on competition and how insights from understanding consumer behavior can impact what we do on competition, how that would feed in. But there was I think a lot more, or equally as much to do, on the consumer protection side. Indeed, actually part of what's happening right now at the FTC is that we're realizing it's time to break down silos and not really think, "Okay, there's consumer protection here, there's competition there." The two are very closely linked and have interaction between them. I think the Bureau of economics has really been a perfect situation to bridge the two. The work that's done both on the Bureau of Consumer Protection, the lawyers that worked on one side, and the Bureau of Competition, again, the lawyers that work on it. And we in a way sit in the middle, but we say and really can bridge the two. Just to give you a few examples, again, you'll see a lot more of this later today, is we've been doing a lot of work on so-called junk fees. Various fees that are added at various points and not fully disclosed.

I know Mike is going to be talking a bit later about our so-called CARS rule and what that does on junk fees. On one hand, that's a very clear consumer protection aspect. Say, " Okay, consumers need to be protected and not surprised by rules." But it has very clear implications for competition because it empowers consumers to make decisions that ultimately discipline how firms behave and disciplines competition. So even though it's coming from the consumer protection side, it really is in many ways a way to enhance competition. Similarly, when we talk about privacy, there's issues there that overlap both consumer protection and competition. I'm sure we'll be talking a lot about that throughout the day. So this is where many of you come in. You might not know this, but we honestly and truly rely on your research and performing a mission on a regular basis, whether it's taking theoretical results or theoretical models and methods that have been developed, or actually using numbers that you've produced and putting it into whether it's a rulemaking or deciding harm or anything of that sort.

So hopefully throughout the day, you'll get a sense of how we use that and how important it is. I therefore, as moving forward, urge you to always look for policy relevance in your papers. So try to find what is the policy relevance. Try to spell it out, try to make sure that's clear. I also urge you to interact with the FTC folks. Again, future, as you produce the works and tell them about the work. Sometimes find they might actually tell you, "Here's the policy relevance, and here's ... If you say this way, that's how it's going to be read by non-economists. But if you write this way, maybe it's clear." Again, you have to decide what you are trying to say, but just to make sure that it's clear and reflected. So hopefully this conference will increase this interaction, and you'll know the people that are working at the FTC and be able to create future interaction. And with that, let me turn to the actual substance. I look forward to a great conference and I'll turn it over to Yesim. We'll talk about the first session. Thank you.

Yesim Orhun:

Welcome, everybody. It's my honor to start kick off the first session. We have three great papers. I just want to tell you how the logistics will work because we are starting a little behind schedule. We're going to give this session an hour and 15 minutes that they were promised. I'm going to keep the speakers and discussants to their time limits with cards. The speaker will come up, do the presentation, the discussant will come up, do their discussion, and then the speaker will come up again for the Q&A session. At the Q&A session, microphones will be passed around for you to be asking questions. So let's get started with Garrett who's going to be talking about COPPAcalypse.

Garrett Johnson:

All right, this is a clicker. Here we are. Well, I want to begin by thanking the organizers for resurrecting this conference. It's a really great opportunity for us to interact with people at the FTC and find out what is the most important policy relevant questions that they're asking. All right, so I'm happy to talk today about some research that we have with my colleague, Tesary Lin, Liang Zhong from Boston University, and James Cooper who's actually at the FTC when this case was going through. So in this room of people that are in marketing, a lot of us study digital marketing and we think carefully about the benefits and trade-offs associated with digital marketing. I'd say probably as a whole, most of us have a cautiously optimistic view. But critics of digitization and of digital marketing take a more pessimistic view. They view the digital economy as being more exploitative and at the top of the list, there's concerns about privacy.

Now, these two views are difficult to reconcile, and we confront them our daily lives every time we visit a website and see one of these consent pop-ups. It's as if regulators are struggling with these opposing viewpoints as well. So we've decided to punt this to the consumer and have the consumer decide this every day on every single website that they visit. So today, I want to talk about a thought experiment. Suppose we were to get rid of cookies on the web. Well, what would happen? Well, we would get rid of advertising cookies. That would mean we wouldn't have personalized advertising. That would mean ad revenue would fall, and this would reduce the funding for the content and services we enjoy on the web.

But if we really go scorched earth, we would also lose personalization on websites. So we wouldn't see personalized recommendations, personalized search. The user would lose the ability to curate their experience through playlists and notifications, and perhaps even lose the ability to comment. So this may seem a bit extreme or the consequences of very strict privacy policy, but it's one that would certainly have consequences for the supply and demand of content. What we're going to do in this paper is try to explore those subjects empirically.

So in order to study this, we're going to look at the YouTube settlement. Google paid $170 million for showing personalized ads to kids without parental consent in violation of COPPA, which is the Child's Online Privacy Protection Act. This is actually the largest COPPA fine at the time, and it is the 11th largest FTC fine in this millennium, so I think of interest to this audience. But the consequences of this really models this thought experiment. So there's no personalization for made for kids content.

This was made worldwide for all of YouTube's made for kids content on January of 2020. It was announced a few months earlier. To decide what was made for kids content, creators would label their own content. But also YouTube had a backstop algorithm to detect made for kids content. So this turned off personalized ads, personalized search, personalized recommendations. There's no comments, no subscriber notifications, and you couldn't even add one of these videos to a playlist. Now, we also think this is interesting to study because the economic importance of the platform. YouTube is a top destination for children. It's the second most popular website online. About a quarter of the marketing spend devoted to children online is on YouTube.

So YouTubers response to this was one of a great fear. So they feared an apocalypse and they termed it the COPPAlypse. So they essentially set records in the number of comments that they sent to the FTC in 2019. Tens of thousands of comments, millions of people signing online petitions. I think a lot of the concern was that if you only have contextually targeted ads, then ad revenue could fall. Some creators had experimented with that found their revenue fell between 60 and 90%. So really, really large. So creators said that they were going to create less content, exit the market, or pivot away from making content for children. Now, YouTubers can be the melodramatic sort. So what we're going to do in this study is try to evaluate this empirically.

So in order to do that, we're going to study a little over 5,000 top U.S. channels, which are in the top categories for made-for-kids content. These are education, entertainment, and film and animation. We're going to split our sample into creators that are only releasing content that's not made for kids. Those that are only doing made-for-kids content and those that are doing a mix of the two. Our sample is going to contain 1.8 million videos between July 2018 and the end of 2020. Our data sources for this, we're getting data from Social Blade on the top channel rankings as well as some demand side data. And the YouTube data API is giving us video level data, like the time that the video is released and whether it was made for kids. So if we look at the chart, the gray non-MF mixer, the dotted line is for the made-for-kids content and it follows the trend of the non-MFK content. In fact, it's pretty similar levels. We see a gap opening up around the time of the announcement that grows quite a bit. So what we're going to do is we're going to do differences in differences and we're going to have two treatment groups. We're going to separately look at the effect on MFK channels and mixed channels. We're going to use the non-MFK channels as a control group, and we're going to allow for a pre-announcement adjustment, so that we compare the pre-announcement period to the post-settlement period.

All right, so the first thing we're going to look at is the effect on supply. If ad prices go down so much and the ability for content creators and users to meet each other in the marketplace is inhibited, we would expect this to fall. This is indeed what we find. So we find about a 13% reduction in made for kids content post-settlement, a similar number for mixed channels. We also see some pivoting away from made for kids content. So a 2.7 percentage point reduction for the pure MFK channels, and a much larger 9.5 percentage point reduction for the mixed channels. One would think that they have a greater ability to pivot away from pure made for kids content. We also see, going to Aviv's point earlier about the intersection between privacy and competition, we see larger hits to the smaller channels. So I've binned the channels by quartile, by the number of subscribers that they have, and we see larger effects for the smaller creators. Now, creators have multiple margins that they could adjust on. Certainly, one is the quantity of content that they create. But another important such margin is the content quality. The challenge is that content quality on YouTube is hard to quantify, it's in the eye of the beholder. So we're going to propose three different quality measures. The first is the originality of content. Here, we're saying that all else equal, content duplication is worse for viewers. We're going to detect this by looking at the transcripts of YouTube videos and looking for duplicate chunks from videos of the creator had released earlier.

The genesis of this was that we noticed that the drop in the supply of content was maybe less than we expected, given the large revenue decreases. When we talked to some creators, we figured that a lot of creators were maintaining the same production schedule, but they were just re-releasing compilations of their old videos. So in effect, it looked like there was new content being released, but there was no original content being released.

The second thing we look at is manual transcriptions. Now, YouTube has an automated transcription service to put closed captioning on videos. But creators can have the ability to upload a manual transcription and those tend to have much less errors. These errors hurt people that are hard of hearing, as well as people that are learning to read from captions, as one might expect children are doing with these videos. And finally, we look at user ratings. So the like to view ratio for these videos. What we find is there's a 7.7 percentage point reduction in the original content share, which goes to this pattern that we expected to find from talking with creators. We see a 3.8 percentage point reduction in the manual caption share. On a relative basis, this is a fall of 26%. And finally, we see a 10.3 percentage point reduction in the like to view ratio. So across all of our quality measures, it appears as though creators are sacrificing quality as well.

Now, we want to turn to our results on consumer demand. So what we see is that there's a large reduction in subscribers of about 29% for both MFK and mixed channels. We see a 22% reduction in the new views for MFK channels and a little bit smaller, but 14% reduction in views for the mixed channels. Turning to the effects by subscribers, we see again, larger effects on the smaller subscribers. This is the channels that are smaller, pardon me. And in fact, the magnitudes here are much larger. So we see a negative 50% approximately reduction on the smallest channels, and something close to a zero or a noisy zero for the largest ones. We think this is consistent with the loss of personalization. Personalization algorithms help move consumers down the long tail and match them to content that's more in their niche interests. When you get rid of that, you probably move to a ranking of recommendations and search that's more based on popularity. And so we'd expect more of the viewership to go to the largest, most popular channels.

All right, so to conclude, we find evidence that the MFK supply was hurt by the YouTube settlement. If you stack all of our estimates together, we're looking at something like a 20% reduction in original MFK video releases. So what that does is it adds the direct effect on reducing the number of videos, plus reduction in original content, plus the pivot away from MFK content towards not made for kids content. We find that quality suffers on both objective and subjective dimensions. As I said, we are looking at three different content categories. I think one of the most important for children is education. We see similar reductions in the production of education content as well. We see similar effects on the demand side. A little bit larger, MFK views are falling something like 22%. We see the views are becoming markedly more concentrated in the top channels when we lose platform personalization.

And so what do we take from this? Well, if we were to apply a traditional welfare logic, this would suggest that if the level of supply and the level of demand falls, then this hurts producer and consumer welfare. Now, we are not attempting a full welfare calculation. To do so, you would need to factor in the offsetting consumer benefit of improved privacy. I think we have a broad agreement in society that children's privacy is important. So that is certainly non-negligible, but we help to quantify at least one part of this. So I think I'm done a little bit early, which the organizers will be very happy about. Thank you.

Yesim Orhun:

We have Guy Aridor offering a discussion.

Guy Aridor:

Cool. Thanks, Garrett, for that. Awesome talk. I'm Guy, I'm going to give you some comments on the paper. So first, as a brief overview of what's the high level research question, this paper aims to study this trade-off between privacy and online content creation, consumption. The context is going to be obviously on YouTube in the COPPA settlement. Again, the two main changes to YouTube's experience for kids was no personalized search and recommendations and also no personalized ads. So the main findings on child content are roughly on the supply side. Again, they find that quantity goes down, quality goes down. Consumption goes down, and concentration goes up on the demand side. Now, the bottom line is that content creators for child's content surely are hurt. On the consumer side, it seems like they're hurt, but there are some questions that I'll get to that later. My main comments on the paper, again, it's very timely.

Again, we're considering revamping COPPA. So it's very important to understand what the economic implications of this law are. I think it's really important. Again, given the set of authors, not super surprising. It's a very well-executed, well-done paper. So most of my discussions is going to be on policy implications and things that I think will be interesting for future work. So the first is to start from the supply side. So I have a very, very stylized model up here that's from a JL paper we have coming out later this year, that thinks of the content creator problem as again, picking some vector of posts, where effectively the content creators trading off the expected reward minus the expected costs. So what do we have with this COPPA settlement? Again, we kind of have overall reduced expected benefits from the search and personalization side. Again, we think of this reducing the likelihood of discoveries, reducing the expectation, the reduced per view revenue coming from the lack of personalized ads impacts the direct utility term there.

So I think the question we have then is if we genuinely cared about fixing this. So I think I argue on the next slide that there may be some genuine value to having production of child content. What can we do? So privacy regulations not implemented in a vacuum. Can we think of pairing the lack of the reduction in expected rewards coming from privacy regulation with other types of levers? So I won't go into too much of the details here. But in this review paper, we highlight all the various incentives that user-generated content platforms like YouTube and various social media platforms, have to offset this. And so if it's going to be that content creators are going to have reduced revenues from personalized ads, UGC platforms rely a lot on non-monetary incentives. Can we think of external funding for this?

Okay, now, from the demand side, the overall motivation for a law like COPPA is again, we get privacy gains for children, which has two benefits. As Garrett said, there's a non-instrumental value to this. We genuinely care about having privacy for our children. The second is obviously to prevent bad targeting. We don't want our kids to go online and get exposed to content that we don't want them to. So the policy objective here is to have a safe environment that has healthy content. The key thing here that we want to understand is what are the equilibrium effects of these regulations and in particular, this settlement? So we want to aim to avoid a thing where we go after YouTube, and then consumers go over to Instagram or Facebook. So I think for future work, there's two things that are interesting to think about.

So the paper looks at the overall reduction in consumption of child content. We know from past work and developmental psychology, that for instance, made for child content on public TV like Sesame Street can have some developmental benefits. So thinking through what are the equilibrium effects in terms of what kinds of content are consumers exposed to, I think the first question we want to ask is, what are healthy content diets for children here? Is there a developmental benefit to this? What are the developmental benefits potentially from online content? Is it the same as the content that was on PBS? How do I think about putting that? And then the second was to have a better characterization of these kind of content diets that children have after this lack of characterization. And so in this particular context, there's a few things.

We can think maybe all the demand that was going from made for child content on YouTube, now, people are going on Instagram or they're going on Netflix, and these have very different implications. So again, there's an increase in social media use. We do have some suggestive evidence that likely, this is probably not good for the long-term mental health of the kids. So again, thinking about these equilibrium effects is important. The other thing is non-children content. So one detail from the YouTube case specifically is that some children use the YouTube kids app or they're constrained only to child content, but other kids will also potentially use the main YouTube application. And here, they may be exposed to non-children content. And so what ends up happening is that we basically get rid of personalization and kids are going organically to this bad content that defeats the purpose of the regulation. So I do think that better understanding when we shut down personalization, what types of content do kids actually go to? Does it have similar or worse developmental benefits, I think is important. But overall, again, great paper. I think there's a lot of interesting dimension of future work here. Better understanding, again, how to incentivize children's content, and the equilibrium change in content diets. But overall, I think the paper is a great paper and a really important contribution to the policy debate. So thanks.

Yesim Orhun:

Thank you. Now, if you have a question, please raise your hand and we will run microphones to you.

Garrett Johnson:

Thanks, Guy, for that really thoughtful discussion. I don't bite. I'm Canadian.

James Thomas:

Hi, it's James Thomas from the FTC. Maybe beyond the scope of the paper, but could you speculate maybe on-

Speaker 2:

... may be beyond the scope of the paper, but could you speculate maybe on to what extent the effects were due to lost revenue through worse targeting versus decrease in personalization of content?

Garrett A. Johnson:

The million dollar question. Yeah, I think the challenge here is that we have a bundled treatment and so we have multiple things happening at the same time and so it's hard to apportion that. My personal opinion is ... So we showed you some evidence that other people did. Other people ran this test where they turned off personalized ads and it decreased the ad revenue for those creators from 60 to 90%. This is beforehand so this is sort of not [inaudible 01:08:42]. We have ad price data from one creator that suggests that their ad prices fell by 72%. So I think of this as an economist and think, "Wow, that is a very, very large effect," and I think that personalization effects are also important, but my personal belief is that a 70% reduction in revenue is first order, but this is something we're actively working on better talking about more carefully in the paper. But yeah, I wish I could say something more precise than that. Sorry.

Speaker 3:

I was wondering if you can talk a little bit about what kind of feedback were these companies using for targeting before this happened. Because on YouTube Kids app, that's a big deal in our household and I don't see anybody clicking on it. So is that happening through past clicks or through some profile of the household that they've made? Because it seems to be a big impact here.

Garrett A. Johnson:

Yeah. Thanks Dabhi. So YouTube kids from the beginning never used any sort of ad personalization. It's all contextual ads. The YouTube main app was using personalized ads and if we think a little bit about what that means, I don't know about your kids, but my kids are probably not visiting a lot of commercially relevant websites.

So I think what was going on and certainly what we've heard a lot from creators is that effectively the targeting was happening at the household level. So maybe mom was searching for a new car and so she was being retargeted with car ads or the child was being retargeted with car ads, and maybe they're making some bet that maybe mom and dad are close so that they're actually seeing the ad or maybe this is just a misfiring of targeting, but I think that's certainly what's a part of what's going on here is that a lot of the targeted ads that they were seeing were targeted at the parents rather than the child. We don't have any ads data to really quantify that.

Hi Andrew.

Andrew:

Hi, Garrett. I have read the paper but it's been a while because you've had this paper for a while so it's good to see it getting more attention. I forget. Did you look at exit? Because it strikes me that one of the issues we might see here is that there's just not a lot of outside opportunity for some of these YouTubers. So if they're kind of stuck with the platform, they're not necessarily going to be reducing the amount of content very much. So it seems like there's some stickiness there.

Garrett A. Johnson:

Yeah, so there is not exit. So exit could mean different things. The simplest way to look at it is when you stop releasing videos. So if you look at the last video that you release, there's a very marked difference between the mixed channels and the non made-for-kids channels and the made-for-kids channels and we have a kind of a coarse figure of that in the paper looking at do you release any videos in 2020. But we plan to augment that with something that looks month by month, and you can see that they're really on different trends. So that's definitely an important part of the story too. Thanks.

Speaker 4:

Thank you.

Garrett A. Johnson:

Great. Thanks everyone.

Speaker 5:

Next, Alessandro Acquisti is going to present Behavioral Advertising and Consumer Welfare.

Alessandro Acquisti:

Thank you. Thank you to the FTC for hosting this event. I'm Alessandro Acquisti. This is joint work with Eduardo Mustri who is in the audience and Idris Adjerid. The motivation behind this work is the following. In the public debate around behavioral advertising, you see two contrasting views. On the one side, privacy advocates point out the invasiveness of behavioral advertising, which relies on extensive tracking of consumer behavior and this tracking is then used to target ads to consumers. On the other hand, industry, especially the ad tech industry, point that this formal advertising is economically beneficial for a number of different entities including the consumers themselves.

For instance, after the Federal Trade Commission asked for comments both on its groundbreaking rulemaking on commercial surveillance as well as the request for information on digital advertising business, it was common to see, among the comments, statements such as this. This one is from the IAB suggesting a direct causal link between digital advertising and significant intrinsic monetary value to consumers.

However, how much causal rigorous empirical evidence do we have about that link between behavioral advertising and consumer welfare effects? It turns out if you look at empirical evidence, the evidence is remarkably opaque. Let me explain what I mean by this. You can think about these effect as both direct and indirect. Indirect is essentially a mediated effect, the argument being that behavioral targeted ads have more value, therefore they bring higher revenues to publishers and therefore publishers can provide more and higher quality content. It is difficult here to disentangle correlation from causation, and typically researchers use event studies to try to tease out the effect. And we just saw an example here by Garrett and indeed there is research suggesting that exogenous events which reduce tracking and therefore targeting also reduce content availability.

But there is also other evidence, including articles that I co-authored with [inaudible 01:14:57], which suggest that, in response to major shocks such as the GDPR or Apple ATT, there was actually no significant negative effect on the provision of content by EU media sites compared to US media sites in the case of GDPR or in terms of apps availability and quality in the Apple ecosystem compared to the Android ecosystem. So this is an indirect effect.

What about the direct effect? And this is the focus of the paper I'm presenting today. The direct effect is the claim that consumers benefit directly from ads. And typically industry frames this in terms of these ads are more relevant to consumers. And yes, there is evidence of that and we know this because behavioral targeted ads do tend to get higher click-through rates and higher conversion rates than non-behavioral targeted ads. However, and this is the key, mainly these higher conversion rates tell us about a reduction in consumer search costs and the utility a consumer derives from a transaction is not just a search cost. It's also things such as the price the consumer ends up paying for trading through a behavior targeted ad compared to, for instance, an alternative [inaudible 01:16:17] search or the quality of the product or the quality of the vendor.

And notwithstanding the fact that in the last few years there has been a growing amount of excellent empirical research on the consumer side, that research has focused mostly on search ads or on specific advertising campaigns. So what we lack is a holistic understanding of the behavioral direct welfare consumer effects of behavioral advertising. And what we tried to do through experiments on the consumer side rather than on the platform side was to try to get this holistic analysis. And I will explain what I mean by counterfactual approach.

So we ran two studies, study one and then a replication study, study two. The studies attempt to compare objective metrics such as the price consumers pay if they buy through an ad versus they buy a competitor through a search and subjective metrics such as relevance, perceived quality of the product, et cetera. We compare the products associated with ads which had been behaviorally targeted to our participants to a competitor found in organic search results on Chrome, the first page of Chrome, to random products. By random products, I am referring to the products which had been behaviorally targeted to other participants in this experiment.

In a nutshell, participants with their own computers and their own browsers, we focus on Chrome users because it's the most common browser and it's also by default the most trackable, would be assigned to visit randomly picked websites where we knew behaviorally targeted ads were likely to appear and were asked to find an ad and then extract it through a tutorial to right click on the ad and select the ad URL and then submit it to us through an experimental platform. After this link was submitted to us, this is an example of a link you can see at the very top the advertising network and at the very bottom highlighted is the landing page, we use script to visit the landing page and capture information on the landing page so that we could capture the price of the product, the vendor, and other information associated with the product.

You may be wondering whether the information we, through our scripts, would found on the landing page could be identical to the information that the participant would encounter as they click on the ad. That's a crucial question for our design to make sense. The answer is yes. We ran several pilots confirming that 100% identical information when we use that long URL. Essentially, all the targeting and personalization happens before, not when the person is then clicking on that long URL. And then, we use script to search for competitor's products online and we use a Google search engine and then we randomly selected one out of the 10 or so organic results on the first page of Google search results. In addition, we also studied not just competitors but identical products. I'm going to explain what I mean by that in a couple of slides.

Finally, our participants were called back about two weeks later after stage one, called back and presented products that they had seen, products from the search and random products in random order and they were asked questions about these products. Is this relevant for you? How do you like this product? And so on and so forth. So in a nutshell, this was the three conditions within subject design. The three conditions are the ones I told you but, because this is a short presentation, I will only focus on behaviorally targeted ads versus search results. And as I mentioned, we capture both objective and subjective metrics. But again, due to the time limitations, I will focus mostly on objective metrics.

The first objective metric I want to share with you is vendor quality. We capture vendor quality through Better Business Bureau data as well as Sitejabber. This table only shows Better Business Bureau data. The results from Sitejabber are consistent with the results I present here. As you know, Better Business Bureau ranks businesses through a letter rating and here we have the distribution of vendors in ads and in search. You can observe a higher likelihood of better rated vendors in search compared to ads and you can also observe a higher likelihood of F-rated vendors, the kind of vendors who don't send the product or the product malfunctions, they want to replace it and so on and so forth in ads relative to search.

The second result I want to show is prices. These are log prices, again, ad versus search. You can see a statistically significant difference. The prices on average are higher in ads than in search for competitor products. I'd hinted to you that we also look for identical products. So some of our products were unique, were only sold by a certain vendor. Other products were sold by multiple vendors. In that case, we call them identical products. We can actually get all the vendors on the first page of the Google search results, organic search results, and we can compare those prices to those of the vendor, which ad behaviorally target our participants. And we can see here that when you compare the price of the vendor in ads to the prices of the identical products sold by other vendors in search, over 52% of the times you could find a lower price in search rather than in the ads.

So this was somewhat surprising, at least for some of us, and we started asking ourselves why and how this can be an equilibrium. Well, it so happens that Hal Varian, way back before the internet, in the 1980s had this lovely paper that I'm sure many of you know about, the model of states, in which he imagined a market where there are two types of consumers, informed and uninformed. And these sellers strategically can end up in a mixed strategies equilibrium in which sometimes they price low to get the informed consumers and sometimes they price high to get the uninformed consumers and get higher revenues from the uninformed consumers. Is it possible if something similar may be happening here where the informed consumers are those who search and there is price competition there? The uninformed consumers are those who click on the ads. So we cannot answer that directly, but we can look at the distribution of vendors in ads and in search. So this is the distribution of vendors in search. Even though the font type is quite small, I'm pretty sure you can imagine what the top vendor is. It's Amazon, right? What is interesting here is that all the vendors appearing in the first page of organic search results are all large vendors, Amazon, Best Buy, Target. It is nearly impossible for a small vendor to appear on the front page of organic search results. Which means that if you are a small vendor and you want to get to consumers, you have to buy the ads.

But, and this is the key, among these vendors in the ads you see the large ones and the small ones, websites I personally never heard before. And among those vendors, there are some who are legit good vendors and some instead who are F-rated high price vendors. This suggests, therefore, that yes, our results could be a long-term equilibrium which we started to test by repeating our experiment one year later. And these are the results one year later. So first study, study one, study two, same distribution of vendors. In fact, you can see, very interestingly, that in the ads we again find this mix of well-known ... in ads you can find this combination of well-known vendors and unknown vendors, only that they are no longer those from one year prior, meaning that the small vendors keep changing, but there is always a fraction of small vendors in the ads.

This was the results for quality study one, study two, exactly replicated. These were the results for log prices for competitors' products. Study one, study two, replicated. These were the results for prices identical, study one, study two replicated. We also test for robustness to multiplicity of different specifications. We ran a latent utility analysis and furthermore, finally, the one subjective matrix that I want to share with you is about relevance. We ask our participants how relevant was the product advertised to you?

And naturally the behaviorally targeted products were rated by our participants as being more relevant to them than the random products. That makes sense. However, once we control in study two for whether the participants had searched for a similar product before, the difference that increased relevance of behaviorally targeted products relative to random products disappears. This suggests that at least within our sample of participants, the informational value of the behaviorally targeted ads was not really about discovering new products but discovering new vendors or perhaps new brands such as small vendors. But again, among these small vendors, we have a combination of good vendors and F-rated vendors.

To clarify limitations, this was done on a very specific setup, Chrome users, Google search, display ads. So we cannot really say much about all other ads that happen online. But in a nutshell, search results are dominated by large vendors and therefore the ads have an important role to play because they allow the smaller vendors to reach consumers. However, among those vendors, we do find a higher relevance of lower quality vendors which sell at a higher price. And the impact on search costs is perhaps ambiguous given what I showed you. Thank you very much.

Speaker 5:

Thank you. Li Jiang will offer a discussion.

Li Jiang:

Hello. Hi, I'm an assistant professor at George Washington University. Thank you for letting me join remotely as I travel for another event.

So today I'm going to talk about behavioral advertising and consumer welfare, the paper by Alessandro Acquisti. Next slide. So, online behavioral advertising tracks consumer data and use the data for personalized advertising. First, I want to highlight the background of this paper. There's a policy debate between different players. It's always about online behavioral advertising and one claim of online behavioral advertising, especially by the advertisers, is that it's a win-win-win situation for advertisers, publishers and consumers. Advertisers tend to say it's online behavioral advertising that funds the free internet. Such claims are often cited to counter regulatory interventions. For example, FTC may want to regulate online behavioral advertising. However, advertisers always have this counterclaim. However, we don't have enough evidence to speak to this claim and whether this is true or not.

Next slide. So there's a tension between consumer privacy and economic benefits. So online behavioral advertising deliver personalized advertising at a cost of consumer privacy. However, the economic benefits are distributed between all those players, publishers, advertisers and consumers. And we know, from the value of data research, publishers and advertisers do get a lot of revenue and a lot of benefits from online behavioral advertising. However, the benefits of our consumers are unknown and we don't have a lot of literature speaking to that. So this paper answered the question, are consumers better off or worse off? And this is a first attempt to answer this question.

Next slide. So for the sake of time, I'm going to comment ... Sorry, for the sake of time, I think my slide get disrupted somehow, so I'm on the slide, Uniqueness of The Research. So I'm just going to comment on the uniqueness of research. Two aspects, very important. One, research is about uniqueness of the literature space. Most research tend to focus on firm outcomes instead of consumers. For example, research related to GDPR. And because of that, most research tend to focus on a global research not tied to a specific platform. However, this research also used a unique approach in which you collect naturalistic behavior globally and measure price and vendor quality as opposed to ad effectiveness and conduct counterfactual analysis.

So I'm not going to go through the design very carefully because Alessandro have done this. What I want to say is that they actually use this very smart design in which they ask people to browse internet websites and submit URL landing page and then use a script to create the competing ... advertise the product and create two other conditions, search condition, and target condition to contrast with ad condition. And what they find is that small vendor actually are shown to the consumers more in the ad condition compared to a search condition, and also behavioral advertising leads to lower quality vendors and potentially higher prices.

So now I'm on the slides of limitation in future direction. I just want to point out a few limitation in future directions. So the first one is how to define consumer welfare. This paper used price and also vendor quality to capture consumer welfare. But consumer welfare also involve other things. For example, are consumers satisfied with the purchase or not? Those are the other things to think about in terms of consumer welfare. Also, this paper also only focus on desktop Chrome users as in the internet people focus on mobile purchasing and also social network purchasing. So if this paper can expand to that space in the future, that would be great. So with that, I will just conclude and leave room for the questions.

Alessandro Acquisti:

Thank you, Li. We were playing catch-up with the slides and you and everyone here handled that very gracefully and masterfully. So thank you very much.

Speaker 6:

Two questions. So search, Google wants to match people to reputable companies with low prices and high ratings. So I guess my first question is do we learn that behavioral targeting is negatively selected or just that search is positively selected? And then, my second question is just one of counterfactuals. So if we get rid of behavioral targeting, when I use Safari, I see a bunch of ads for earwax cleaning and tooth whitening. So I guess I associate the non-behaviorally targeted with low quality sellers. And I'm curious how you can tell which of the ads are behaviorally targeted versus not behaviorally targeted. Thanks.

Alessandro Acquisti:

Thank you. One clarification about the second question. Are you asking technically how did we know that the ads were behaviorally targeted? Or you're asking an economic question?

Speaker 6:

Yeah, the technical question, like how can you tell that the bad ads are actually coming from behaviorally targeted versus they would've still shown up to a user if they didn't have cookies? Thanks.

Alessandro Acquisti:

Thank you. For question number one, exactly. It's an excellent question and we do not know. We cannot really disentangle the element of how good search is essentially Google doing a good job through its algorithm in filtering out bad vendors versus how bad behaviorally targeted ads are, other than sticking to the results that do suggest the higher prevalence of F-rated vendors on those ads.

On the technical question, on the second technical question, we did a number of things before, during and after experiment. Before, we selected websites where we tested for high prevalence, about 80, 85% of ads behaviorally targeted on their front page, and that tended to be news and media sites. We also look at data such as the icons which appear next to the ads, which tell you, on the Google Advertising ecosystem, whether the ad is in fact being behaviorally targeted. We also select the users who use Chrome and we check their settings to test whether or not they were using tools which would make them less trackable. And, finally, the self-reported answers, which suggest that the ads were indeed relevant to them more so than the random ads, does support the view that our efforts were successful in that regard.

Speaker 7:

Oh hi. So when you run these experiments on Facebook with several million users and you turn off personalization, you see that some of the metrics seem to go the other way. So consumers spend less time on the platform, they purchase fewer things, they report less satisfaction with purchases that they do make. If you look at incentivized willingness to paid surveys, people are actually willing to pay more for a personalized version of Facebook. This is just sort of from ad personalization. So I was wondering how you could rectify your findings with those findings. What are differences across the settings that might speak to that difference?

Alessandro Acquisti:

Sure, great question. They are essentially different metrics because some of the metrics you mentioned are metrics that we do not capture and some we do capture in ecosystem with our results, but our key results are not included in the metrics you were suggesting. So let me go in order.

Some of the metrics you mentioned which we do not capture are how ads may lead to more or less time spent on a platform, more or less time spent searching. That we cannot capture. We cannot talk about that. We do find indirect evidence that consumers do find those ads more relevant, suggesting that there is, as other researchers found, a reduction in search costs. So in that regard, our results are either consistent or not directly capturing. Or, let me rephrase it, not directly capturing total search time but, for the element that we do capture, consistent with other results.

Second, we find, again, in terms of satisfaction with the ads, we do find that our participants do claim that the ads behaviorally targeted to them were more relevant. However, unlike private work, we focus on these objective metrics, price and vendor quality. So we don't just look at whether the consumers are saying, "Yeah, I find this ad more relevant and I find this product relevant." We are actually searching, looking for whether the price they would end up paying and the quality of the vendor they would end up buying from are in fact different from what they would've experienced had they searched for the product. That objective metrics, I believe, is what makes this study different from the studies that were mentioned. Thank you.

Speaker 4:

Thank you. We have time for one more question.

Mike:

Hey Alessandro. Super interesting. Quick question-

Speaker 5:

Sorry, the microphone is with Ken. Oh, there are two microphones. Go ahead. Go ahead. Sorry about that.

Alessandro Acquisti:

Mike, please. And then I will discuss with you offline. Thank you.

Mike:

So just trying to get a sense of this in the context of some of the other sort of steering issues that are kind of lurking about, right? So if you think about there's advertising as one thing that Google might be using to sort of draw attention to things, then there's also, I'm thinking of the Google shopping case and then uses of defaults. Are these all the same sort of underlying behavioral mechanism or do you think these are all sort of slightly things that are happening in these different aspects of creating a choice architecture online?

Alessandro Acquisti:

Very interesting. It's a great question. I'm afraid I do not have a good answer other than that will be something worthwhile studying further. We were asking ourselves, as we found these results, whether they may be indicative of a longer term equilibrium. And that's why we decided to repeat one year later. But, even so, although we found evidence that there may be a long-term equilibrium, we cannot extrapolate our results to other platforms where other types of nudges or choice architectures may occur. So this may be potentially a valuable venue for further research. Thank you.

Speaker 4:

Thank you for a great discussion everyone.

Alessandro Acquisti:

Thank you.

Speaker 4:

Next, Malika Korganbekova will present Balancing User Privacy and Personalization.

Malika Korganbekova:

Hi. Thank you so much to the organizers for including our paper in the program.

So this is a joint work with my fantastic co-author, Cole Zuber, who used to be at Wayfair when we were working on this project. So, as you all know, many online platforms, they personalize user content. For example, platforms like YouTube, Twitter, they may personalize advertising and retail platforms like Amazon or Wayfair may show personalized prices. But all of these platforms, they also personalize the actual content, the products that are shown on their premises, which is exactly the focus of my paper. So, to give you an example, imagine you go on Amazon or Wayfair and you're searching for a particular style or color of furniture. Then the next time you arrive on the platform, the platform will show you other products that are similar to the ones that you previously browsed.

So, on the one hand, it is nice that the platform is trying to be relevant. They're trying to show you something that you might like as a consumer. But if we stop and think what makes personalization possible, it is online tracking. So when you browse online, very detailed individual level data is being collected about you. And so this started raising regulatory concerns over user privacy.

So first of all, there is a concern about general data usage where, as consumers, we might not know what kind of data is being collected, how this data is being used, what is the objective function of the algorithms that is using this data. And then, there is more personalization specific concerns. So regulators such as CMA in the UK and FTC in the US have said that personalization can be harmful because it can manipulate consumer choices even when the consumers are not aware of that. And so, as a result of these privacy and personalization related concerns, although US does not have federal level regulation, tech companies such as Apple and Google started limiting online consumer tracking. And what that means is that the platforms are no longer able to track consumers within and across websites.

So, in this regard, the research questions that we ask in this paper are the following. First of all, is it empirically true that personalization hurts consumers and sellers? Are regulatory concerns justified? And I will show you experimental evidence that suggests that actually personalization can benefit consumers, sellers, and the platform overall. Then, the next question then becomes how will privacy restrictions that limit the ability of the platform to recognize consumers impact these benefits? What will be the underlying impact on different types of consumers, sellers and the platform? And, finally, in the last part of the paper, I will start thinking what can platforms do to mitigate the losses from privacy regulation.

Malika Korganbekova:

Platforms due to mitigate the losses from privacy regulation. In terms of methodology, we answered the first question about the impact of personalization using a two-year-long field experiment that I ran together with Wayfair where we randomly turned off personalization for a sample of consumers. And so this allows me to make causal statements about the impact of personalization. To answer the question about the privacy restriction impact, what I do is I take Wayfair's actual personalization algorithm, I retrain it using lower- quality data to mimic different types of privacy regulations, and then I use the structural model of search to simulate the counterfactual market outcomes. And finally, in the last part of the paper, we propose a very simple probabilistic identity recognition algorithm to help platforms recognize consumers in a probabilistic way. In terms of the empirical setting for this project, I collaborate with Wayfair. It's an online retail platform, and one of the things that Wayfair does is they personalize product ranking pages.

And so just in terms of what it means at a high level is that if as a consumer you are looking for blue chairs, then the next time you arrive on the platform, the platform will make more prominent other blue chairs. Similarly, if you are looking for white chairs, they would make more prominent white chairs. This is a very trivial example, but you could think of multi-attribute personalization. And so the way personalization works on a platform like Wayfair is the following. Again, at the high level, the input to the algorithm is the individual level consumer search data. So basically it is the clicks add to carts and purchases that each consumer makes. And then this gets fed into the deep learning based algorithm. And the algorithm's job is to understand the similarities between the products and between consumers and to output personalized set of rankings for each individual consumer.

And as you can see, the important part of this whole process is the input data. So if the input data is not good enough, then the algorithm simply cannot output good personalized results. And so one of the challenges for platforms like Wayfair is that it is hard to track and recognize consumers across time. And usually consumers arrive for multiple sessions, they use multiple devices, multiple browsers, and on average only 40% of consumers voluntarily log in, which makes it super easy to connect all the sessions. And what that means is that for the remaining 60% of consumers, the platforms have to use online tracking technologies such as first and third party cookies. And this is exactly where the privacy restrictions are going to impact platforms like Wayfair and many others. And so in terms of data, I have data from Wayfair where I observed the device, the browser, and the source of traffic when consumers arrive on the platform, which means that I know whether the consumers arrived directly to Wayfair.com or did they arrive from advertising channel.

And I know the product rankings that were shown at each consumer page load. I have very detailed pixel-level clickstream data that tracks consumers actions on the platform like clicks, scrolling behavior, et cetera. I observe final purchase decisions and product returns if any. And so basically on the consumer side, I know where consumers are coming from, what they do on the website and what they purchase on the website if anything. And on the seller and platform side, I observe the daily retail price and seller-set wholesale prices, which allows me to calculate metrics such as seller revenue and platform revenue and profit.

So now I hope it's very clear in terms of the empirical setting. And now I want to start answering the research questions. So the first question is about the impact of personalization. So to understand that together with VTR, I ran this field experiment, there were more than 9 million consumers that were randomized either into treatment or control condition for a period of two years across all product categories. And so basically in the treatment condition, consumers would see personalized set of rankings, same ideas before the rankings are a function of your previous browsing history versus in the control condition. Regardless of what consumers were browsing previously, they will always see bestseller non-personalized set of rankings. And so this experiment allows me to make causal statements about the impact of personalization.

And so what I find here is the following. So first of all, consumers in personalized group filtered less, they scrolled less, they decided what to purchase faster. So they saved two days of time searching and they were also more likely to purchase. So the first three sets of results, they're telling us that personalization might be decreasing consumer search effort. However, the fact that consumers are purchasing more does not necessarily mean that they're better off. And to understand whether the consumers are better off or not, we have to look at the purchase outcomes. And so when I look at that, what I find is that consumers in personalized condition, they were actually purchasing slightly more expensive items, but at the same time they seem to get better match values, which I measured using the product return metrics and repeat visits. So product returns decrease significantly in the personalized condition. And also consumers were more likely to repeat visit and repeat purchase on the website. So the takeaway here is that it seems that on average personalization is benefiting consumers on the website because they're getting better match products.

Now on the platform and seller side, what I find is that the platform's revenue increases significantly and its profit also increases, which is kind of natural in the sense that because conversion rates are increasing and consumers are purchasing slightly more expensive items, revenue should go up. But what I think is very important to document is that the profit increase is not driven by the fact that consumers are buying higher-margin items. The profit increase is solely driven by the fact that consumers repeat purchase. So per consumer profit is increasing over this two-year long period. However, consumers are not purchasing higher-margin items. And on the seller side, what I think is also super important to document is that it is less experienced smaller sellers that are benefiting from personalization, where they're getting more revenue because they're now shown at the top of the search page results. A.

Nd I think it is important because from the personalization algorithm standpoint, it is not important who is the seller as long as the product is matchable to the consumer, right? And in that sense, both small and large sellers get equal opportunity to be shown at the top of the search results. And so the takeaway here is that it seems at least like in this case study of Wayfair, it seems that personalization can benefit consumers, it can benefit smaller sellers and the platform. So it's not a zero-sum game.

And now the question becomes what will privacy restrictions do to these benefits? And just to remind you, the challenge with privacy arises because the platform has to track consumers across different sessions and only 40% of consumers usually log in. And so one of the privacy restrictions that's in discussion right now is the Safari first-party cookie expiration policy where Safari wants to set a seven-day limit for first cookies. And so imagine there is a consumer who arrives on the platform for the first session, and then more than seven days later, he arrives for the second session and then again they arrive for the search session, but within a week. So what this type of Safari policy will do is the following. So as a platform, the platform will not be able to recognize consumers who arrived more than a week later, which means that the data becomes fragmented.

The platform will think that the second session is generated by a complete new consumer. So instead of having full search history for each consumer, the platform now starts working with fragmented data. And to give you the scale of the problem, on average, 30% of consumers arrive from Safari browser. And so to understand what is the impact of this type of privacy policies, what I do is the following. So here is the original personalization algorithms that Wayfair uses. Now what I do in the counterfactuals is I retrain this algorithm using lower quality data. So I basically take the input data because I know the source of traffic. I manually delete data for Safari users and then retrain the deep learning algorithm again to generate the counterfactual set of rankings that would have been generated had the privacy policy been in place.

And afterwards I developed a search model to help me understand what are the counterfactual market outcomes. And the outcomes of interest are changes in consumer choices, seller revenue, platform revenue and profit. And due to time constraints, I'm going to be very quick on this. So on the search, the way search model works is that consumers arrive on the platform, they have valuation for each product, they form expected value about the utility of the product that they see, and they basically decide whether they want to click on a product, they want to scroll down to discover additional products that they have not seen so far. They can leave the platform to go to an outside option or they can purchase the best products that they've seen so far. So basically this is what the model is trying to capture.

And afterwards, I can basically use the model to calculate the counterfactual market outcomes. And what I find is the following. So on the consumer welfare side, consumer welfare decreases by almost 20% as a result of Safari policy restrictions. And the main group of consumers who are impacted are the price-responsive consumers. So in the data, what I observed is that consumers who are more price-responsive. They also tend to search less, which means that the platform usually has less data to train on for those type of consumers. And by imposing the privacy restrictions, we are kind of taking the data away from the consumers who already have little data to train on. And that's why their welfare is decreasing significantly. And on the seller side, the smaller sellers are impacted more. And again, the same logic applies because for the smaller sellers, the marginal value of a data point is very high.

If they start losing data, the algorithm just deprioritizes them in the ranking results and the larger seller's revenue starts to increase. On the platform side, I observed that both revenue and profit will go down. And I think one of the limitations of this type of welfare calculations, as Garrett mentioned, is also that consumers have intrinsic privacy valuation, which is not captured in this type of welfare calculation. And this is something that I'm working on right now. And in the last part of the paper, we start thinking what can platform do to mitigate the losses from privacy regulation? And we come up with a simple probabilistic identity recognition algorithm, whereas the idea is to use the device-level IP address and behavioral data to predict the user labels in a probabilistic way. And so essentially it's an XGBoost-based algorithm. And the idea is if there is a consumer for whom we don't know who the consumer is, can we probabilistically guess who among the existing consumers could be the closest to the unknown consumer?

And so in the last part of the paper, I show counterfactual results simulating using probabilistic identity recognition in the privacy-restricted world. And the blue bars are the previous Safari results. The green bars are the results from the probabilistic algorithm. And as you can see, some negative consequences of privacy regulation can be mitigated if the platforms are allowed to use probabilistic identity recognition. So to summarize, what I showed you today is that personalization can benefit consumers and sellers. It is not necessarily automatically harmful. Privacy restrictions will primarily hurt more price-responsive consumers and smaller sellers. But however, there is a way out where platforms can partially mitigate these losses if they use probabilistic identity recognition. Thank you so much.

Speaker 8:

Thank you, Malika. Scott Shriver will offer discussion.

Scott Shriver:

Green button. All right.

Speaker 8:

[inaudible 01:54:57].

Scott Shriver:

All right.

Speaker 8:

Before you hand off, advance to your last slide.

Scott Shriver:

Okay, thank you very much. Let me just jump straight in to the summarizing the paper. So at its core this is an empirical measurement exercise, trying to value consumer clickstream data both in terms of product search and product purchases, both to firms in terms of revenues and profits as well as to consumers in terms of the impact on their welfare, where there's explicitly not only just the product utility but also the search costs involved. How it's done, there's both a very nice field experiment plus a structural model that is developed and a series of very nice counterfactuals, more than enough material for probably two papers here, in my estimation. And big picture, it's informing the public policy and firm policy regarding consumer privacy elucidating this trade-off between the use of personal information or consumer privacy. So this is a very nice paper. So almost everything I'm going to be saying here is just refinements to a very solid piece of work.

It's obviously policy relevant anytime you're dealing with experimental variation as opposed to observational data. We're very happy about that. Fewer concerns for indigeneity, even though the environment is not fully controlled here, the data set itself is extremely rich. It's very rare we get complete clickstream data. Also encoding of product image data. The structural model is then calibrated using this experimental variation, which is a very nice aspect of this paper and it uses some state-of-theart modeling in the form of these Gaussian processes in the modeling itself. Okay. The field experiment's been described at Wayfair.com, it's this dining chair product category. The observation window is 2020 to 2021, so it's during the COVID era and it would be nice to have a little bit more discussion about how that kind of supply shock might impact things. The treatment specifically is whether or not the product ranking page has been customized based on prior browsing session history. And by nature of that the treatment is going to be both. And of course, so this makes me wonder a little bit about the SUTVA assumption. Neither stable nor uniform.

So a little bit more discussion, that would be nice I think. So what do they find broadly the key outcomes are increasing one to 2% under personalization versus not with the exception of clicks. Just a few things about the structural model itself. So it's very nice. We have the modeling of the complete search behavior, both clicking and scrolling. Again, we very rarely see scrolling data incorporated as well as leaving the site. Underneath the hood here, the search process is a heuristic type of search process. It's not a Bellman equation, it's near optimal. And the Gaussian processes are used to incorporate this Bayesian learning. So this very nice feature, and it's very natural how when you click on it it reduces the uncertainty about the valuation of a product.

Something that I did have a little bit of difficulty following is there are a lot of different variance components in this model, both for product utility as well as the search costs themselves. So I'm wonder about using different scales of utility for search costs and product utility and how that impacts the welfare calculation. Similarly, the scrolling costs appear to be deterministic and I wasn't quite sure why that was for that choice and proportional to the product rank. So I'm not sure whether that implies that you have to go all the way back to the beginning of the search results for each time or not. So that's just a clarification that would be helpful. The counterfactuals, there are three of them, they're very nice. The first two are again are looking at what happens under various types of browser policies. What happens if we were to rule out the use of first-party cookies after seven days or blocking third-party cookies entirely.

And the third counterfactual looks at mitigation strategies for these types of privacy policies. So instead of using personal identifiers, can we probabilistically predict who the consumer is rather than identifying them explicitly? So the counterfactual results on the consumer side, the welfare side, they find fairly large effects here, 20 to 30% welfare losses under these various browsing policies. But the good news is that some extent these can be mitigated by this probabilistic prediction algorithm. On the seller side, we see that there's comparable reductions in terms of who is impacted on the smaller sellers. And this is a feature that we've seen in the literature elsewhere that privacy policies tend to disproportionately impact smaller sellers. That's a nice thing, would be I think nice to emphasize more. So just a couple of quick comments here. The counterfactual is this 20 to 30% seems a little bit on the large side to me, just intuitively I'm wondering whether it's an artifact of the data or the modeling.

And the probabilistic recognition algorithm. I find this very interesting idea, but to the extent that it's no silver bullet here because if we have a highly predictive algorithm, is this really privacy preserving? And so it begs the question, what would be an acceptable probability threshold for identification? So just to conclude, this is a very nice paper. There's plenty of meat on the bone here, both rigorous methods and novel data. Just in terms of improvements, it would be nice, I'd like to push a little bit more on some of the robustness of the effects and trying to streamline the narrative because a lot going on in this paper. But it's all very good work. Thank you.

Malika Korganbekova:

Thank you. We have five minutes of Q&A and then the break. Thank you so much Scott, for such a thoughtful discussion. I think there's a question.

Speaker 10:

Thanks, Malika. This is a great presentation. I was just curious, and maybe I missed this. Why would personalization in this setting necessarily be bad? And if whatever your answer is, is there heterogeneity in consumers where some people in some cases actually get hurt if it can potentially be bad? So I'm just curious on the other side of why that might be.

Malika Korganbekova:

Great question. Thank you. So typically a regulator standpoint is that personalization can be bad depending on what is the objective function behind the algorithm, what is it they're trying to optimize? And largely this is not auditable and they don't know what the algorithm does. So in that sense it is like an economic question of whether it is bad or not. In this particular setting, the objective function is to maximize conversion rates and it turns out to be good. At the same time, in terms of whether personalization can be bad or not, one of the maybe challenges working with this data is that it is very large, so I'm only concentrating on one product category, but I can look and that's why sometimes I just lack power in terms of understanding what has a different effect. So I think personalization can be bad for consumers for whom the personalization algorithm just simply does not have enough data to understand what is it they like and they might be showing something irrelevant.

Speaker 9:

Hi, I think if we think about this in terms of advertising, usually a platform would have to pay more to be able to serve personalized advertisements. So do you still think that the profit margins would be higher for the advertisers if they had to pay a higher price to get those targeted advertisements?

Malika Korganbekova:

So actually personalization is different from advertising in the sense that the sellers do not pay to be shown higher up on the ranking page results. And I think your question is basically whether sellers will have an incentive to advertise. And I think that is true. So during my sample period, the advertising was not as big, but now it is bigger. And I think in the long term it is a very important question that some sellers, when they see lower quantities sold, et cetera, they might try to advertise more and that would lead to more competition for the advertising slots, which potentially could increase the bid, the bids. And then again, it's not very clear whether, again, the larger sellers will have better positioning because now they are capable to pay higher bids versus smaller sellers cannot afford to pay for the advertising slots. Leon.

Leon:

Hi, really great presentation. Just following up on what was raised earlier, I'm really wondering what the downside of personalization here could be. And it struck me that you said something interesting in your response there. You said the algorithm maximizes conversion rate, what a peculiar objective, right? You would think normally the algorithm maximizes profits and one way of maximizing profits is of course showing less price-sensitive people the really, really pricey chairs, whether they be blue or white and less price-sensitive people, they get the cheap chairs. So do you have any insight on why that's not the case or do you know whether it's not the case or was this conjecture?

Malika Korganbekova:

So I know the algorithm because I can observe the code, like literal code that serves ranking pages. At the same time, Wayfair is not a monopolist, right? There is a very heavy competition on this market, which is why showing higher margin items will simply lead consumers to go to an outside option. So in that sense it is not necessarily good to show higher-margin items. At the same time, I agree with you, there are price-nonresponsive consumers who will have higher willingness to pay, they might be leaving some money on the table. That is true. But to the best of my knowledge, they are not at maximizing profit. They're maximizing the conversion rates. Thank you so much.

Speaker 8:

Thank you. We will be back at 10:35.

Nellie Lew:

I wanted to introduce myself. I'm Nellie Lew. I'm the Assistant Director of Consumer Protection within the Bureau of Economics at FTC. And I'm excited to chair this lightning round session called Regulations in Consumer Protection. We have four speakers and each speaker will have eight minutes to present. And at the end of everyone's presentation, we will have a question and answer. And to start us off, we have Bernard Ganglmair.

Bernard Ganglmair:

All right. Thank you. My name is Bernard Ganglmair. I'm at the Center of European Economic Research in Mannheim, with a joint appointment at the University of Mannheim. This is joint work with Jacopo Gambato at the University of Vienna and Julia Kramer, who is at Erasmus University in Rotterdam.

I'm going to speak three times as fast as everyone else here to get my eight minutes, but I'm going to continue in the fashion we just had, talking about cookies, talking about consumers, deciding whether or not to share data, what to use with the data and firms' incentives or firms' intentions to use the data whichever way they intend to use them.

The focus of this paper though is not so much on that decision and the personalization itself, but it's on this link at the lower-left of what you see here in that screenshot, it's about how can users, how can consumers learn about how the data is used? Or even more so, how do firms communicate that information to consumers or regulators?

And they do so through privacy policies. And privacy policies are inherently bad and there are some death wishes of people who have worked with privacy policies, I hope that's figuratively and not literal.

What this paper is all about, it's about the GDPR and the GDPR's intentions or the GDPR's goal to enhance transparency and accountability by firms, and all firms, with the main goal of giving users accessible information on how firms use their data. And we focus in this paper on two vehicles in the GDPR, one is the disclosure requirement in Article 13-14 that prescribes what firms have to disclose, what data is collected by how and by whom.

And the second vehicle in this transparency principle is the readability requirement that states that these disclosures have to be made in a concise, transparent, intelligible, and easily accessible form using clear and plain language.

And now, we can already see that one of these seems very objective, it's a checklist, regulators can check off whether they disclose what's necessary to disclose. And the second requirement comes in a rather subjective or vague fashion, in a language that might be a little bit more difficult to enforce.

So the question we're asking in this paper, and I want to give us just some teasers about our results, is how do firms respond to the GDPR's transparency principle and that asymmetric enforceability of these two vehicles or these two requirements? And we also ask how does the stringency of enforcement of these rules affect firms compliance decisions?

So first, in the paper, and also, I think for the discussion here and the presentation, we want to think a little bit about what actually should we expect. And a very simple way of looking at this is, well, if we have a budget-constrained regulator who is tasked to enforce these rules, we would think that that regulator focuses on what's easy because he can't enforce everything, so focuses on the easy rules. So that's the objective disclosure requirement and it's not so much the vague or subjective readability requirement.

Rational firms will then respond by complying with more disclosure, but not so much with more readability. So this is the theory model really in a nutshell. In the paper, we formalize this and we have three predictions in the paper and I just want to summarize them quickly.

Here, the first one is, well, with the introduction of this transparency principle or the strengthening of that in the GDPR, we should expect to see longer privacy policies with more disclosures, but not necessarily more readable privacy policies.

Second, if a firm expects to be under special scrutiny or attention by the regulator, it's going to exhibit more of that readability compliance because it knows it's going to be under the focus of a regulator.

And third, given that if a regulator were not budget-constrained, that regulator would have the resources and capacity to enforce a lot more rules. If a firm knows that it faces a better funded regulator, it will also exhibit better readability compliance relative to disclosure compliance.

So we ask and answer this question using data, and for that, we've constructed a panel of privacy policies of German firms with one limitation or one restriction, we want to make sure we contract firms over time, so we have a filter where we require that every firm has at least one observation before the introduction of the GDPR, the enforcement of the GDPR, and one observation after. We end up with an unbalanced privacy policy panel at the quarterly level of roughly 600,000 firms or privacy policies posted by 75,000 German firms. Thanks to the Wayback Machine, it's back online. So if you're doing work using the Wayback Machine, you're good to go, I think it's been back online for a couple of days.

We know the firm identifier so we can complement our data with firm level information. We know something about the size. We know where the firm is active or the industry the firm's active in. We use UK enforcement data, and I'll mention in just a minute how we use that. And we hand-collect state government budgets or the budgets for the German state data protection authorities for one of the results I'll show here in a second.

How do we measure compliance? We have to measure firm's compliance on the disclosure side and the readability side. On the disclosure side, we use a relatively straightforward or standard approach from natural language processing. In the simplest form, we just look at the length of privacy policies. We do a little bit more, we want to look at the volume of disclosure, the number of words and paragraphs we would expect to disclose information that's required per Article 13- 14.

For the readability side, we dip into the linguists toolbox and we use what linguists have been developing in terms of readability scores or readability indices. In the paper, we use two, for the presentation today, I'll just use the German version of the Flesch Reading Ease score, it's a very popular readability score that measures how readable text is. We use it because there's regulatory history in the United States, it's been used to regulate the text of insurance contracts.

All right, so if we just look at the data, we see an enormous increase in disclosures, not really surprisingly. Disclosure increases by about 80% if we also control for a year fixed effects, firm fixed effects and some additional controls. So longer privacy policy is a lot more information is disclosed. Not surprisingly, there's relatively less on readability, so the readability hasn't improved much.

Does this mean now the readability compliance is dead? Our answer here is no. Firms still do respond to the readability requirement, but it depends on how much they're exposed to the rules, how much scrutiny they can expect, and what's the capacity of the regulator.

On the exposure side, very simple, we just look at firms that, before the GDPR, had relatively little readability in their privacy policy, they see more exposure to the GDPR. Those are also the firms where we actually see an improvement in the readability of privacy policies, whereas firms that had very readable privacy policies didn't really improve much.

Second, when firms expect to be targeted firms in industries where regulators pay specific attention, we also see better readability compliance by these firms. We don't know much about the enforcement on the German side because the irony, privacy, but we have relatively good information for the UK's privacy enforcer.

And the last result I'm going to show is we exploit a feature of Germany. Germany has 16 plus two state data protection authorities that regulate the firms in their jurisdictions. They all have their own budgets. They regulate the same rules, but we see different enforcers. So we exploit variation across states and variation over time to get back to this prediction that a better funded regulator would induce firms to exhibit better readability compliance. So again, the idea, state level regulator's budget, we use that variation for enforcement capacity. And we do find firms in better funded states exhibit better readability compliance.

All right, just to summarize, firms write a lot more in their privacy policies. Those privacy policies are, on average, as incomprehensive as before. But the rules are, to some degree, effective, those who were behind caught up, firms did respond to more stringent regulation. And regulators funding does matter. And I'll close here. Thanks a lot.

Qiyao Pu:

Okay. Hello everyone. My name is Qiyao Pu. So this is our joint work, me with Dr. Jagdip Singh and Dr. Robert Widing. And we're all from Case Western Reserve University. So today, I'm here to present my work, title is Notice and Choice Framework for Fostering Marketing Trustworthiness and Ensuring Customer Engagement.

So with the increasing use of customer data, regulators and firms are having a challenge to winning customer trust and try to engaging customer. So what they currently use as an interface to build customer trust and to foster customer engagement to communicate privacy issues is what's called notice and choice framework. And these are all the screenshots of real, different websites that varies in terms of this notice and choice framework that they use in communicating with firms' data practices.

So conceptually, this can be seen as notice and choice framework where, as we can see, that there is a notice section and there is a choice section. So as we can see, usually, the firms are using these interfaces, these pop-up windows as a compliance to the current regulations. But we are arguing that can these interfaces be a trust- building strategy that the firms can leverage to build trust with customers?

Therefore, we are conceptualizing, under the notice and choice framework, with two strategies. The first is the notice strategy and the second is the choice strategy. So how these strategies will work, we argue that they will work under the theory of information flow norms. So the theory of information flow norms indicated that the individuals, the customers, maintain established norms regarding what is acceptable or not acceptable considering the flow of personal information online or offline interfaces. And the marketers or the firms who respect that individual preference of their information flow norms will win customer trust and will have more customer engagement.

So conceptually, these are two strategies, but practically, we need to consider the operationalization of these strategies. As we can see in the first slides that they actually are operationalized in different ways, although they are all under the same framework, which is notice and choice. And in current study, we consider two operationalizations of notice strategy. The first one is the level of openness where we contrast detailed notice versus global notice, where detailed notice had a higher level of openness. And we argue that this operationalization will win customer trust compared to the global notice. And the second operationalization of notice strategy is the framing effect where we are contrasting the negative framing, where the firms tell the customers what they will not do with their data, with the positive framing where the firms tell the customers what they will do with their data. So that's the two operationalization of the notice strategy.

So in terms of the choice strategy, we considered the operationalization of the active choices versus no choice. And we argue that the active choice will gain customer trustworthiness. And building on that, we also argue that there's a dependency between notice and choice, where notice lay the foundation of choice and the choice will be no meaning without the notice. And we are arguing that this is a trust building strategy and this will work when people have privacy concerns.

So in terms of the experimental results, so in the first study, we actually found that detailed notice, especially for the negative framing and active choice, are building customer trust. And when considered together, as we can see, in the first graph, when we considered notice, we can see this framing effect in the left. And after we factor in the active choice under the negative framing notice, we can see that this active choice actually significantly increased customer trust.

And in study three, we employed a more holistic design of this notice and the choice framework where, as I just discussed, that there is a dependency between the notice and choice. So it takes that dependency into consideration in setting up the experimental design in study three.

However, counter to our expectation where we expected that the detailed notice pair with active choice will help with building trust, we didn't find that. What we find is that the detailed notice, which is a high level of openness, paired with no choice, will win customer trust, and especially under the people who have privacy risk concerns rather than they care more about personalization benefits. And we also find a mediating mechanism that the trust is kind of the link between this framework and the engagement behavior.

Therefore, we keep doing the experimental work, and we are updating with our experimental setup to be more realistic and updating our dependent measurement. And just some takeaways with the public policy in terms of our results, the first one is that, not only the principle matters, but also, the actual design, the operationalization of those principles matter very much. And although we didn't find that detailed notice paired with active choice promote customer trust, we need to educate the customer that this is very important to you, although, probably, you do not want to read them. So that's my presentation. Thank you very much.

Nellie Lew:

Next we have Itay.

Itay P. Fainmesser:

Okay, so thank you very much for accepting my paper. This is a theory paper, so I'm not going to have any data. I think that might be the only one today. So we'll talk about consumer profiling using information design. What I want you to have in mind is just the broad observations that there are rich data sets online and platforms can sort consumers and do that, sort consumers along all kinds of characteristics. We'll think of them as like bins or something like that. They can monitor and record data on purchases. I think even just from today's talk we kind of get the gist of that.

And then, there is a question, well, can they really learn the willingness to pay of consumers through that? Can they do this mapping between this bin and the willingness to pay? Which came up in some questions today already. There was a question, well, can they actually see who's sensitive to price, who's not, and price like that? So this would be a theoretical question of feasibility.

Okay. So the setup will be very, very simple. We have a platform that's long-lived. Think about marketplace platform that doesn't set their own prices, so the sellers are coming to the platforms and are pricing in this platform. Then every period, you'll have a short-lived seller. There will be one seller who tries to maximize profit in that period. And every period, there's going to be many consumers who are short-lived, they don't think about the implications of whether they buy and what that will do to pricing in the next periods or stuff like that.

And again, I just want to emphasize, the platform does not set prices here, the sellers will be the one setting prices. And we'll allow, here, the sellers to price discriminate in any way that they want. So whether they do that or not, we leave that question open, but they can here.

Okay, speaking about the consumers, we'll have a lot of them, so unit mass, index C will be those bins if you would like. That will be a type of consumers, all the observables that the platform sees about the consumer. Each type of consumer, each bin, we're going to assume that it has a willingness to pay, otherwise, there is nothing to learn. So they have a willingness to pay. VC, it's going to be drawn IID from some distribution that the platform, everybody know, they know what the demand is at this point, in general, they just don't know who is willing to pay what. And over time, consumers from the same bin that are coming in different period have the same willingness to pay. So there's this environment where, really, it's really beneficial to learn and it should be easier to learn because well, the values don't change over time.

And then, in terms of timeline, at the beginning of time, those values are being drawn. Thank you. The platform shares with the seller segmentation of consumers. It doesn't have to give the seller all the information. It can give the seller any information that it wants based on the information that it has. And then, the seller posts prices. Each consumer observes their own willingness to pay and the price and make a purchase decision.

Now, I'm going to do everything that I can in pictures, the statements, otherwise, we'll take the entire rest of the time. So if we look at the demand curve, and this is like the simplest demand curve that we can think from Econ 101, the main result will say that if you take the demand curve and set the uniform price, that the seller does not know anything about whose consumer has that valuation, that creates an important partition for us. Everything that all the consumers that have valuations that is marked in blue, the platform will never be able to learn their valuation apart from the fact that they are above that uniform price, that they buy in that uniform price. That means the platform will never be able to convince a seller to price above the uniform price in order to get that information.

All the consumers that are in the red area, so all the consumers that have a valuation that's below that uniform price, over time, the platform, through segmenting the market for the seller, and then, the seller setting profit-maximizing prices, the platform will be able to learn everything about them. So pick a level of accuracy, within that epsilon, they will know the valuation of these consumers.

Now, I'm going to give a little bit of intuition, and to get the intuition of why they can learn one and not the other, think about the platform that's kind of just an open source platform, just give the sellers all the information from all the previous periods and all the purchases of all the consumers.

What does the seller that comes first to the platform do? Well, there is no information at that point, they just price the uniform price. Well, then a seller comes in the next period, well now there is information, knows everybody that bought at uniform price will buy again for that uniform price, everybody that didn't will not buy for that. So for the segment above, they actually price the same. They price the same P1. There is no incentive to raise it because when they set it, it was conditional knowing that these are the ones that's going to buy. But for the one below, they want to price lower. So they're going to price some price that maximizes the profit there. Now, you can imagine what happens in the period three, the seller comes, has a partition to three parts. Well set P1 to the group above with a high valuation P2 for the ones with the medium valuations, the one with the low valuation, they set lower. And that continues. And they learn a lot about the bottom, very little about the top.

But my result, what they can do even better than that, they can learn below P1, they can learn everything, not like this big gap between P2 and P1. So if I want to think about that, well, what will be one way that the platform can do it? Let them first reveal all the information to seller, get some partition, and then, what you want is to try and divide each of the big segments. The top one, you can't divide because we just said if you give all the ones that you know are in the top segment to the seller, the seller is not pricing above P-hat 1. So unless we bring buyers from somewhere else, we can't really [inaudible 02:47:29] anything here. But for the segments below, what we're going to do is we're going to just want to divide it in way. Right? Now, the fact that we want to divide it, that's nice, but can we divide them and how will we do that? Well, the way that we do that, let's say that we want to divide the consumers between P-hat 2 and P-hat 3, so this group B. Well, if we give the seller only group B, we know that they're going to price P3. If we give them group A and group B together, we gave it to them before, we know they're going to price P-hat 2.

So what we do is we just give them, every time, some quantity of members of consumer from group A, mix them together with group B and tell the seller, "Well, that's the mix that you have." And with that mix, we can essentially shift the demand curve by adding this mass of consumers from the higher and divide each interval and get the seller to price all of those pricing. So this relates to literature that was an experimentation, but this is indirect because the platform cannot actually price directly, so they really have to convince the sellers to do that.

Okay. There are lots of implication and I'll be happy to talk about why this is interesting. The main reason I think that it's interesting to this audience is that, when they finish learning, because they don't know about the top, no matter what the platform optimization is, if they maximize consumer surplus because they want to attract consumers, if they maximize seller's revenues because they take a fee, pick whatever that they're maximizing, as long as they can't learn more than that, it's a Pareto improvement. So everybody benefits.

But if you have a platform that prices by itself, that sets the prices by itself, and they can learn also the top, then they can really get into an area that's really not a Pareto improvement and can be harmful for consumers. So that was the motivation to look at that particular learning. Anyway, thank you.

Ken Wilbur:

Hi, my name is Ken Wilbur from UC San Diego, co-author of work with Wayne Taylor at Southern Methodist, and Dan McCarthy-

Kenneth C. Wilbur:

... work with Wayne Taylor at Southern Methodist and Dan McCarty, who just moved to Maryland. And just want to start by saying thank you just to everyone and we're going to talk about sports betting. So in 2018, the US Supreme Court repealed a federal law that prohibited new state gambling regulations and after it repealed that law, most states have now legalized sports betting in some form. 36 states have legalized. These are the online ones, only 30 of those. And 14 states have not legalized, but it's fair to speculate some of them probably will. As you might imagine, when more states legalize sports gambling, we get more sports gambling revenue, and in particular, the share of online sports gambling revenue, according to the industry association, went from about 0% six years ago to about a quarter last year. And when you compare it to the time series of retail gambling revenue, it's not obviously cutting into that revenue stream much yet.

On the right-hand side there, you see the number of monthly calls to the National Sports... Sorry, the National Gambling helpline. And you see a similar uptick at about the same point, and it's hard to interpret normatively. It's probably bad if more people need help with problem gambling, but it's probably good if those who do need help reach out for help. So I encourage you to resist simplistic interpretations there, but it's at least motivation that we should be thinking about what's going on in the marketplace. Now, there's been a lot of research on gambling, most of it outside the US, much of it correlational. It's perhaps unethical to randomize consumers to gambling treatments. In Canada, they've come up with a very interesting public health message focused on prevention of problem gambling.

The advice is not to gamble more than 1% of your pre-tax monthly household income because when they look at surveys, they notice that the incidence of financial, relationship, psychological and health problems, the incidence of those types of problems are higher when people are spending over 1% of their income on gambling. And similar increases in risk have been found in Norway on the left and in the right from a UK bank, which observed digital gambling spend, income, and mortality of its customers. It found that those spending higher fractions of income on gambling were substantially more likely to die within the next few years than those who are not.

Now, all of that is correlational. What we are going to do is investigate a credit card analytics data set, which enables us to observe transfers between consumers and online gambling providers, which are primarily DraftKings and FanDuel. Highly concentrated market, increasingly so in recent years. We're going to select a balanced panel of about 234,000 people who use their cards frequently over a five-year period. And for those people, we'll observe all of their transfers to and from 42 different online gambling services. We also observe their income, which is inferred from direct deposit records, and we'll classify them into low, medium, and high thirds of their state's income distribution. Now, importantly, daily fantasy sports expenditures were available to pretty much all of these people even before their states legalized sports gambling. So we're observing increases in gambling activity rather than the beginning of gambling activity. Now, show you what this looks like on the left-hand side, those are all of the gamblers plotted by total transfers into gambling services. And then net transfers on the Y-axis and the 95% quantiles nearly overlay the axes. And so this is a story of whales and minnows. The large majority of industry revenue is coming from the people who are spending the most from the tails of the distribution from the whales. And when you break industry profitability by customer rank across those income tersiles, the gaps between the tersiles are surprisingly small. The high-income gamblers are more profitable for the industry, but not as much more profitable as you probably would've imagined. Now, when we look at the distribution of percentage of income spent on gambling in gambling months, that's on the left, the median US gambler barely met that 1% guidance in the months that they gambled.

However, that is only in gambling months. When we include all the zero gambling months and we look at changes in percentages of these gamblers over time, we see an increasing rate, but only around two to 4% of these 234,000 gamblers exceeded 1% of income. Now, we do a lot in the paper, but in my remaining two minutes I'm going to try to show you something relatively simple, but I think important. We're going to use those 14 states that have not legalized as controls to predict counterfactual outcomes within a group of nine states that legalized online sports betting and retail sports betting at the same time. And our question is going to be, how much did the legalization increase the fraction of gamblers spending more than 1%, and how does that co-vary with income?

The answer is that overall, when a state legalizes sports gambling through all channels at the same time, it lifts the fraction spending more than 1% from about 2% as a baseline of the control states up to about 5.4% after the treatment. Now, the covariance of that with income is substantial, where low income gamblers are experiencing an increase in irresponsible gambling, about double the increased experienced by high income gamblers. Now, we have to be careful, those income groups are not randomly assigned. Many other unobserved variables correlate with income, but still, we think this is an important pattern to be aware of.

So three takeaways. If you remember anything, one, because online gambling is digitized, we're able to observe it, which is new in the gambling world. Land-based gambling typically relies on cash to avoid allowing addicted gamblers to go into debt that they can't pay off. Two, we find a substantial increase, about 150% over baseline in irresponsible gambling after states legalize sports betting. And three, gambling taxes are primarily progressive because the majority of gambling is done by these whale customers. However, the incidence of irresponsible gambling, therefore exposure to gambling problems appears to be regressive. Low income gamblers appear to be much more likely to get themselves into trouble than high income gamblers. And there's a lot more in the paper. Thank you.

Speaker 11:

Thank you, Kenneth. Okay, next we'll open it up to questions.

Speaker 12:

So maybe I'll start with the gambling paper. You mentioned at the beginning that the fantasy sports in particular are very highly concentrated with FanDuel and so on as two major players. Would that actually be beneficial for cutting off the irresponsible gambling if they have a lot of data over the customers because it's harder to switch to the lots of range players?

Kenneth C. Wilbur:

So to a substantial extent, what we see in the marketplace right now is the major platforms enabling consumers to limit the times, numbers and sizes of the bets that they place. In other countries, platforms have either been compelled or voluntarily enabled further controls than those. And the state regulators that are tasked with regulating operator behavior typically to this point in the US, as far as I'm aware of, have not required more serious consumer self-constraining technology than what the platforms are currently offering. And there are probably a lot more things that they could offer if it were demanded.

Speaker 13:

Ken, I wanted to follow up on this. As I saw the aggregate data with the pandemic being a low point in online gambling, largely because of supply, I guess, I was struck by an anecdote that came up about someone betting on Korean baseball games who had no interest in Korean baseball, but obviously wanted to gamble. Is there a way to use the longitudinal data as a natural experiment that would identify people probabilistically who are addicts and use that in lieu of a database of gambling addicts to assess characteristics, patterns of behavior, and other things that might allow for identification?

Kenneth C. Wilbur:

That's a great idea. We haven't done that. That's a really good idea. Thank you.

Speaker 14:

So my comment is about the first paper we saw about enforceability of GDPR. The thing that comes to mind when I saw that is in the deterrence literature, the idea of a random crackdown where there's a small chance that we may just do a sweep and test for readability of these terms and conditions and the mere threat of just randomly just ramping things up for a month and doing a sweep of a random subsample of websites, it can be a pretty powerful deterrent. And I wonder if that's sort of comports with your understanding of, not necessarily what they're doing, but what might be desirable?

Bernard:

I'll pass it on to the state data protection regulators. My understanding of what they're focusing on, given the continuing complaints from them also to the public, that they're completely underfunded, that readability is not something... The regulation of the readability requirement is currently not something they have on their radar. Something that you described might indeed work. And given that it sounds or it seems very low cost or a very low budget experiment, it's something that could try either in the domain of readability or in other domains. My understanding, again, low budget, you're trying to find ways of regulating within their means. Thank you.

Jagdeep Singh:

My name is Jagdeep Singh and I'm from Case Western. I have a question for Bernard. This is about readability again. I wondered if you checked, what do regulators do? Do they have some practices that they enforced in order to check for readability, and if those standards or practices vary by different regulators across the EU?

Bernard:

That's a very good question, because how do you measure readability? We take a stance. We use some readability scores and we pick two because one has some historic best practice evidence for it. And the other one that we use and we describe that in the paper is because it best explains text comparisons. The best I can say to that is the European Data Protection Board, or its predecessor, alluded to readability testing. So one reason we use readability scores for those papers is because the European Data Protection Board suggested, "Oh, if you try and enforce that, use something similar to the readability testing."

Whether or not that's being picked up by the actual... Across the EU and within Germany by the regulators, I don't know. I have my doubts. We've been talking to NGOs in Germany that sort of keep an eye on what the German state data protection regulators are doing. One message we get from them is they're completely understaffed and underfunded and don't have the IT to try and do something. There's even been proposals going in stating, "Well, why don't one of the state data protections in Germany do that? It just takes some IT infrastructure and run 200 different readability scores and then check what the firms are doing in that." But the federalist structure of regulation Germany apparently doesn't allow for that.

Jagdeep Singh:

I have a follow-up. Does any of the regulators actually disclose what they will use for readability testing?

Bernard:

I don't know the answer to that. We should probably ask them what we do know, the limited information we have about what they actually do, it's the irony, privacy, is information we get from the GDPR tracker. That's a website that tries to track cases in Europe. There's very, very little on readability or an Article 12. I think we've found a handful of these cases, but we don't know what they actually used and how they try to enforce that.

Speaker 15:

This is a question for Dr. Fainmesser. It seems like the crucial bit about not being able to discriminate above the uniform price is that the platform isn't sending prices themselves. Is that the only thing that drives that or is that more subtle than that?

Itay Fainmesser:

So that's definitely the main thing that drives that. The other thing is kind of the idea that the other players are small enough that they don't think that their learning will change everything. If there was a seller on the platform that was big enough and was selling all the time on the platform, then they might want to explore by themselves. But that kind of turns them into the platform in some ways.

Speaker 15:

And then the follow-up is, there is a literature on how different players in the market, the platform and the sellers have different preferences about how much information is shared. And there's some question about trying to make the segments thicker to change the competition across segments. Did you mess around with any of the kind of information sharing with the sellers or was it just sort of a blanket it all gets shared?

Itay Fainmesser:

No, so if I understand correctly, the platform shares completely different information. So in order to implement that, they share different segmentations over time to the sellers. And then at the end, the part that they didn't present, once they got all the information that they can get, then starts the question, "Well, what do they want to do with that?" And then we have in the paper, what they didn't have time to present is characterizing the all set of market outcomes they can implement by sharing different information with the sellers. So yeah, there's a lot more on that.

Speaker 16:

So this question is also for Bernard. To sort of follow up, I'm wondering if firms in Germany maybe don't even understand what readability means. So there's a difference between not knowing what the standard should be and compliance, right? So I'm wondering if you see more variability in the readability of disclosures as opposed to the listing of, "Oh, here's the data we collect," amongst firms as a way to sort of show that.

Bernard:

To your first point, yes, I do think that lots of firms don't pay attention to that or don't know how to measure it or don't know how the enforcers or the regulators would try and measure it. Because we have at our disposal a couple of hundred of these readability scores and it depends on... So our results our... The average results are very sensitive to which score one actually uses. Some scores say readability has improved. Those scores suggest readability has gotten worse. Majority of scores suggest it's gotten worse, but there is quite a variation also in the scores. As to do we see a higher variance in the readability, yes, we do compared to disclosure. And I think that not really knowing how to implement it might explain the average effects that we have. The readability is a little bit all over the place, but it didn't really change that much.

And if anything, it's gotten down a little bit, but some firms still do pay attention and some firms still try and figure it out. Those that had very low readability seem to have paid a lot more attention and improve theirs. And those who were really good decided not to mess with that. And on average, with more disclosures, readability mechanically goes down. There's this thing called a transparency paradox. They disclosed more, that means readability kind of mechanically has to come down. So some firms did pay attention and try to figure out how to best implement the rules.

Speaker 11:

Oh, there's one more. Okay.

Speaker 17:

Question for Itay. The platform in your model seems to have very limited information, just whether or not consumers purchased. Of course, one of the concerns about privacy is that platforms are collecting a lot more information than that. So have you thought about what would happen when the platform has better information? Would it just result in finer segmentation or would something different happen?

Itay Fainmesser:

So I actually think in this model, the platform had tons of information. So first, I'm saying that hidden within the C, that means that they know the set of all individuals that have the exact same value. So that means they have really, really, really fine characteristics. They don't just know my age and my geographic. They have to know really, really, really a lot of me so they can pair me with twins all over the world that have the same valuation. So basically, all the information that's not pricing information I give them for free. I'm just saying, suppose a world that they have all that.

With respect to pricing, they don't just see what consumers purchase, they see what price each consumer is presented with and whether they purchased or not. So in that sense, I know you could think of more information, but it's quite a lot of information. What they don't have here is information outside of the platform. And that one depends on how other platforms are run and what information other platform. So that limitation is in terms of where one could give them more, but we actually intended to do it the other way. We intended to say, "Give them a lot, create an environment that's easy to learn." And then still they can't learn everything. That was at least a theoretical approach.

Speaker 11:

All right, well, let's give this panel another round of applause. Thank you. We're going to take a very short nine-minute break and try to come back at 11:40. Thanks, everyone.

Devesh Raval:

... everybody. I'd like to start. So I'm the chair of this session. But before we get to the paper talks, I'd like to talk a little bit about what's new in consumer protection of the FTC. And the normal disclaimer applies. These are my views and not the views of the FTC or any of these commissioners.

So I want to start with the kinds of tools that we have to protect consumers. So I'm going to put these into four main tools. The first is law enforcement, so that's suing people with our current laws. And the second is rulemaking, which is writing regulations that out of the force of laws, so essentially new laws.

The third is research, which is what most of you guys do in your day job, but we both have some special powers as well as maybe special interests. And then the last is consumer education. And so I'm going to hope to show you in the next 15 minutes some examples of all of these.

But the biggest, I think, change has been the second one, which is rulemaking. So there's been, I think, an unprecedented push of rulemaking in the past couple of years, at least unprecedented since the '70s, but I think it probably beats the late '70s.

So this is something, just to give you a scope of how this works, the lawyers are going to write a rule. They're first going to give a proposed rule. We have to put out a preliminary cost benefit analysis trying to both explain and quantify to extent possible both the benefits and the cost of the rule. Then it's going to be issued out. People like you can give us comments about you like, what you don't like, and then we issue a final rule and a final cost benefit analysis.

So this is something that we didn't have much experience in before a couple of years ago. So the spirit of what's new at the FTC, we're very happy to have Nellie Lew, who's sitting over here who joined us as assistant director. She has a lot of experience in rulemaking, both at the FDA where she was acting chief economist and then at the FAA.

So I'm going to give you a whirlwind tour of all the rulemakings that we're doing. So I'm going to put these into three different buckets. The first bucket is information disclosure, so rules that are intended to improve the kind of information that's in the marketplace. So the first one here is the CARS rule. So this is about auto dealerships. So Mike LeGower, who's sitting up here, was a lead economist and then manager on this rule. And he's going to give an extensive discussion of both the benefits and costs that we try to calculate in the rule as part of the cost benefit analysis.

Second is the unfair and deceptive fees or junk fees rule. So this is about partitioned in dripped pricing, trying to remove partitioned in dripped pricing. So for example, resort fees for hotels or the Ticketmaster service charge.

Now some of these, like the negative option rule is really a rule review. So this is an old rule. But every 10 years we do reviews to try to update them given changes in the economy. So this just came out on Wednesday. Unofficially, it's the click to cancel rule because the goal is to make it just as easy to cancel as it was to sign up for negative options or subscriptions, and it's also aimed to produce misrepresentations on subscriptions.

And then finally here we have the funeral rule, which again is an old rule that required funeral homes to have price lists and give you price lists if you ask for them. And now we're updating that potentially to make those price lists be online as well if you have a website.

So next is privacy and data security. So here, by far the biggest rule is I think the commercial surveillance rule. So this is about potential privacy and data security regulations. Unfortunately, I can't tell you anything about these except that Dan Wood, who's somewhere in the audience, is managing it, and several of the economists in the room are working on it.

We also have the COPPA rule that just went over a review. This is the subject of Garrett's talk about children's online privacy. And then the health breach notification rule, which was about notifying people when there's a breach of their health information. And we recently updated this to also extend the health apps on your phone.

Lastly, there are a bunch of rule makings that are about fraud. So this is the impersonation rule, making it illegal to impersonate a government or a business. So a lot of people trying to impersonate FTC officials as part of their scams, for example.

Second is the reviews rule. So this is going to make it illegal to have fake reviews to suppress real negative reviews to buy and sell reviews and so on. And there's also some stuff about influencers and things like that with endorsements.

And then there are two here. One is a new rule, new potential rule. One is a rule review, but these are about business opportunities lying about the kinds of earnings you can get. So again, trying to reduce the kind of fraud in the marketplace where someone comes to you with this great business opportunity that doesn't turn out to be there.

So that's the rule makings. So the second thing I would want to highlight is digital economy. So if you look at what the FTC was doing in the 1980s, a lot of it was goods in the physical economy. So deceptive claims about washing machines or healthy claims on cereal. If you look at what we're doing now, we're really being focused on the digital economy and trying to be the regulator for the US for the digital economy.

So I want to give a few examples here of recent things that I think highlight some of the problems that we have. So the first is a recent case against Amazon over cancellation flows and Prime subscriptions. So for consumers that were going on Amazon, maybe you put something in your cart and when you click to check out, you see this screen here. So if you click the yellow button, you can get free two-day shipping. And then light blue on the left-hand side, it's like, "No thanks, I don't like free things."

So you're maybe wondering what's wrong with this. So thankfully Amazon wrote a memo in 2018 detailing all the problems with this. So I'm going to give a couple of them. So first of all, it said the button to enroll in Prime doesn't make it clear that consumers are signing up for Prime. It also doesn't tell you what the price is, the fact that this is an auto-renew subscription. And they have several other issues.

So you may wonder why is Amazon writing documents about the problems with their own service. So this turns out to be a very interesting fight between one group and Amazon that was trying to improve its cancellation flows and its subscriptions for Prime, and did a lot of work documenting the problems and potentially showing the share of consumers that didn't know they were on Prime and how that could be improved. But there was another group on Amazon that was the Prime group that was incentivized to have lots of Prime subscribers. So these conflicted. When it was escalated up to the top, management sided with the Prime group and didn't do anything. But the losers wrote a memo and they said, "There's some consumer watch dogs. They're going to have problems with these dark patterns and manipulation." And we ended up investigating and then suing Amazon for this.

But the reason I wanted to bring this up is that a lot of digital economies are now the size of economies and are huge bureaucracies, we know very little about platform governance and how to improve platform governance and how potentially our enforcement tools affect platform governance.

So the second two pieces I want to talk about are two of the pieces of maybe the implicit price of digital services. The first, this is our Kochava case on privacy and data security. I would guess most of you have not heard of Kochava, but Kochava has heard of you, or at least it's probably you are probably in their massive geolocation data set. So this company, his business model was selling geolocation data.

So you're maybe wondering what are the issues with selling geolocation data. So first of all, it can tell where you live, so then a household mapping service to figure out where you live. Once it knows where you live, this could be mapped to offline data on addresses that data brokers also sell. They could use this to market to people with sensitive characteristics. So for example, based on the apps on your phone, they can try to detect whether or not you're an expecting parent or not. They could also market based off the sensitive locations you go to. So for example, they advertise during Covid. We can tell if somebody went to Covid testing site, we can tell whether they went to the hospital and so on. If you look at the FTC's complaint, they single out abortion clinics as well as places of worship as potential problems. Maybe we won't want anyone to figure out that you're doing those things.

And then finally, Commissioner Holyoak, in a concurring opinion, said, "This is maybe a way around the fourth amendment." So normally, if you're a law enforcement, you'd have to go to a judge with a warrant and with evidence to get this kind of data, but now anyone can buy it. And that's really reducing our freedoms as Americans. So this is maybe one part of the implicit price of digital services.

The second one is fraud. So this is coming from a data spotlight by Emma Fletcher in the Bureau of consumer protection. And so this is part of our consumer education to try to warn consumers about things that happen online. But this is data on reported fraud losses from our consumer sentinel complaint data over two and a half years. And what Emma documents is that by far the largest amount of fraud losses is coming from people saying that they were originally contacted on social media. So a lot of this is things like fraudulent advertisements on Facebook and Instagram. It's very related, I think, to Alessandro's talk earlier today.

So I guess in two and a half years, $2.7 billion were lost coming from this kind of social media contact. And so we're now doing a six piece study on social media ad fraud. So we have special authorities under the FTC Act to try to get information and data from companies. And so here we're looking at Meta, TikTok, YouTube and so on to try to learn what they're doing about social media. Now in the final three or four minutes, I want to talk about the intersection between competition and consumer protection. So historically, we have a dual mandate with both unfair measures of competition, so antitrust policy as well as unfair and deceptive action practices. So this consumer protection policy. But for most of the time I've been here, these have been siloed. So one group of lawyers is competition, the other group of lawyers is consumer protection. That's now been changing. So Aviv talked a lot about this, and he is always reminding me about how our consumer protection rules are really competition rules.

But anyway, I'll give a couple of other examples. So first is our recent case against the PBMs over insulin, over unfairly-preferencing high wholesale cost insulin on formularies. This is a very complicated case, so I don't want to try to explain it, except that this is a case coming out of the Bureau of competition, but two of the three counts are from our consumer protection authority.

And similarly, if you look at consumer protection cases, we are thinking about whether or not we want to include competition counts if the facts would warrant them. So that's one example.

I think just more generally, often competition and consumer protection are looking at the same big digital platforms. And you may want to think holistically about both of our authorities instead of sort of siloed of competition thinks about competition and consumer protection thinks about consumer protection.

So I have two or three minutes left, so I want to just highlight what the panelists are going to talk about. So first we have Mike LeGower, who's going to talk about our cost benefit analysis on the CARS rule. Then we have Eric Spurlino and Shitaro Nakamura. So they're both fresh PhDs who came aboard a year ago from grad school. One is looking at a lab experiment, one on a field experiment about information disclosure and the effects of information disclosure.

The next two are about the intersection between competition and consumer protection. So I think Dan is maybe the opposite. He's a wealth of knowledge and experience at the FTC in particular on merger retrospectives. So if you think about what we know about mergers, a lot of it's coming from exposed evaluations of what happened to those mergers. If you look at the data, I think the FTC has a 22% market share by my calculation of all the merger retrospectives that have been done. And Dan has contributed a lot to that. And today he's going to talk about extending that using data and reviews, which is what marketers like you have thought about.

Lastly in the Hawaiian cat shirt is Ben Kasner, who is a theorist in our antitrust group, and he is going to be talking about digital curation on app platforms and some of the potential tensions between our competition consumer protection missions. So with that, I'll give it to Mike.

Mike LeGower:

Okay. First I want to thank everybody for coming out. I want to thank the organizers for inviting me up here to talk about our regulatory analysis in the CARS rule. So the CARS rule is officially the Combating Auto Retail Scams rule. As you might expect, it is a regulation that covers motor vehicle dealers, specifically automobile dealers. It was first proposed in 2022. And then it was finalized in 2023 with the final publication in the Federal Register happening early in 2024.

Now, it has since been challenged in court by the National Automobile Dealers Association, and so the enforcement of the rule has been stayed apropos of nothing. I'm going to be following very closely to my notes here so that I don't say something incorrect about it that the NADA can then use in litigation.

So the rule has three substantive sections with provisions that prohibit certain misrepresentations about automobile transactions. It requires certain disclosures including the disclosure of a defined offering price. It's defined in the language of the statute. So it's by definition comparable across dealerships. It requires disclosures that add-ons are not required for a vehicle purchase if that is in fact true. And then it has requirements that the total of any payments for financing or leasing transactions be disclosed at the time that you're talking about such a transaction and a disclosure that whenever a comparison between monthly payments is made, the dealer also has to acknowledge which of the two options has the higher of total payments.

The third substantive section has a prohibition on charging for certain add-ons. So this includes add-ons that provide no benefit to the consumer. And it also has a broader prohibition on charging consumers for any products or services without their express informed consent. And in addition to that, it imposes some record-keeping requirements on the firms. For more details, you can look at the rulemaking docket for the rule. It has the whole text of the rule in there.

So the Bureau of Economics was tasked with the preliminary and final regulatory analyses that were required by the statutes that authorized the rule. We identified two main sources of quantifiable benefits that would result from the rule. The first of these is time savings or in the jargon of economics, a reduction in search costs. And then the second would be a reduction in deadweight loss from an expectation that some of the provisions of the rule will lead to lower prices.

For consumers, the costs were broken down by provision. So different provisions have different compliance tasks associated with them. They're primarily going to reflect the resources that the dealers have to expend in these compliance tasks. So they're primarily going to be labor costs, but there are some nominated costs. We also note in the regulatory analysis one major transfer that results from the price decrease that I just talked about with the reduction in deadweight loss. So that price decrease is also going to come along with a transfer from dealers to consumers. There are other benefits and costs that we note in the regulatory analysis, but we were unable to quantify.

So before I get into the substance, I just want to note that this was a huge effort from individuals across the agency. From the conception to the finalization of this rule and now beyond as it goes into litigation, it probably used hundreds of hours of economist time to sort of button down all the regulatory analysis involved. It involved staff scouring publicly available data and trade publications for data to apply to the questions that we had about the impacts of the rule. We also had to read academic articles and other agencies publications to find theory that we could apply to the regulatory analysis.

And then finally, we had a lot of cross-pollination from attorneys to the economic analysis as they refined what the different provisions would require, and then from the economists back to the attorneys and developing the regulatory text.

So first I'm going to talk about the benefits of the rule. I'm going to talk in detail about how we thought about the reduced search costs and then the reduced deadweight loss. I'm going to talk briefly about costs and give you just an example of how we thought about costs for the rule. I'm going to discuss the sensitivity analysis that we conducted to sort of stress test our recommendations. And then I'm going to conclude briefly.

To start, we thought about time savings from reduced search that we expected would result from the rule. We know from trade publications that consumers spend an average of 15 or so hours in the process of completing an auto transaction. And we expected that the greater transparency that results from the provisions of the rule will eliminate some time spent pursuing misleading offers. So the key idea here that we use to develop our analysis was that digital consumers, and these are consumers who do a lot of the tasks associated with shopping for a car over email or online, spend less time in the process than non- digital consumers.

So here you can see on the chart that, for example, a consumer who's negotiating the purchase price of a vehicle over email or online spends about 43 minutes less at the dealership of purchase than a non-digital consumer. And so some features of our rule are expected to emulate some of those features of shopping digitally, right? The fact that a digital consumer can maybe get a price quote in writing from a dealership over email and then take it to another dealership is expected to sort of reduce time negotiating the price of the vehicle. The required disclosures in the rule include comparable prices up front, right? So we expect that that's going to emulate some of this digital price negotiation function.

So the main approach here was to attribute a portion of the time savings that these consumers who do tasks digitally under the status quo receive to the non-digital consumers once the rule is in effect. So here I've listed all of the replacement rates that we use in the base case of our analysis. We expect the price negotiation piece to be a big part of it, selecting add-ons and doing paperwork we're going to attribute slightly time, savings to. And then we don't expect the rule to operate as much on getting a trade-in offer. So we in our base case, don't allow for any sort of time savings from the trade-in.

In addition to this, since the data that we have on time savings is from the dealership of purchase, we're going to account for time spent at non-purchase dealerships as well, since that's also part of the shopping process.

And so here you'll see the resulting table. Our base case estimates correspond to a roughly two-hour per transaction reduction in search costs. On a base of 34 million covered transactions per year at a value of time of about $24 per hour, that results in a benefit over the 10-year analysis window of between 12 and $15 billion in present value terms.

The main text of the analysis also walks through alternative assumptions for the time savings replacement rates that I showed you on the previous slide. The upshot of that is that our low-end numbers roughly cut the savings in half. But importantly, if you hold the costs of the rule constant, there's still net benefits to the rule. The high-end savings estimates increase time savings by about 50%, and so obviously is more beneficial. Next, I'm going to talk about the reduction in deadweight loss. So because of the search frictions involved in the automobile market and the shrouded prices, dealers are able to mark up products higher than they would under the rule. This idea is supported by models and structural empirical research that shows that allowing consumers to observe prices before they engage in costly search reduces prices, and in the auto context specifically reduces dealer profits.

So here on the slide, I have just a simple graph that illustrates the idea. The reduction in price that results from this increased price competition is going to lead to a quantity expansion. And then as a result, a reduction in deadweight loss.

Now we're going to need a more complicated model than just the graph on the slide to deal with this because of the used car market and how it interacts with the new car market. So our approach is going to be to follow the sufficient statistics approach in Henry Kleven's 2021 paper to sort of calculate the welfare effects when you have related goods.

So after a little bit of algebra from the Kleven paper, we find this equation that gives you the welfare change that results from the rule. This equation has seven parameters that we need to calibrate from the literature. So there's two baseline markups. The tau is actually baseline markups in this setting. There's the aggregate cost of all new and used vehicles sold. There's the change in that results from the policy. And then there's the elasticities that give the responsiveness of quantities to the price changes.

So we're going to get elasticities from an EPA report by Jacobsen et al that has a model of the new and used vehicle markets that incorporates scrappage of vehicles and was designed to answer this specific question. When the price of new vehicles changes, how do the quantity of new and used vehicles respond? We're going to get baseline markups for new vehicles from a Grieco et al working paper. They model the US new car industry and find that their most recent estimate of industry markups is about 15%. We don't have a similar source for used car markups, so we're going to assume that it's either 0% or symmetric with the new car markups. Aggregate costs are going to come from the national transportation statistics tables. So we have the quantity of new cars sold, we have the average prices of new cars sold. That combined with our markup data is going to give us the costs, the aggregate costs of those cars.

And then finally, the crucial piece is the change in price. This comes from a Murry Zhou paper in 2020. They simulate a full information counterfactual in the Ohio Auto dealer market that essentially eliminates all distance- based surstritions. And they find that prices for new cars drop by about $333 as a result. That's about 1% of the average car transaction.

We're going to plug those values into our equation here, and we get about $152 million per year reduction in deadweight losses. Over 10 years, the present value of that is just over a billion dollars. As with the time savings, the main text explores alternative assumptions. And we find that in all of those situations, the net benefits of the rule are positive.

So I'm going to move on to our costs now. I'm going to talk about this very briefly. Mainly, our costs deal with the value of time spent in compliance tasks for automobile dealers. It's mostly going to be labor hours. Some of them happen one time when the rule is passed. Some of them happen every year. And some of them happen per transaction. The quantity of hours spent was determined by the judgment of how long various compliance tests would take and referencing similar previous rules that had similar provisions.

The cost per unit of time was given by occupational wage estimates from the Bureau of Labor Statistics by SOC code for the various occupations involved in the compliance task. There are some non-labor costs, mainly optional paper or e-disclosures that the dealers might choose to provide, or storage space for record keeping that was imposed by the rule.

So an example of a typical cost table from the regulatory analysis. Here we have the compliance costs for financing cost disclosures. There's an upfront cost components, which is basically the compliance managers at the dealerships having to draft the language for the disclosure. And then there's ongoing costs where they have to train the sales staff to deliver the disclosure appropriately, right? And so we find that adding all of those up, we get between... What does it say there? 98,000 and $116,000 million. So that's just an example of a typical cost table.

And finally, I'm going to talk about our sensitivity analysis that we conducted in our regulatory analysis. Now, we don't have a crystal ball. A lot of our assumptions, some of our assumptions might be incorrect. So we're going to characterize the uncertainty around our base case estimates and stress test our recommendations a couple of different ways.

The first I've already talked about. In some of the critical assumptions that we made in the benefits analysis, we're going to, in the main body, look at how it responds to alternative assumptions. But in addition to that, we're going to do two sort of comprehensive sensitivity analyses in a dedicated appendix. The first is we might be making correlated errors with the labor costs, right? We might just misunderstand something about how auto dealers operate. So here we're going to say, "What if we're wrong by a factor of 10 on the labor costs," right? Everything takes 10 times longer than we say it's going to. Importantly, holding the benefits side constant, we still find net benefits from the rule even if you multiply the labor costs by 10.

The next thing I'm going to talk about is our Monte Carlo simulation. So here we take all of the parameters from our analysis that aren't based on hard data. And we're going to allow them to vary mostly symmetrically via a triangular kernel around our base case assumption. Some of our estimates are like our proportions there. We're going to let them vary uniformly because it didn't make sense to do a symmetric distribution.

Obviously, this assumes that our underlying model is all correct. But assuming that our underlying model is correct, we're going to make these draws independent from all of these distributions and simulate the net benefit outcome under each draw. So what you see here on the slide is that the net benefit distributions that result are roughly symmetric. But the important part is that every single simulated outcome provides positive net benefits for consumers.

So just to conclude, our regulatory analysis used work from inside and outside the bureau. We combined theory and data to generate defensible estimates of the benefits of the rule. We made reasonable assumptions about labor costs and stress tested those assumptions via sensitivity analysis, and ultimately concluded that this proposed rule is highly likely to increase total wealth.

Devesh Raval:

[inaudible 03:56:55] Eric Spurlino who's going to start the paper session.

Eric Spurlino:

Okay. So thanks everybody for coming to the session. I'm writing a paper called Clear Disclosures with some academic co-authors, Stefan, Andrew and Ryan. We all have eight minutes here today, so I'll just say it for everybody. The views expressed in this presentation and the other three presentations are those of the author and do not reflect the commission, the usual disclosure here.

So what's the motivation behind this paper? So as we've talked about earlier today, I think Bernard had a paper that is highly relevant to this. Mandated disclosures are kind of ubiquitous regulatory tool that people use to try to reduce information asymmetry between consumers and firms. But there's been a very big critique in both the legal and the marketing literatures that these disclosures are often complex to the point of being essentially unreadable or unusable. Ben-Shahar and Schneider have a big paper called The Failure of Mandated Disclosure that gets into this a lot.

So a key question-

Speaker 18:

Disclosure that gets into this a lot. So, a key question, then, is, obviously, how do we make these disclosures simpler? And it turns out that, before we even answer that, as was alluded to in Bernard's presentation, we have to figure out, how are we going to measure complexity disclosures, in the first place? So, what we're going to do here is we're going to use some of the theory that my co-authors have made over the last decade about complexity and how economists model complexity in the lab, and we're going to try to bring them into the real world. So, we're going to introduce a theoretically motivated experimental test to measure complexity of a given disclosure format. We're going to focus on the privacy policy setting, just because it's very relevant to the work we do here. But there's nothing particular about privacy that makes this method work.

So, just a really, really quick overview of what the theory looks like here. So, we're doing a really simplified model of what a consumer's decision problem is. A disclosure is just going to be d, and it's going to convey some information theta. And we can think about this vector as just being a vector of different things that a consumer might care about. For example, one of the thetas could just be a variable that's true if the firm is going to collect information on your gender. We can just think about disclosure as modeling a whole bunch of this information.

We're only going to be examining disclosures that contain the same amount of information. So, what we're not interested in is saying, should there be more information or less information. What we're talking about here is, how do we frame this information? After reading the disclosure d, the consumer isn't going to make a binary action a equals zero or one. We can just think about this as being accept or reject for the privacy policy.

So, for the consumer's problem, we're going to suppose the consumer gets some utility delta for accepting the privacy policy and utility zero for rejecting the privacy policy. It turns out that complexity or cost of information of reading this disclosure can be revealed by looking at what we're calling the cognitive economics curve, which looks like this, p of delta. So, this is just the probability you're accepting a given privacy policy, conditional on the utility of accepting that privacy policy. It's exactly analogous to something that's 200 years old in the psychometrics literature.

So, the idea here is just that things on the right of this axis are going to be things you should accept and things on the left side of the axis are going to be privacy policies you should reject, and then how far away you are from that vertical axis is essentially how much you care about making the right decision here. And the idea is, if you don't really care about making the right decision, maybe you don't care about privacy that much, you should just accept with probably 50% and not learn that much. As I start caring a lot about making the right decision, I should accept good options more often and I should reject bad options more often. That's all it's really saying. Oops, I just pressed a, there we go. Wrong button. So, it turns out that you can use these to rank disclosure formats. So, we have a very simple way to say that one disclosure generates this cognitive economic curve, another disclosure format will generate another cognitive economic curve. And essentially, a cognitive economic, or a disclosure is more clear than another disclosure if you are accepting good policies more and you're rejecting bad policies more. That's it. So, we can just graph these and we can rank disclosure clarity. I won't get into all this detail, there's a lot of math behind this, but you can actually recover the cost of information from these curves, as well as functions of the area underneath these curves or the area above these curves. This is going to allow us to do things that, maybe, readability scores can't do, where we're actually going to be able to say, okay, what is the actual cost of this information? Which is obviously something regulators carry a good deal about.

So, there's a lot more to this that I would say in more than eight minutes, but the idea is that these curves give us a lot to work with in terms of analyzing welfare and analyzing how clear disclosure format might be. The question, then, is how do we get these cognitive economic curves? These aren't just something that we can observe super easily in natural data. So, we're all experimentalists on this paper. So, we're going to argue that you can use simple economics experiments to elicit these curves directly.

So, what we do here, again, we're sticking with a privacy policy setting. We choose 20 relevant brand to binary privacy policy variables. You can see them in the small print at the bottom that could be used in a privacy policy. We, then, used a large language model to read actual HTML codes of a lot of real-life privacy policies, and we use that to generate a Python code that takes the vector of these 20 binary privacy binary variables and generates a HTML code that looks like a real privacy policy.

So, we have two different versions of this. Again, this is just a methodological paper, so we're not saying that this is the privacy policy we should follow or not. But the first one we do is just a control. We made Claude make it very, very ugly. They're about 1,300 words. They present them in random order. There's a lot of legalese. It's really annoying to read. We then ask Claude to make it simpler and make it more user-friendly, just this is an obvious first step to do. And yeah, it's 250 words instead of 1,300 words. There's a table of contents, there's links, there's color coding, it's very pretty.

Then, what subjects do is they are choosing a hypothetical accept or reject decision for each disclosure. We're not interested in doing the work that all of you are doing, trying to understand how much people care about privacy. And instead, we're going to experimentally endow them with incentives. So, we're told that we're telling subjects that they care about five randomly selected red flags, and if there are more than two of those flags, they should reject the policy, and if there's two or less of these flags, they should accept the policy. And they get some prize z for making the right decision, prize zero for making the wrong decision. Or if they just reject, they get something in the middle. You should accept good policies, reject bad policies.

They play 30 rounds where we vary z from zero to 10 probability points for a cash prize. The type of the disclosure is going to stay constant. We vary that between subjects. And each red flag is a 50% chance of being correct.

So, then, we're actually able to draw these cognitive economic curves. So, this is very early data. I want to mention that. Still, this is very ongoing. But this is the control cognitive economics curve. You can see that it's fairly flat. People aren't much better at accepting good deals when it's really incentivized to accept good privacy policies. In fact, it's within the bounds of having no incentives at all, right? So, people aren't really learning anything from these disclosures. And just the first pilot that we've run on this, the simple curve does much better. People are rejecting bad deals or bad policies with a very high probability, and they're accepting good disclosures with a very high probability.

So, just to wrap up, we introduced a method to measure disclosure clarity through revealed complexity. We have an application of privacy policies. And the broader agenda here is just that cognitive economics in general. This whole field has been really interested in measuring things like complexity and cost of information, but has been stuck in a very, very artificial setting. There's a lot of work to be done converting this into things that people actually care about like privacy policies. So, thank you.

Speaker 19:

[inaudible 04:06:03].

Shotaro Nakamura:

Sure. Hi, my name is Shotaro Nakamura. I'll be presenting on work I started during grad school with... do you have a clicker?

Speaker 18:

It's there.

Shotaro Nakamura:

Oh, I see. The green button?

Speaker 18:

At the bottom, yeah.

Shotaro Nakamura:

Okay. That I started with Ali and Adeel at Lahore University of Management Sciences. This is a development digital economy and marketing paper in that... oh, I see. We're interested in how information interventions or platform design could reduce and improve market frictions in developing economies. And these are particularly important concerns, given that online marketplaces have been becoming more relevant across different contexts in countries, such as large e-commerce platforms and ride-sharing or other forms of gig work.

The idea is that the existence, all of our further improvements in these online platforms, would further reduce frictions and then that would induce welfare impact as have been documented in previous work on ICT, more broadly. But the challenge is that there are residual friction in these markets and the extent to which there are residual frictions may be functions of the platform design, as well as other confounding market failures that may exist in developing economies. So, we're taking this information intervention as effectively a case study to try to understand, in such context, how do these interventions reduce information and search frictions and/or it would have spillovers and other adjustment mechanisms.

Our context is listing platforms for used vehicles in Pakistan where there are large search frictions and noise and beliefs about the demand that folks face, in that these types of platforms are increasingly common, particularly in places where they import used Japanese vehicles in countries where they drive on the left-hand side. And what these platforms do is they let transactors expand their search and partners for transactions beyond their existing social network. And there you go. Oh, and then we partner with this platform called PakWheels.com, which is the leading platform in Pakistan. And consumers or sellers and buyers on these platforms face limited information. They have noisy beliefs about the demand they're facing. And as a result, they face high search friction. To combat these challenges, the sellers in this context can promote their ad by basically paying money and having them be on a a more visible portions of the search results.

And given the publicly available pricing information, there's still a fair amount of discrepancy between what is publicly available in a form of listing prices and the eventual transaction prices in that being. So, what we do is that we, in our experiment, we provide pricing information privately to sellers in the process of they're making a new post. We construct a machine-learning-based price prediction model that we call the price calculator, and we provide this information privately to sellers when they're creating posts. We randomize in two stages so that we can detect both direct treatment effects, as well as spillover effects.

And this is a screenshot or this is what you would've seen if you are a treated seller. You type in the information about the vehicle that you're trying to sell, you've set your listing price, and then you're prompted with a pop-up that shows you the price estimate, and you can then, if you choose to do so, update your listing price.

We follow the choices that sellers make on the platform in terms of the deviations of their listing price away from the price calculator estimate that they would've received or they received or they would've received transaction outcomes in terms of sale, occurrence of sale, and transaction price, and several categories of advertising tools that they had access to bumps, which are bringing up your post up on a search result, featured ad sections, and vehicle inspection requests to basically signal quality. We can also track buyer attention on a given post by tracking our click-throughs.

And I'll talk about results in words as opposed to going through tables. But what we find is that the information intervention reduced price deviation from the price calculator estimate by about 7 to 11% for both directly treated and the spillover groups. And we also find that the intervention increases transaction probability by one percentage points for the spillover group, but not for the treatment.

And we pre-specified two mechanisms or two outcomes. One is advertising usage and we find that there's reduction in ad usage by the directly treated folks and accompanying that increased page views for the spillover groups. So, table.

And to understand and validate our claims that we're making about the mechanisms, we also have a pre-specified framework of static search in which there is this component of information friction and beliefs updating process as a mechanism. And we conduct an inline survey of about 3,000 sellers to capture their beliefs about the demand that they're facing, as well as their experiences on the market.

And we find effects only on the directly treated folks on several measures of beliefs. One is that they believe that they shift their beliefs about the demand that they're facing closer to what the signal that we sent in the form of the price calculator had. They also believe that they experienced less search frictions or that they believe that the market conditions are slightly better than folks who are in a control group. Sorry, this is an old slide.

But what we have found is that, if you're directly treated, you update your beliefs about what kind of market conditions you face, you adjust your pricing, and you substitute away from advertising. Whereas, if you're a control group or rather a spillover group, you only respond to the choices that are... or publicly available information that are available on the platform and the choices that are made by your competitors.

So, in conclusion, we find... sorry, this is an old slide. I'll conclude with what we can attribute are smaller effect sizes to be, in that the small effect sizes that we identify may be the results of spillovers, but also, the extent to which the interventions can shift market participants' beliefs about the demand and search frictions that they're facing. And our work also ties into the broader literature about micro-entrepreneurship training in developing economies in that these are fairly low-cost interventions that platforms or policymakers can engage in through existing online platforms. So, thanks. Sorry, I had bajillion appendices that I didn't realize were going to go up here. Okay.

Speaker 19:

Very optimistic, I think.

Shotaro Nakamura:

All right. Now, we have Daniel Hosken.

Daniel Hosken:

All right. Well, thanks. I'm very excited to talk about a project I've been working on with Frank Pinter and Devesh for a couple of years now. As Devesh indicated when he was describing me, I have spent a very large portion of my career estimating the effects of consummated mergers. The typical merger retrospective study focuses on price and output, typically. And that's because... and I trust economics is normally focused on things like consumer exit, measures of substitution, price elasticity, trying to think about how pricing incentives change when you change competition.

There's been much, much less information looking at what people think, what their opinions are of stuff, things like voice. So, what this paper is going to do is it's going to be a merger retrospective. We're going to look at what's called a divestiture, and I'll describe in some detail what that is. But it's a big antitrust action where, when some companies were trying to do a supermarket merger in order to have the merger cleared, they had to sell a ton of stores. What we're going to measure is what consumers thought about the divestitures, how the consumers who use these divested stores reported their satisfaction using the sample of, well, using the universe of Yelp reviews.

We think this information can be really helpful in measuring things like quality. One of the big problems, particularly, in retailing is trying to come up with good measures of quality. It's almost impossible to do. You can barely measure price. And then, more generally, this is a promising area of research to try and improve antitrust enforcement to take advantage of what consumers are actually saying to learn about the effectiveness of policy. So, let me just provide a really super-duper brief description of divestiture policy and the logic underlying it. So, particularly, in the retail setting, if you think about supermarket mergers, historically, the way regulators thought about these deals is that all the competition that mattered was highly localized. So, you can imagine two large supermarket chains merging. And let's say that they directly compete in 10% of their stores. Like, those stores are in the same regions, they're overlapping drawing from the same area. What the government would do is they would tell the merging parties, "You have to sell the stores in that overlap area to eliminate the competition that would be lost. You have to sell it off to some firm not participating in the market. And then your deal can go through."

The logic underlying this divestiture policy was that the new buyer of these assets will be able to operate them in a similar way, provide a similar quality of service, similar prices, have similar costs. So, if magic happens, it's like the merger didn't happen and your competition is maintained. We're going to look at two big mergers that took place in 2015 and 2016. One is Albertsons-Safeway, the other one, Ahold-Delhaize. These were big mergers that had large divestitures. So, the first merger, 168 stores were divested and a bunch of individual markets, primarily in the western part of the U.S., 146 of them went to a single firm that later went bankrupt. The Ahold-Delhaize deal, 81 stores were divested to a bunch of buyers in these different areas in which the firm operated.

A key thing to know about these divestitures is what was sold where the retail outlets, so the physical locations. The banner, so it would be like Giant or Safeway, that was not transferred. So, when these assets are transferred to a new firm, they have to come in with a new product, which most consumers would be unfamiliar with. So, there's an element of risk coming from that. Here's just a map showing you where the merger took place. The green dots are the primary stores sold in the Safeway Albertsons deal, all on the West Coast. There were some other stores sold in Texas and Wyoming and Montana. And then the Ahold-Delhaize was all on the East Coast.

So, we have three research questions we're going to get at here, and I'll show you some pictures showing you the answers in a minute. The first thing we are interested in looking at is to see if people were more likely to review stores after these divestitures. And the data screamed at us, "Yeah, the consumers really started reviewing stores on Yelp after these divestitures." Were the consumers happy or were they sad? The consumers were really, really sad about these divestitures. The reviews were just overwhelmingly negative. And then we did some work looking at the text in the Yelp reviews and we saw that what most of these negative complaints focused on were prices. So, real quickly, what's the data we're using? This is standard Yelp data. Devesh somehow magically got the people at Yelp to give us the universe of the supermarket reviews that Yelp had from the beginning of Yelp until, what, 2020, something like that. Our estimation sample has about 600,000 reviews. We're going to take advantage of the star rating to determine whether review is positive or negative. We're going to look at the text that the person left in the review. And then we have information identifying the store. And what we're going to do... and we had a research analyst help us with this, is to really identify retail locations, physical locations. And that's going to be our unit of analysis. So, we're going to look at stores that changed ownership because of the divestiture, so the same physical location, and see how stuff changed in that store. We aggregate the data up to the quarter, because I think unlike restaurants, there aren't a ton of reviews for supermarkets. So, they don't take place very frequently.

Empirical analysis, like most of these merger retrospective type studies, it's a difference in difference analysis. We're going to compare these divested stores to stores that are in the same state. So, hopefully, experiencing similar demand and supply conditions, similar trends in Yelp reviews, but in different three-digit zip codes, so, they're not directly competing with the divested stores. And it's not on the slide, but we also look at stores that are owned by firms not directly involved in the merger or the divestiture.

We're going to use Poisson regression. It's standard two-way, fixed effects, fixed effects for the retail location and time. We're going to look at three outcomes, the number of reviews, the share of negative reviews, and we're going to look at the content through magic. We use this zero-shot classification that goes through and looks at these reviews and assigns a value between zero and one if a certain topic is mentioned.

I got to read a few hundred reviews to try and identify what topics typically come up. And then we specify those topics and they're identified by the algorithm. Prices, products, and customer service are the primary topics we focused on. These are not mutually exclusive. A review can talk about multiple topics.

All right, I'm going to show you everything in the paper. Well, there's lots of other stuff in the paper, as you have to be written. But the key things for the paper can be seen in three event study figures. First one, does the divestiture cause people to start reviewing? So, in the pre-period, we have these control stores and the stores that are going to be divested. They look the same. In these three quarters, after the divestiture, there's a huge spike, like a factor of four or five. So, after the divestiture, the divestiture is causing tons of people to start reviewing these stores that were divested. So, then the question is, are the people happy or sad?

So, we look at the share of negative reviews, huge spike in negative reviews. So, people are sad. Then we ask what are they sad about? So, we focus on the negative reviews. We see a huge spike in people having reviews, essentially complaining about the pricing at the divested stores, some complaints about the products, product quality, things like that. And if anything, the share of people talking about customer service is pretty much unchanged.

So, that's a sketch of what this paper's about. We think this is a really promising area for research to really start focusing on using things like reviews or customer comments to learn more about the efficacy of antitrust policy. Thanks.

Speaker 19:

All right. And last, we have Ben Casner.

Ben Casner:

Hello, everyone. Thank you for coming. This is the second theory paper today. Will there be a third? We'll have to wait to find out. So, I'm going to be talking about the relationship between market platforms curating the quality of sellers that they host and the of these sellers to try and disintermediate.

The motivation for this is coming from a recent regulatory trend that pertains to a reduction in anti-steering clauses. So, essentially, these regulations are saying platforms can't make it harder for sellers to try and steer consumers towards transacting directly instead of transacting via the platform. And the primary motivation for this paper is actually the Epic v. Apple case. So, Epic's one win in this case is that app developers are now free to include a link directly in the app that will take you or take consumers to direct signup flow that avoids the app store commission. There's an asterisk there. This is an eight-minute talk. I'd be happy to talk about that asterisk offline.

And so, the main effect here is that there's going to be a marketplace leakage risk for app stores. So, it's going to be easier for the buyers and sellers to disintermediate. Apple's response to this ruling is that they have started putting in these scary warnings essentially saying, "Don't come to us asking for a refund if you're not satisfied with your purchase." There's obviously a liability angle here, but I was also curious whether this could be an attempt to scare consumers into not transacting directly. But this is only going to be effective at reducing marketplace leakage if there are low-quality sellers present that you might want to get a refund from. So, my question is, does Apple have an incentive to reduce screening in order to reduce disintermediation? To answer this question, I create a model. The model has a monopoly platform. It's setting an entry fee row. You can think of row here as being the price of an iPhone. And an ad valorem commission xi for sellers, this is the 30% commission that everybody complains about when they're an app developer. And it's also going to set a proportion alpha of low-quality sellers. And I'll talk about what low quality means in a minute. There's continuum of sellers, there's mass one. Their only decision is whether to allow transactions via the platform or to steer consumers towards direct transactions. And a continuum of consumers, they decide whether to pay row and participate on the platform or stick with our outside option. And if they do participate on the platform, they face a Wolinsky-ish search problem.

So, if a consumer transacts with a low-quality seller, they're going to get negative utility. If they encounter a high-quality seller, there's going to be a probability that that seller is a taste match. And if the seller is a taste match, their payoff is going to be v minus mu. So, mu is an exogenous price charged by the seller. But consumers can't observe whether a seller is low quality. So, if they encounter a low-quality seller, they're just going to get a fake quality signal that tells them whether or not this product is a taste match. So, they can't observe whether a seller is low quality, but ex-post they have the option of asking for a refund.

If they transact through the platform, the platform is going to grant this refund. If they transact with a low-quality seller, the low-quality seller is not going to grant this refund. This means that, low-quality sellers, because this is an ex-post refund, are not going to make a positive profit if they transact through the platform. And so, they're only going to steer consumers towards direct transactions. So, no low-quality seller is going to be willing to go through the platform.

So, I'm going to define this object psi. So, psi here is, given a proportion beta of high-quality sellers steering to direct transactions, the probability conditional on encountering a seller who is steering towards direct transactions that that seller is high quality. So, if you encounter a seller steering towards direct transactions, the expected utility is going to be the probability that they're high-quality times that payoff probability that they're low-quality times that payoff. This is decreasing in alpha. So, as alpha increases, the expected value of transacting directly goes down, unsurprisingly.

I can plug this into the value of participation. And there's going to be two different strategies that consumers will adopt, depending on the value of alpha. So, they could search and buy from the first positive signal having seller that they encounter, in which case... so, the first two terms are going to be roughly that expected value of transacting with a seller, so that it's going to be some possibility that they encounter a low-quality seller, get scammed, get a negative payoff. But s here being the search cost, that's going to be lower than their other option, which is they search for a seller who is transacting through the platform. So, if they transact only through the platform, they know they're going to transact with a high-quality seller, but they're going to have to search more. There's going to be some cutoff alpha where this second strategy is superior.

So, consumers are going to participate if they're indifferent, which means that the optimal row is going to be the max of these two value functions. It's pretty trivial to show that the transacting with the first positive signal you encounter is going to be superior at alpha is equal to zero, and it's going to be decreasing in alpha, which is going to be decreasing consumer willingness to pay. So, it creates the question, why would the platform set alpha greater than zero? And the answer is the commission on sellers. So, if sellers don't steer, then they're going to have to pay a commission. If they do steer, they're not going to pay a commission, but they're going to pay a steering cost, which I'm going to denote t. If the consumers are willing to transact directly, if consumers aren't willing to transact directly, then they're going to get zeroed profit. The platform is going to set xi such that sellers are just indifferent between steering and not. And this is increasing in alpha.

So, my main proposition, the platform is going to set positive alpha if mu is not too small, essentially the sellers get enough surplus such that it's worth degrading the platform in order to extract more surplus from sellers. And this positive alpha is going to be decreasing in the steering cost. So, if we go back to this motivation, we can think of the Epic v. Apple ruling as decreasing the steering cost. So, the answer to my question is, yes, Apple has an incentive to reduce its screening as the ability to prevent this intermediation goes down.

So, just very quickly, my last slide, the solution for a regulator is to introduce competition. If you have competition, the competition for consumers is going to drive alpha equal to zero. Consumers and sellers are both going to benefit. And this trade-off from lowering t is going to disappear. Thanks.

Speaker 19:

All right. At the risk of keeping people from lunch, I think we have a few minutes for questions.

Speaker 20:

Question. Yeah, there it goes. I guess the question is, we would expect customers at a store that has changed brands to be unhappy and to comment on it about something. And the question is, do you have any evidence that prices did increase, as we might expect, if they were-

Speaker 21:

Prices did increase as we might expect if they were separated from a big chain, or was there something else going on?

Daniel Hosken:

This is a very involved question. The Haggen transaction, which is a lot of our data, Haggen was a huge failure. The company went bankrupt six months after the divestiture, and there was litigation between Haggen and Albertsons Safeway about how they were treated in this divestiture. And one thing that's interesting is if you look at the public complaint, you can see that there's lots of complaints by Haggen that they didn't know the prices that were charged at the Albertsons stores during the changeover so they used the price vector from their stores in Seattle and were imposing it in other parts of the country where those prices were inappropriate. And in looking, I've looked at hundreds of these reviews, you see a lot of complaints like, "They sold the store to this other company and the price quality match just doesn't work. It's inappropriate." And that's backed up in the topics thing that we see when we look at the entire collection of reviews. This data, we only have the Yelp data, we can't observe prices, but it's certainly consistent with the price quality relationship being out of whack with what had been there previously.

Speaker 22:

I have a question for Eric. Very interesting work. I wonder if your method requires the researcher to observe the ground truth of the utility of making the choice. And if so, how would we deploy that to actually examine the complexity in privacy policies? Because, well, in privacy choices, oftentimes we don't observe the ground truth.

Eric Spurlino:

Yeah. I think that's a very valid question. I think you do essentially need to understand what the benefits are so that's why in all of our work so far, we're just experimentally endowing people with incentives about privacy. We do have hopes to be doing things in the future where we ask people about their privacy preferences and then try to use those to incentivize them.

But in general, I think what we can do here is if we assume that the regulator has some beliefs about what consumer's incentives are from the research that everybody here does, we can use those and then experimentally implement them in a lab kind of setting. And then, hopefully, things won't be too off from reality. But that is a weak part of the design.

Speaker 23:

This is a question for Dan. I am really interested in what the policy implication of your work is. In particular, should we take the unhappy customers to mean that the grocery stores should have been able to keep all of their stores so that the merger should have happened and divestitures wouldn't be necessary? Or should we have prevented the stores from merging in the first place?

Daniel Hosken:

Well, that's a policy question for people like our commissioners to decide. What I think I take away from this study is the FTC has done a number of studies evaluating the efficacy of divestitures and the divestiture policy, and one of the factors that was identified there was the importance of having a divestiture package that could allow the new firm to replicate the services provided by the previous firm.

In this transaction, one of the assets that was not transferred was a retail brand. And I think the retail brand, the pricing strategy, those kind of things are probably very important to transfer to a new buyer to kind of replicate the certain level of competition that was there before. That's my interpretation of the kind of customer complaints we see is that the price-quality relationship the customers are seeing, they really don't like, and I think it has to do with the change in retail format that is coming from the deal.

Speaker 23:

Do you have any information about what the customer voices were at the stores that were involved in the merger? Were people even more unhappy?

Daniel Hosken:

Well, the average review falls at the stores that are divested pretty significantly. Are you saying... I don't think we've explored yet the population of individuals who would've observed the store pre and post. I don't think we've got a lot of data on that. The supermarket reviews are fairly new, so most of what we're seeing are new people reviewing stores. It's hard. That's something we want to explore. That's part of one of our ongoing projects is to see if we have enough power to see if specific consumers are describing their experience going down. Because for those people who do multiple reviews on the same store, we observe them.

Speaker 24:

All right, I guess we can go to lunch.

Speaker 25:

From...

K. Sudhir:

... for the next session. Before I talk about the session, I just wanted to say, anyone wants to join for dinner and you haven't still done it, I think we still have a few spots. And if anyone wants to register, you can still do that.

Okay. With that, let's get started on the next session, Advertising and Information. Our first speaker is Sylvia Hristakeva, and we are just going to follow the same rules that we've been following.

Sylvia Hristakeva:

Thank you. All right, so today, and right after lunch, I know everybody's just digesting and I'm going to change things a little bit and talk about the analog world. Move us a little bit away from the digital world. I'm going to study the effects of drug injury advertising on prescriptions, and more importantly, on health outcomes. Drug injury ads are used by lawyers to recruit plaintiffs for mass tort lawsuits. These mass tort lawsuits alleged that companies knew about a problem with a product and they just did not adequately disclose the information to their clients. More or less, they're targeting companies that allegedly injured a large group of people. In order to reach potential plaintiffs, the lawyers use drug injury ads. A lot of these lawsuits target drugs and medical devices, hence the term drug injury, more or less. Hopefully, most of you have seen one of these ads, but if you haven't, here's how the script goes.

If you or a loved one has taken this drug and suffered one of these many, many, many, many, many conditions, please call this number because you might be eligible for compensation. I have two main research questions. The first one asks, do drug injury ads affect prescriptions? We have a lot of information about the effects of direct-to-consumer ads on demand, but these ads are very different in both in their intent and their content. They're emphasizing a lot of the seriousness of the side effects and the severity of the risks while you're taking the drugs. To the best of my knowledge, this is the first paper that actually talks about negative advertising in the pharmaceutical world, in the healthcare world.

So I'll find, as you can expect, that yes, drug injury ads do affect prescriptions, but by using only that piece of information, we still can't say anything about public health because ads might be misleading. They might be scaring people a little bit too much or leading them to take the wrong decisions. At the same time, we still want to say something about public health, and this is where my second question comes up and asks, are there measurable effects on public health from drug injury ads through any changes in-

]

Sylvia Hristakeva:

... We add through any changes in drug utilization. Here, what matters is not only the direction of an effect, but the magnitude of it. And we would see, hopefully, I'll have enough time to show you that the effects are relatively large in the case study I studied, and that would give us some clear policy implications both for public health policy, as well as more about just continued oversight of drug injury ad content. And I'll revisit that at the end.

I'll study all these questions in the context of anticoagulants, which are drugs used to treat to prevent the risk of stroke. Let me jump straight into this. I suspect most of you haven't seen much about the lawsuits. Most of these lawsuits are played as failure to warn cases, right? The company knew about something, and they just failed to disclose it. That means that the lawsuits and the ads can happen even if the FDA doesn't change their recommendation. Yes, it has to do only with warning people.

Quick note here is that what I'll study is not just a one-time event. Just last week, I was looking for new drug injury ads and surprise, surprise, we already see the playbook playing out for Ozempic and Wegovy, right? We know this is going to be a huge market with a lot of potential plaintiffs. And more or less now, we're just trying to figure out, are there certain side effects that are prevalent to us and haven't been disclosed before, that we can have a good case against companies?

Today, I'm not going to say anything whether the lawsuits are good or bad. I don't have the data for that. I'm not going to touch it. The only thing, though, we can say for sure is that there are high financial gains from representing a lot of clients. And this is why the companies advertise quite a bit. Usually, the ads are triggered by new information, either about the product, such as a change in labeling, or about the financial benefits from a lawsuit, such as a settlement or a positive verdict.

The ads heavily emphasize the serious side effects of the product. And in this case, we'll see, I have an ad there for Xarelto, an anti-coagulant drug, where the ad is highlighting or at least suggesting that the drug might be causing stroke, even though the drug is used to prevent stroke. The more unkosher way, things that companies usually do is that they present the information as a medical alert or an FDA alert. It creates a little bit of attention there. I'll have to be honest that I don't have information about the content of the ads, so I'll be very cautious about interpreting my results in that direction.

All right. Now, to tell you a little bit about anti-coagulants, I had to learn a lot about health in this project. They are used to prevent the risk of clots, stroke, and embolism, and AFib. They're primarily taken by elderly people, so that's why I can use Medicare data without any problem in this setting. There are two types of anticoagulants. One of them is warfarin. It has been there in the market since the '50s. And these new guys entering in the 2010s that are called NOACs, non-vitamin K oral anticoagulants. You can see in the graph here how the patterns of prescriptions for both of these two types of drugs changing over time. And you can see NOACs gaining a lot of market share over time, mainly because clinical trials show that they're more effective and safer than the older drug.

One quick note, yes, they are not coming without their risks. There is a side effect of severe bleeding and that was the actual reason for the lawsuits. So in the lawsuits, they alleged that the NOACs did not adequately disclose the risk of severe bleeding. That doesn't say that they are worse or better than warfarin. According to clinical trials, they actually have less severe side effects. The comment here is all about disclosure. Very quickly, to tell you how the timeframe went, is that in 2012, there was an initial wave of ads. That's what I'm putting on the left-hand side of the graph. So these new initial ads definitely generated some lawsuits. The first company settled in 2014, and that's when you see all of the ads coming up, right after the settlement because now you can advertise, get more plaintiffs, and then submit them towards the original settlement. And then go after the next company.

I'm going to study that second wave, not because the first isn't interesting, but because that's where my data are. And straight to my data. I have three types of data sets. The first one is the typical advertising data, that's the DMA-month. And then we have prescription-level data and inpatient hospital... Sorry about this. Inpatient hospital visits from Medicare patients. These data sets are publicly available, and that's one of the main drawbacks and concerns about my approach, and we have to be very careful with it, because my data are annual.

My first analysis we want to see is, hey, do prescriptions actually change in response to ads? Here, we can think my unit of observations is at the drug-physician-year level and want to regress this on ads, maybe throw in some fixed effects, and very quickly, you'll tell me this definitely has an endogeneity problem, because law firms might be specifically targeting certain areas and spending more in those areas where they expect more people to file a lawsuit.

All right. We're going to use in this case, the border strategy introduced by Shapiro, whereas more or less the main idea there is that we're just going to compare counties that are bordering each other, and assuming they share the same demand, but just randomly, they end up being on opposite sides of DMAs, so they get different advertising. That just gives, the way we'll operationalize it, is to just use data only on physicians located in this border counties. So you see, I lose a lot of power because of that. And we're going to control for border-time fixed effects.

My first set analysis looks at prescriptions of NOACs, warfarin, total anticoagulants, and a nice clean placebo. Because of time, I'm going to just focus only on the NOACs. We do see that big decrease in prescriptions. But what does this mean? Let's put it in perspective a little bit. Consider in this case a $0.19 increase in advertising, which is the difference between the 25th and the 75th percentile of the ads distribution. That would lead to about 2.6% decrease in prescriptions, and that is massive for this market. That's about more than 300,000 decrease in prescriptions annually. This can have very large effects on health outcomes, but again, because we don't know if the advertising was steering people in the right way or the wrong way, I can't use just these numbers to infer anything.

At the same time, I can't use clinical trials as well. In this case, people might be making the right decisions for themselves and that's why we'll use the second question. Are there measurable effects on health? Here, I want to regress health outcomes on prescriptions, and then again, control for fixed effects. And we'll need to instrument in this case for prescriptions, our instrument using drug injury advertising.

For the sake of time, I won't go into why this is a valid instrument, but quick comment here that it's a late, like a local average treatment effect, but that's exactly what we want to know in here. We want to know what's the impact of the prescriptions on patients that were influenced by the ads. So it's the right thing to estimate. We're going to see this for relevant diagnosis, side effects, and the placebo. Again, I'll focus only on the relevant diagnosis today. Where here, I'm showing in the first two columns, the results for the OLS and an IV clearly showing the importance to instrument for prescriptions. And we see that prescriptions do prevent the risk of stroke, but again, what does it mean?

All right. So let's pause on this one, and then calculate the same thing. Those same $0.19 increase in ads that will lead into 2.6% decrease in prescriptions, are leading to about 1.4% increase in inpatient hospital visits, for a relatively serious condition like strokes, AFib. This is a large number, even if I look at the lower bounds of my confidence intervals. And if I extrapolate it in terms of number of people, we get about an increase of about a thousand additional inpatient hospital visits for the border counties themselves. Right? In this case, we come up to this point of like, magnitudes do matter.

All right. My last minute, let me just summarize this and bring us back to the policy dimension. Yes, we see novel results of negative advertising on demand, and these are having real impact on health. It is true the bar for proving that the ads are misguiding and misleading people is high, and I don't even know whether individual level data can help us answer these questions, but these results can still give us clear policy implications. We know that ads can affect consumer behavior, and that's why the FDA carefully regulates the content of most direct-to-consumer ads from pharma companies. At the same time, the ads I'm studying are regulated by the FTC and the state bars, and until recently have been operating with little oversight. In 2019, the FTC sent seven warning letters to lawyers. And I guess what this results are clearly telling us that yes, keeping an eye on these ads and making sure they don't have completely inaccurate information in them, makes a lot of sense in this, because they can have really large negative impacts on public health. Obviously, my results have even clearer implications for public health policy, and the idea that we might want to have a little bit more of these informational campaigns that we already have for vaccines, and just having those more or less playing out whenever we have more of these ads, with the assumption that they're going to go after different drugs coming up. Thank you.

Sarah Moshary:

Okay, terrific. Hi, I'm Sarah Moshary, and I'm delighted to be here today to talk about Sylvia's really excellent paper. Her paper touches on a broad concern that marketing practices, and I think advertising in particular, can lead consumers to make suboptimal choices. But usually I think of this as ads convince you to buy something that you end up regretting, that you don't really need or you don't really like. But in this case it's going to be concern about the opposite, that this particular type of advertising can convince people not to buy something that actually would really enhance their utility. In particular, I think this is a context we care about, Sylvia emphasized the public health aspect. So I think that's a real strength of the paper. To reiterate, she's focusing on these ads that law firms are running to recruit additional clients to come in on these mass tort suits against pharma companies.

One of the great things about the paper is that she's not just looking at the purchase decision, which is the prescription, but she actually has then the data on hospitalizations. I think that's really special, among many other nice things to about the paper, including the lengthy and serious discussion about the endogeneity of advertising. She barely scratched the surface on that. There are a bunch of robustness checks. She's very clear about not just thinking at the point estimates of the effects, but at the top or bottom of the confidence interval, how important are the effects she's finding? So there's a lot to like here. Most of my comments are really just about how we interpret the findings, and sort of where you go next from the results that she presented. And so there's sort of this flavor of deceptive advertising, that these ads are leading consumers to have incorrect perceptions about the benefits of anticoagulants.

But it's actually quite difficult to nail that question about misperceptions, because it could be that what these ads are doing is that they're giving information to consumers about side effects that they were unaware of before, and actually leading consumers to have more accurate beliefs. And then since we're not directly measuring beliefs, it's hard to say before or after the advertising, are beliefs more or less accurate. And actually I was sort of just even curious to know, what did the drug ads look like for the anticoagulants? Did they stress the benefits? Did they give enough sort of airtime to the potential side effects? So I think some more background would be helpful, but there's a sort of fundamental measurement problem. We also, of course, don't know exactly what the patient utility function looks like. So we're not measuring their beliefs, but we're also not measuring how they would trade off side effects from the benefits of taking the anticoagulants.

And it could be that the belief system of these patients, or the preferences of some of these patients, differs from say what an expert like a doctor who prescribes the anticoagulants, would trade off the risks and benefits. And so if we want to then sort of take that next step to measuring consumer welfare, that's going to be a critical input to try and measure that. And I think that also goes to, some of the same is true for trying to understand the potential benefits of things like a public service announcement about the benefits of anticoagulants. That could be highly effective if people have misperceptions, but if really people are pretty well-informed about the benefits, now they see these ads, they sort of learn about the costs and they make a decision that's good for them, PSAs might not be such an effective policy tool. So I see this as a really excellent paper that's sort of opening up a bunch of questions around this important topic.

Speaker 26:

So regarding this question of, are the ads actually misleading people or not? You did measure the hemorrhaging incidents, right? That was supposed to be a side effect, and you didn't see any effects, correct?

Sylvia Hristakeva:

It's just too noisy. But even in that case, I see that again, the newer drugs are better than the old drug, so they are decreasing.

Speaker 26:

So that's what leads you to believe that these ads are misleading?

Sylvia Hristakeva:

I agree with Sarah that using aggregate data is going to be hard to do that. So my only way of doing it... Okay, I'll pitch you potentially my out of paper. If I get the individual level data, can we say something better? In that case, I think I need to talk to cardiologists, but I think there is... We're foreign, the old drug is strictly dominated by the new drugs. So seeing substitution in that direction, I think can alleviate a lot of the concerns. But that's conditional in talking to cardiologists first. But that's my main way of doing this.

Speaker 27:

I was wondering, is it possible that this could also be happening because doctors are reducing their prescribing to avoid liability?

Sylvia Hristakeva:

I know for sure they're not doing that. I definitely checked in on that. I've talked to the lawyers engaged in these lawsuits that the doctors are just not being engaged, they're not liable for any of this. So we kind of got on that. And also, you're only responsible for things that happened in the past before the new information was revealed. But it's also a great question of whether the changes are happening at the patient or the physician level, against something that hopefully individual level data can help me split up.

Speaker 28:

I have a question regarding the RQ-1. Actually is a comment, I want to see what you thought of it, as I wonder to what extent it is a mere exposure effect that decays over time. So did you see if it becomes stable, or does it decay?

Sylvia Hristakeva:

It's annual level data, so I don't have enough power for that. But once, if I can manage to get individual level data, that'll be very interesting question to understand.

Speaker 28:

But you should still see that [inaudible 05:57:58] they were, right? [inaudible 05:58:00] time.

Sylvia Hristakeva:

Yes, they do recover a little bit, right? After the ads are gone.

Speaker 28:

My question is how much is the stable effect? How much is the decay effect, which is due to mere exposure?

Sylvia Hristakeva:

I'll have to look into that. Thank you.

Speaker 28:

Thank you.

Speaker 29:

Thank you. Very interesting. Very interesting work, thank you for sharing. I'm wondering down the line, do you feel that this would drive a lot of pharmaceutical companies to sort of blanket package an entire list to protect themselves? To the point where it doesn't really mean anything for a consumer, and some of those side effects could be highly improbable, just as a defense mechanism?

Sylvia Hristakeva:

I cannot speculate on that, but the one clear thing with these pieces is that, I didn't have the time to mention this, companies... And I love to talk about these lawsuits with everybody. We don't have much work on this. I don't know how important they are for the economy, but it seems like they target many companies, especially in the pharmaceutical world. Most companies, even though you're winning in court, you settle. So law firms can still, as long as you have enough people, you can still go after them and the company just chooses to settle.

Speaker 30:

[inaudible 05:59:27].

Speaker 31:

I have the mic, maybe I can ask that. So very nice work. Very interesting. I have one thought, it's also kind of related to some of the other work. It might be interesting to check how the drug companies respond to it using detailing efforts. Another comment I have is, obviously the result is very interesting, but also it could be context specific in the sense that if the lawyers go after another drug that turns out to have some other very close substitute that are also very effective, then the public health problem may be not so serious.

Sylvia Hristakeva:

Yeah, first question, great. I did look at physician payments. That didn't change. I can't touch anything from Nielsen, because this studies lawsuits and I chose to disclose the names of the companies. Second question, as any empirical study, this is case specific, there were three drugs, that only two of them were heavily targeted. So I can potentially start splitting them up between the three drugs. But all three of them were by definition, affected by this, even though one of them was not targeted at all by any lawsuits. All right, thank you so much.

Mingduo Zhao:

Yeah. Hi everyone. I'm Mingduo Zhao. I'm a PhD student in Berkeley Economics, and I got a diverse background also in computer science and statistics. If you are curious, just a brief introduction. There are also a poorly-written bio there if you want to check out for more detail, I also got a LinkedIn page. So I'm honored to be here to be invited, and also I'm happy to see all those familiar faces. As a graduate student, I mean, in most of cases I just say in audience to press my idea or hold my opinion. But now on your circumstance I just share with you all those this great project with you. So this project is with Yunhao and Miguel is called "Masking and Deception, the Interplay Between Fake Reviews, Rating Discrepancy, and the Consumer Demand." So there is a huge component for these three terms, which is a fake review rating, discrepancy, and the consumer demand.

And I feel fake reviews, or in general fabricating information on a platform, is very harmful for people. Could include fake reviews on Amazon or any bots in social media, and in general, platform is trying to tackle it by teaching their or giving more information to their consumers so that they can figure out which piece of information is fabricated. And then we will talk a little bit more about that piece in a second.

So just look at the distribution, two distributions we have here in the classic Amazon settings. Although the average score for this two products or for this two distribution are exactly same. But the left panel is a little bit discrepant in the sense that at the same time they're high percentage of five star and also one star as well. So here we measure the discrepancy by standard deviation, but you could imagine in other scenario where you can use other measures to describe this discrepancy. The first research question we got is basically whether rating discrepancy is correlated with fake review, and why that could be theoretically true. Because we got two piece of evidence. The first one is, we know that if a company want to buy purchase fake review, what they did is they just reimbursed a lot of consumers if they leave five star for them, right? And the second case, there is a famous paper called "Why Consumers Talk," basically says if the expectation of the product quality is lower than the rating, they probably will retaliate by leaving one star review. That could contribute to a bipolar or a highly discrepant review distribution. And the preview of the result with our full study shows that empirically, that indeed true, because rating discrepancy strongly correlated with fake review purchase, conditional on product characteristics. We will see the empirical result in a second. Second research question regarding this project is whether rating discrepancy will influence the demand. And there are a lot of literature talking about how average rating will influence the demand, and whether discrepancy will influence demand. They have mixed evidence for both empirical paper and theoretical paper. And in our context the preview is, we got a novel identification method that on the percentage of rounding, and we found that 0.01 increase in standard deviation will cost minus 6.5% in demand. That's a preview of a result.

The third research question is related to how the platform going to educate their customer to figure out, or give more information for them to figure out, which part is a fake review. And the context we have is we use the experiment, which is information treatment we provide to the consumer. And what we found is 1% more suspicion, I mean we can hold all those definitions for later, but 1% of the suspicion will decrease the demand by 6.5%.

And one important thing is, socially disadvantaged group actually benefits more from the information treatment. This information treatment is very similar to the one you can see in developing economies, right? In a lot of development economics, what they did is they go to certain developing country, and they offer some opinion or some information for people so that they can manage their wealth, for example. And then that could help the people to smooth their consumption, or whatever.

So the theoretical framework, I'm going to jump a little bit ahead, but the page I want you to see is basically, there are two distributions. The one on the left is the observed rating, which is denoted by Rho-j. That's the actual rating you observe when you encounter the product, and you have certain beliefs about what percentage of the product of the five star review are being faked. You could come to that belief based on certain machine learning algorithm, based on belief you have, based on some of your reading of verbal text, whatsoever. But you have certain beliefs about a certain percentage of the five-star review are being fake. What do you do is, basically you get rid of that part and re-weight the distribution, so that the actual distribution that enter into your utility is a Tau-ij, which is the parties after you unmask the deception or you re-weight the distribution.

So here we are using two data. The first one is the data we got from He et. al. 2022, which is a [inaudible 06:07:19] science paper. What they did is they got a great label about whether the product purchased fake reviews or not. And another data set is what we scraped from Amazon. And one main variable we want to have is that sales data. What we did is, we calculated by inventory change and also we imputed it based on the methodology provided by that paper. And then by combining these two data, we can see how those three things, namely fake review rating, discrepancies, and demand, are interacting with each other.

The first empirical result I want to show is based on the raw data. All those plots are based on different average ratings of a product. And then the X-axis is a value of that standard deviation for that distribution. And the Y-axis is density. And you can see the one with green color is a product that purchased faked review. And the one with pink color is the one without purchasing fake review. And across all averages, you can observe that the one purchase fake review actually have a more discrepant review distribution. The second piece of evidence is also similar in a sense, that if we fix average rating, for example 4.0, and we go vertically, and we can see that as we have larger and larger standard deviation, the probability of us having fake review products is actually higher. Those two are all from raw data. I just want you to show that actually, the rating discrepancy has strongly correlated with a fake review.

The second piece of empirical evidence we want to provide is how rating discrepancy influenced the demand. What we did is, we just utilize adjacent day, the product in adjacent day with the same rounded average rating, but different rounded rating distribution as identification. So in certain sense it looks like RDD, but actually, we can kind of causally identify what's the impact of the Sigma KT on this YKT. By doing this exercise, we can find that actually, the sales is negatively impacted by the standard deviation or standard deviation of the rating. And the second column also show very consistent evidence, because the larger the sales ranking are, the lower lower the sales is, just like rhetorically.

So the last piece of evidence I want to show is actually how the suspicion of the fake review were going to influence the demand, and also how the platform going to educate their consumers so that they could be less suspect to the fake review. And so the experiment designs, we have three products. And those three products have exactly the same average rating, but their rating discrepancy are different. The first one has the largest, the third one has smallest. And then we ask the consumer to make choice for each pair, and then they also want to elicit their reasons as well.

And then after they make the choice in the first iteration, what we did is we give them information treatment. This information treatment basically give them information about how those fake review works. Basically we told them, "Oh, the company will purchase a lot of five-star review." And then after doing so, the consumer will retaliate. It is basically how the dynamic works. And they are 712 participant, and roughly half of them are in treat group and half of them are in control group. And what we found, this is huge table, but just use one sentence to summarize. Given that information treatment, will let people to be more concerned is it mechanically true. We'll let people to be more concerned about fake review, and also they will less likely to choose a product that has a higher variance. So they basically are more aware of the existence of a fake review, and they make choices accordingly.

And the part that a little bit more interesting, although things are two-stage lead square regression. What we are doing here, if you denote information treatment as Z, and your suspicion for the fake review as X, and your demand as Y, then Z is an instrument for the suspicion of fake review. And all those table are kind of showing different measures of your concerns for fake review. For example, if we have a least fake review as a reason, as a measure of people's concern for fake review, then we can see is positively impact the chance of you choosing a lower variance products. Basically, if you are more concerned about it, there are higher chance for you to choose a lower variance product. And those two tables showing very similar evidence, just two-stage least square regression.

Last piece probably very importantly, we find that by providing that information treatment, which could be interpreted as information provided by the platform or the education provided by the platform, you can see the gold one is actually from a pre-treatment. That is a measure of people's awareness or their preference for lower variance product. And then, this access we first have the Black, the second group is Asian, the third group is white. For pre-treatment. The awareness of fake review, or the preference for choosing a lower variance product, is actually increasing when we go from Black to Asian to white. But after treatment, they're very similar.

So what that is showing is, actually providing information treatment or providing education for the consumer is actually very important for the people who are actually from socially disadvantaged group. They probably have less information in prior about how this Amazon system work, and why they are fake reviews. How do I identify this fake review? By providing those piece of information, this could help them to make better decision. And then this effect is robust across education, as well. Earlier is about ethnicity-

Mingduo Zhao:

... Robust across education as well. Earlier is about ethnicity and here we have education. If we have a high school education, college education, graduate school education and unspecified, we also see very similar pattern. And then we also do the version with the income. We still have very similar result. Basically people who actually from a minority group or a social disadvantage group or people from a lower income group or people from a group that receive less education will benefit more from platform, giving them more information about fake review and they can make better decisions. And also this could be true if we have certain extension like providing their estimate about fake review percentage, I think similar result we're also holding that scenario. Basically just giving more information about it will be helpful for a lot of people who from socially disadvantaged group.

So then with the last one minute, I'm going to summarize. The first result we got is rating discrepancy negatively effect cells. This is not only talking about average scores or average ratings effect cells, the distribution of rating also influence the cell even if we condition on average rating being the same. And the channel we propose, this is not only channel, but one channel we propose is suspicion of a fake review could be the channel. And then we also see the evidence that rating discrepancy strongly predict the fake review purchase and also fake review concerns also affect amount as well. By providing information treatment actually a platform can benefit the people from more socially disadvantaged group and then those people could make better decisions. And that's it for my presentation.

Speaker 32:

Let's see. Great, thanks everybody. And I enjoyed reading this paper. So I thought what I would do is just give a little bit of the landscape on the way that platforms deal with fake reviews and think about some of the decisions at hand here.

So how might a platform mitigate the influence of fake reviews? Basically every platform is dealing with some version of this. One thing you could do is just remove or label spammy reviews. So platforms typically will do some version of this. So on Yelp for example, they pull off about a fifth of reviews that they now label as not recommended, but are essentially the ones that they think are spamming. Here if you're the platform designer, you'd want to think about how tolerant you are of false positives versus false negatives because it's pretty tricky to identify which ones are fake.

Another approach that people think about is verifying reviews that you know to be real. So think about someone who has an authenticated purchase or something you're confident in. So now that sort greatly shrinks the number of reviews that you have a lot of weight that you could put on. But then the ones that are there you sort of know are from somebody who's actually purchased the product or kind done the thing. You could have guidelines for reviewing. If you look across platforms, there's actually pretty different guidelines that they set. So different guidelines for example about even whether a business could directly ask for a review or whether they just could say, "Hey, people can review us here, but they're not allowed to go out and solicit." And this could sort of change the distribution of reviews across businesses. And then you could also have transparency at the platform level. So platforms sort of vary whether they just sort of pull a bunch of reviews off versus pull a bunch of reviews off of but tell you exactly how many they're pulling off and why. But individuals also decide which ratings to trust. And I think this is kind of where this paper plugs in. You could think about factors like how many reviews a product has, what's the average rating, who are the reviewers and trying to pull some inference off of that or are the reviews verified.

And this paper is thinking about, you could also look at the distribution of ratings. And if there's a suspicious looking variance, whatever that would mean to you in this case it might be high variance, low variance, then you could become skeptical of the body reviews overall. Here the main finding is that users seem to prefer lower variance products and that if you sort of associate that with... High variance products tend to have more fake reviews that have been caught.

So what I thought I'd do is just kind of reshow, this is one of the figures from the paper. So the basic version of the question we're thinking about is if you see two of these, which one would you prefer to have? So you see the same type of product, but two different rating distributions. So how many people would prefer the product on the left? How many of you want the product on the right? Sort of a mix. So now let's do the same thing with a slightly more skewed distribution. How many people think bottom of the left now, the bottom left? How many think the right? Okay, so more of a mix. And we could think about so what does that mean and how do we interpret the results? Kind of in you're looking at the variance of reviews. So if variance really is just indicating spam, then you would always say, "Hey, I want the thing that has less spam on it, so I'm going to take the lowest variance distribution of reviews that I see on the platform.

But there's other things that also could be going on and other inferences people may make, and this is where the different theories sort of come into play. What if it's high variance in quality? So the Sun paper that was cited was about thinking about there's product differentiation. So if it's high ratings you want low variance, but if it's medium ratings, you may want high variance because you think you may know if you're a good fit for the product. What if it's just kind of spotty quality and you're trying to figure out they either get it right or don't get it right. So I think these are the different factors that are going on.

A couple of concrete suggestions. We could think a little bit more about what would you do with this? So a platform could use this as a signal to sort of change the way they're flagging reviews or identifying stuff or it could just be, hey, this is something you want to tell users. The information treatment's a pretty strong one. Essentially it's kind of putting everyone into the room, letting them see this talk and then say now which thing do you kind of trust? It's a heavy way of telling people which reviews are more or less trustworthy. I think both could be useful. There's a little bit of a whack-a-mole problem on if you start having platform policies around that, people may adjust the types of reviews that they're leaving, which may make it challenging to do. So It'd be helpful to see a little bit more in the relationship between variance and spam.

So the standard deviation is interesting, but I think it would be nice to dive into more types of distributions and figure out what exactly is it that's triggering somebody to get suspicious. Think a little bit more about what the other factors are that affect variance and how that would affect the interpretation and the predictions. I thought in the data you could probably get a little bit more at what these competing predictions are and what you would see at different ranges of, say, average rating and distribution. More direct on the beliefs of users. I was interested to get a sense of whether people are well-calibrated on how many fakes there are and exactly what the magnitude of that is. So it's slightly different from the way that you've run the experiment, but I think it'd be a nice compliment for that. So let me just stop there for a time. Thank you.

Speaker 33:

So I know there's a bunch of-

Mingduo Zhao:

Thank you for fabulous discussion by Mike. Yeah, sorry.

Speaker 33:

I know there's a bunch of marketing papers that find that you end up with a U-shaped distribution of reviews in a lot of situations because people are more likely to write a review if they have a more extreme experience. I'm wondering how you're differentiating between that and fake reviews.

Mingduo Zhao:

Thank you for the question, it's great. In general, we're comparing between products in the sense that more or less the distribution are like U-shaped, but in some scenario, for example for the plot that Mike just show, they're more likely for the product who purchased fake review will have extremely U-shaped or whatsoever. Basically people kind of update their beliefs in the sense that they are not perfectly sure whether the product purchasing fake review or not, but if they see some U-shape or dream U-shape probably will have put more likelihood on their posterior about whether this product purchasing fake review or not. But there could be competing theory or competing reason why that's the case, but just people might slightly update their belief in certain way.

Speaker 34:

I was wondering how you deal with experimental demand effects and how this wasn't sent to us.

Mingduo Zhao:

Yeah, nice. Thank you for the question. The version you saw it in the paper right now indeed have that concern and we run one with exactly the Amazon interface and then we still see some effect is on appendix. But thank you for the comment.

Speaker 34:

So I have two questions about ones. So how much of dispersion might just be explained by fraction of reviews that are one star and why wouldn't I as a competitor want to leave fake one-star reviews on my rivals products?

Mingduo Zhao:

Yeah, thank you. Thank you so much for the question. Sorry, I don't know all of you so sometimes I don't know the names, but thank you so much. And my comment for that, this is a very good question and the version we run, the later environment we run, basically hold a one star to be constant. And also one for our thought is, we still see that effect by the way, after we hold a one star constant. But you still could argue whether you want to fix two star to be constant, right? This argument could go on and on, but I only have four degree of freedom for the distribution because they need to sum as one, right? But sorry. But I would say if people emphasize a lot on one stars, that's also emphasized on dispersion or discrepancy. Standard deviation is not the only measure to describe higher dispersant, like product reviews.

Like, have a lot of one star, at same time have a lot of five stars. Another way of describing it. And for robustness check we did, if we use the sum of five star and one star will hold the average to be constant, we still got that effect. So there are just different ways of describing the same thing basically. You can use standard deviation, you could use some of five star and one star, you could use any measure you like to describe the discrepancy. That's the reason why discrepancy is a more general term to describe the situation instead of we are so pedantic about just using standard deviation only metrics. Yep. Thank you for the question.

Speaker 35:

If you're going to have fake reviews for yourself, you might have fake reviews which increase the average as well as the standard deviation. So if I have a choice between two suppliers, one with a higher average and higher standard deviation, the other one lower average and lower standard deviation, don't you think I'll still go with the average?

Mingduo Zhao:

Thank you so much for the question. For the version we had kind of illustrate the point and we want to actually see the trade-off between the average and standard deviation. Basically we'll come up with nine distributions. In the paper some of average of 4.0, some of average of 4.4 and some of average of 4.8, each of them have different distribution. We do observe sometimes people are indifferent from two distributions that have different means at the same time have different distributions. But we want is triple the result to some numerical trade-off between the mean and the variance because we feel extremely context dependent because that depends on what actual average you have, what actual standard deviation you have. It's not like a linear extrapolation, but that's a very good question. We thought about it, but yeah. Thank you.

Speaker 36:

Thank you organizers, FTC for this great conference. Diego Aparicio from Spain. So I'm here to talk about concealing prices and what is it really concealing prices, right? So suppose you just, you're moving to a new place with your partner and so you come up with this beautiful ceramic table, you're ready to buy, right? You take out your credit card and let's buy it. Okay, but wait a second, I actually can't buy, I actually don't know what the price is of this item. Right?So to know the price, you got to click inquire and actually send an email to this firm to get the price. And then you start wondering as a consumer, as a manager, as a business person, does this make any sense? Why do I want to make it difficult for consumers to know the price and to buy?

This is what we're going to study in this paper. Now I tend to think that 95% of the things that we see out there, we see the products and we see prices right alongside the product, either online or offline. Now there are some firms trying different stuff. So you want to buy the sofa, well, you got to fill out this form and get the price. Some other luxury firms for instance, Chanel, most of the price are online. But for this most exclusive item, you actually have to set up an appointment with a boutique store, and once there you will be able to see the price. Even on Amazon, there are instances in which you have to click to see the price. Well firms are doing this and there's some debate out there and there's a lot of articles, a lot of speculation about whether you should do it, you shouldn't do it.

Endless Q&A's online from different users, even articles in New York Times Forbes about it. But I would tend to think that these views are backed by intuition, rather than systematic research. What are these intuitions saying? Well maybe there's a sticker shock, prices are painful so you better not show it. But on the other hand, well own the price conversation, be transparent and so forth. Some of these cases that we saw were from the online setting. But you can also think of these happening in the offline store. And this will also try this in the field. So you can have a situation where you go as a customer, you go to the store, you see the products on the shelves. But as a firm, what you can do is ask yourself, should I delay the presentation of the price? Okay.

And so from an experimental point of view will be considering these two situations. And I guess conceptually it is possible to first show you the price and then the item. But this is not a situation that we're going to analyze here. So we'll try to argue, okay, what's the literature saying about this topic? And I'll try to argue that it could go either way, it could go either way and it has not been directly studied or try to show some exploratory field experiments. And then we'll go to the lab to test a theory and validate the field.

So if I start thinking of this topic from a theoretical point of view, I will go and think, go back to Glenn Ellison's paper on price obfuscation. And these papers are saying, well it could be beneficial for the firm to make it difficult because in equilibrium you could charge higher prices, higher margins. Now when we go to the empirical side, there's a lot less evidence and most of the evidence relates to what we call the drip pricing. And I see some of the faces from the work here. [inaudible 06:30:37]. And some are saying, well if you make it difficult to know the fees, the add-ons, it could backfire in terms of reputation, reviews. But there's some evidence saying that it could actually benefit the firm to make it difficult to not show up, be upfront with the fees and so on. There's a cherry paper, right, it's a beautiful paper showing that just by not showing the taxes, this has meaningful effects on sales.

But again, kind of conceptually I think this is different from two ways of price obfuscation. So Michael before talked about the car setting. And in that setting we tend to think that maybe it's because of negotiation, haggling, price discrimination. So the price will be creating interaction and the other form of price obfuscation is more on what we call the junk fees, the junk fees or the add-ons, the drip pricing. And again, I think this distinction matters because some of these papers, the drip pricing papers saying, well consumers sometimes have a hard time calculating things. But now if I think about my phenomena, well there's nothing to miscalculate because I haven't told you what the price is.

So we'll go to the field, collaborate with a consumer electronics retailer, do an experiment in the online store, the physical stores, and with their email campaigns and then we'll go to the lab. Obviously the field matters because it will allow you to say, okay, this has meaningful, sort of high stakes real world purchases and this matters. So let me go into a little more detail perhaps on the online store. And so the first thing that we got to ask as we search is how are we going to operate to analyze the price delay? Because you could do many things. You could fill out the form, contact someone.

So here we have to make a decision. In our case, what we did was show prices with or without prices, products with or without prices in the category. So let me give you a figure of this. Right? So we will split the traffic randomly and some of you will see the price, some of you will not see the price. So then we can just compare conversion rates, purchases and so forth for the online store of these materials. And what do we find? Big boost in sales by removing, no, by delaying the presentation of the price. These effects are substantial. The people are talking about plus 20% increasing in sales.

Then we'll go to the physical stores and we also have to make a tough decision as to what is a price delay. We could make it harder as to hide the price tags. Sometimes you go to the stores and you have to open the item to see where the tag is. Ask someone, check a register. In our case, we decided to print new tags without the prices. And that's our manipulation in terms of delaying the price. So we'll do a crossover design here is not as perfect perhaps as the online store because we cannot randomly assign a customers and so forth. But what we do is some stores will be on the same treatment, some stores no treatment at all and some stores on the same treatment over time.

So we have four conditions. One is all prices will be delayed in the category and controlled business as usual. And then we thought it could also be interesting to see what happens when you hide the prices of the cheap items or of the premium items. Okay, so we'll analyze these four conditions. Then we estimate a fixed effects model, controlling for store category dates, et cetera. And what do we find? Well, let's consider the first delaying all the prices in the category. Again, big boost in sales in the store in those categories. And this is consistent with the online store experiment, which is nice. What happens when we hide the prices of the premium options? Big boost in sales as well for the retailer. So this is consistent with what some managers think in terms of the sticker shock, hiding for the premium items, you want to delay that. And for the cheap items, actually it backfires. The firm is worse off by delaying the presentation of the price.

Now with the final experiment with an email campaign, and we'll send emails to users, again, two conditions, prices or without prices similar to the online. It's kind of setting you open your email, you either see the prices or you don't see the prices. And the firm is using the email campaign to communicate promotions. So these are deal settings. It's kind of situation in which hey, here's an amazing deal but I don't tell you what the price is and here's an amazing deal and here's the price. What we find again is a decline in sales. So the firm by doing this now is worse off. So now the story is getting perhaps a little complicated, right? So what we have seen so far is that the timing, the price should be an important question from a theoretical point of view, go either way and from a managerial point of view as well.

The second thing is that it's economically impactful. The magnitudes that we're finding is like plus-20%. So this is consequential for the firm. And then it makes us wonder what is driving these results because it could go both ways. So in some settings all prices, high prices, boosting sales, whereas in low prices or promotions is declining. And so then we'll turn to the lab to several labs to test a theory and validate what we found in the field. What is the theory about? Well the theory has two building blocks. The first is that a delay amplifies price expectations. And the second one is we tend to be pessimistic about prices. When we don't see prices, we tend to think that prices are going to be high.

And so we have two predictions kind of. When the delay makes us think that prices are going to be very high, it could actually be that we expect the price to be so high that the real price is lower, so I'm more likely to buy. I thought you were going to kill me with the price, but it's actually not so bad. Whereas in a situation which if we think that this is leading us to a very, very cheap price and it's actually the price is higher than what I thought, then I'm less likely to buy. And so when we can think through the lenses of these three in the field is in the low prices or the promotion, these settings in which we can infer that prices are going to be cheaper, whereas in the other settings we can infer prices are going to be expensive.

So let me show you two pieces of evidence with respect to the theory. The first one is about expectations being pessimistic. And so we go to the lab and we ask participants to give us price estimates for a bunch of different categories. And then we ask them if prices are not there, do you think it is because they're higher or lower than what the prices just told us? And 40% of the people think prices are going to be higher than what you just said. Again, we first asked them for their price estimate, what do you think is the price of a coffee maker? Now if you don't see the price of the coffee maker, why do you think that is? So this is the first building block.

And the second one is about amplifying what this delay does. So now we do a more sophisticated now knowing which we have three contextual conditions baseline in the sense of business usual, all prices. Premium setting, the premium setting of all products being very, very expensive. Or cheap setting. Kind of deals and cheap options. And so these are three times two. So we ask participants to select the preferred alternative. And again, remember our predictions are that well in the settings of baseline or premium, right, the context is kind of premium, well we expect higher prices and therefore I'm going more likely to buy once you show me the price. Whereas in the cheap setting I expect a lower price and when I see it, actually, it's higher than next.

So the first thing that we establish is validating the findings of the field, baseline and high premium items being more likely to buy, but it backfires on the other setting. And then this effect is mediated. But what do you expect the price to be, right? So I guess from managerial knowledge, we put our marketing hats, what do we do with this evidence? Well, I guess you can first understand, know what your expectations are of your customers, knowing that you can manipulate those with the context. And then you can think of doing price delays when you can bring processes of prices. But also even if it pays off, you could ask yourself, do I want to do this? What are the implications of doing this? Are the extensions short-term, effects, loyalty? Is this going to hold up over time?

And some final thoughts. Well it also opens up the avenue of thinking of personalizing the delay, personalizing this. We're in an area. I may want to end with two kind of policy questions and I don't have the answers to these questions, but I think they are very interesting. One is that if I think from a morality point of view, which are the customers that are over buying or under buying and what is this doing? What is this manipulation doing with those customers? And I guess the second question is do we think that this price delay can be a beneficial friction? Because when we think about the junk fees, even the word, denotes that this hurts the consumer, but what if the delay actually makes consumers to think more before they buy? Thank you very much.

Speaker 37:

Okay, thank you very much. I don't have slides, but I have thoughts. So first of all, I think this was really an excellent paper, interesting first comprehensive look at this pricing strategy that we're seeing more and more in the marketplace. This paper was about price transparency. So first let's make sure we're all on the same page about what we mean. What is price transparency? In general it means that buyers have information about sellers prices in the marketplace. And why is that important? Well, it allows buyers to predict and judge the relative appeal of one firm's offer to a competitor's offers in the marketplace. The FTC and other consumer protection regulatory bodies have long been concerned about price transparency. For example, the FTC has proposed rules that we heard about earlier prohibiting businesses from misrepresenting total costs. For example, by omitting mandatory fees from advertised prices, banning hidden or misleading fees, including an entertainment and hospitality and e-commerce and other regulators in other countries have done the same thing.

The Canadian Competition Bureau outlawed drip pricing, which they define as the separation of mandatory fees or surcharges that aren't local or federal taxes. And then in other regulation, the hospital price transparency has been required in the U.S. since 2021. Hospitals have to provide clear accessible pricing information and today we learned about cars and funeral pricing as well. So if we look at the regulation around this and the academic literature, as Diego noted, much of it is focused on when transparency is absent because parts of the price are unknown. For example, fees and surcharges. And some of this work, including my own on partition and drip pricing, has looked at mandatory fees and other has looked at optional add-on fees. And this body of work has shown that consumers react differently when total prices are transparent, either because all-in pricing is used or if it's optional fees because they're fully disclosed up front versus being dripped.

And as Diego pointed out, there's also research about tax salience, consumers purchase behavior even for surcharge we all know about in the U.S. that there'll be taxes at checkout is different when that surcharge is more or less transparent at the time we're making our purchase decision. Now the hospital legislation and the funeral price list that we learned about, they're different and they have to do with about knowing the whole price versus not knowing it. And Diego's paper really fits in the spirit of that regulation when a whole price is either made transparent or not. So I really enjoyed Diego's presentation and reading the paper and I also very much appreciated the approach of demonstrating the core effect with field data and then testing process with follow-up lab studies and tests of moderation. So as we just learned, this research shows that in general, whether holding back on disclosing prices benefits retailers depends on consumers price expectations. And this begs the question, where do these price expectations come from to begin with in this kind of context? Or even in other contexts involving fees and surcharges. For example, are consumers more upset about fees and surcharges, the ones they notice at least, when those fees are not expected or they're larger than expected. In the current context, we can wonder whether consumers expect to see this practice in certain situations or for certain types of products. Diego talked about negotiations, maybe we expect not to see the price when we expect that this is a price that we negotiate for. Or more generally even for products that tend to have fixed prices, for a given consumer, when do they think this practice is more prevalent for certain brands or certain products high end versus low end? And that's probably what creates the price expectation. I also wondered about how this work relates to the finding of Kokomar and others JMR 2015 on the neural and behavioral impact of price primacy. I know you talk about that in the paper.

They showed the different parts of the brain are activated when consumers see the product and then price or vice versa. They found that if you view products first before price, then evaluations are more related to the product's attractiveness. But if you see prices first, then consumers are thinking about the worth of this product and that's really what drives their purchase decisions. You didn't examine cases where consumers see the prices first because you said that doesn't really happen in the marketplace, but maybe in some context it does. If you're shopping online and you screen for a certain price, then maybe that price is sort of the first thing that comes to mind. And so in this spirit, if showing prices later makes price comparisons or assessing worth more difficult, I wonder if consumers end up being more price sensitive in those cases, whether it leads them to buy more expensive products. And I wonder if you see any differences in what types of products are bought.

Or another related possibility is that order might lead to different mindsets, one of either pain minimization or quality maximization or maybe consumers who have those different mindsets react differently to whether price is salient or not. Your findings make clear when this benefits and when it harms sellers, but more generally we can think about three parties we want to know about, sellers, consumers and competitors. And you hinted about consumers at the end, whether this practice harms consumers or whether it helps them, whether it affects some consumers more than others, maybe more vulnerable consumers end up buying products that are expensive for them. We can also wonder what does it do for competition? Prices aren't only transparent to consumers but also to competitors. So how does this practice affect competitors price setting and also whether the choices that consumers make when they see prices immediately versus later, how does that affect the competitive nature of the marketplace?

Another thing I wondered in a broader sense, what are the implications of your findings for price transparency in places like healthcare, where in U. S. where this regulation has been put in place where patients should know total costs in advance. That's a complicated context where there's different prices for insured patients versus other ones. My understanding is that compliance among hospitals has been low. So that might provide an opportunity for further tests of your theory to compare what happened before and after. And one very quick last comment. You discussed at the very end that the popular advice to companies is to eliminate all friction. And you said maybe here there's a benefit, that friction could be helpful. And I think more generally, if we think about consumer protection...

Speaker 38:

... and could be helpful. And I think more generally, if we think about consumer protection, friction is often there to protect consumers. So for example, if we think about products out there with embedded finance from non-financial firms that skirt regulation, friction is reduced, but is the consumer better off? So it might be interesting to think more generally about different costs and benefits of frictions to different parties in the consumer journey. Thank you very much.

Ken:

Diego, thank you for your evidence. I've been teaching price salience as a consideration for many years and I've been needing more evidence to inform that for a long time. The thing I kept coming back to during your presentation is that in the consumer search literature, there's this classic distinction between simultaneous search and sequential search. And when you remove the price or delay it, you disable simultaneous search on the price dimension, which is probably the easiest product attribute to compare in many cases. And I was curious if you had a sense of what the search theoretical literature predicts when you disable that option?

Diego:

Ken, you can now use the paper for your classes. So I think that some of these papers on search cost and obfuscation tend to predict that by introducing this friction you can charge higher prices because I know at some point you will stop searching for the next best alternative. But as far as I know, those are theoretical papers. From a empirical point of view, I don't have any search evidence or anything I can say about how people are behaving in the website.

Ken:

Yeah, thanks.

Speaker 39:

So, I was wondering-

Ken:

Hey.

Speaker 39:

... so I was wondering whether you've looked at spillovers on other products. And the second question would be, consider this alternative theory. You consider a product if you, whatever, like the picture or it looks nice. And now with cheaper products, you wouldn't even look at them because they don't look nice, for example. But the more expensive one, once you've looked at it, you're more invested and now you have loss the version and therefore you're still buying the product even if the price higher.

Diego:

Great. Yeah, so we could examine that kind of theory, but we don't really. It would be consistent with these findings. For the premium options, you want the customers to start to love the product and then you raise the willingness to buy, which won't happen for the cheap options. And then what was the first question about?

Speaker 39:

I forgot.

Diego:

Oh, if you forgot, I feel better.

Speaker 40:

Great. As I was listening to your presentation, I was thinking about the cost from the consumer's perspective and it evoked this trade-off between incurring a cost in order to get that price and the sunk cost effect of that. And so I was thinking when that trade-off might be beneficial, and I wondered about the cost on the seller's side. So if the selling costs are significant, would they be willing to get rid of customers or ignore customers who aren't willing to incur the cost to get a price draw so that they evoke the sunk cost effect and have a better lead that they can use?

Diego:

Yeah. So I think there are many avenues in what is said, the consumer costs of acquiring this information, the screening costs as well. Because the firm, it is costly to do these things depending on the setting, but it's costly. And then if you have many customers entering the store just to know the price, it is more costly also for the firm because you have to filter out the right audience. Whereas if you are upfront now, then only those that are willing to buy above this price will enter. So interacts with that, but it's not something that I explore here.

Dan Wood:

We're going to try to get this Panel started.

Okay. So I'm Dan Wood, I'm a Deputy Assistant Director in the Bureau of Economics, and I have the pleasure of chairing this session with four wonderful papers in it. I think they're all very interesting. So without further ado, I'll turn it over to the first presenter, Mimansa.

Mimansa Bairathi:

Hello. Good afternoon everyone. Thank you to the organizers for letting the paper be a part of the schedule today. So today I'll present a paper on social media and its cross-platform spillover effects, particularly in the context of the music industry. So social media platforms, specifically TikTok, which is a platform where users go to look at music videos, allow users to use songs as background music in their videos. The question is, what's the effect of these songs, which are used in TikTok videos, on consumption of these songs on other platforms, particularly streaming platforms like Spotify and Apple Music? So on one hand, TikTok would actually help consumption on other platforms. Users might become aware of songs while looking at videos on TikTok. They might search for them and then listen to them on Spotify. On the other hand, TikTok might actually harm the consumption on other platforms. So if songs are already present on videos in TikTok, users might not feel a need to go to Spotify and listen to them. So it's important for artists, platforms, and record labels to understand whether social media hurts or actually helps consumption on other platforms.

In this paper, we examine does presence on social media help or hurt consumption on other music channels? And the channel that we look at here is Spotify because Spotify is the most popular streaming platform in the world with more than 30% market share. We then ask whether social media aids music discovery, and to understand discovery, we focus on Shazam, which is the most popular music discovery app in the world. We also look at heterogeneity across songs, but in interest of time, I'll not discuss this today, but I'm happy to talk about it offline.

So the context for this study is Universal Music Group's withdrawal from TikTok. At the end of January this year, Universal Music Group withdrew its entire song library from TikTok. What that meant was Universal's songs were no longer available to be used in TikTok videos. Pre-existing TikTok videos which were already using universal songs were muted. So this exit provides us with a nice quasi-experimental setting such that users on TikTok were no longer able to access universal songs on TikTok for reasons which were unrelated to their preferences. So this exit enables us to causally understand the effect of universal songs leaving TikTok on their consumption on other platforms. And like I said, the platform that we focus on are Spotify and Shazam.

So using Soundchart's API, we collect data on worldwide Spotify streams and Shazam discovery of songs at a daily level for a period of five weeks before and after the exit. The focus of this study would be on the average treatment effect on the treated, so the ATT. So treatment group comprises of those Universal songs which had at least one TikTok video prior to Universal's exit from TikTok. And the control group comprises of Sony and Warner songs, which had at least one TikTok video prior to the exit. So here we are focusing on the big three record labels.

Now, one concern with this design could be that in the absence of Universal songs on TikTok, users on TikTok might substitute towards Sony and Warner songs due to which exposure of Sony and Warner songs might increase, and we might end up overestimating the effects. So to get around this possibility, we construct an alternate control group of songs. These are essentially Sony and Warner songs which were on TikTok at some point of time in the past, but had no TikTok videos during our observational study. So essentially, because these songs had no TikTok videos during the period of observation, we know for a fact that users were not substituting towards them. We did check and these songs are not songs which are not popular on other platforms. They are typically popular on other platforms like Spotify.

So to understand the effect of Universal's withdrawal from TikTok on Spotify streams of Universal songs, we use a synthetic difference-in-difference estimator. So this plot shows the treatment effect estimates over time relative to the period right before Universal's exit from TikTok, which is shown using the vertical dashed line. So pre-treatment, all the estimates are zero. So providing some evidence for parallel trends. Post Universal's exit, we find that the Spotify streams of Universal songs actually decreases relative to Sony and Warner songs. On average, this decline is about 553 per day, which amounts to a 2.2% decrease in Spotify streams. Now, these results are also robust to using this alternate control group of songs, Sony and Warner songs, which were not present on TikTok.

So broadly what we find is as a result of Universal's exit from TikTok, the Spotify streams of Universal songs actually reduces. Question is why is this happening? So one reason could be that TikTok acts as a channel of music discovery. While browsing videos on TikTok, users might come across songs they were previously unaware of due to which they might try and identify these songs somewhere, and then later on listen to them on Spotify. If indeed TikTok acts as a channel of music discovery, we would expect that as a result of Universal's exit from TikTok music, discovery of Universal songs would suffer relative to Sony and Warner songs. And the way we measure it is using discovery via Shazam.

So Shazam is a music discovery app. It's the most popular music discovery app in the world. And what Shazam does is in case there's a song playing in the background or on any other app, you can just fire up this app to identify that song. So again, using synthetic difference-in-differences, we-

Mimansa Bairathi:

Again, using synthetic difference-in-differences, we find that as a result of Universal's withdrawal from TikTok, the Shazam discovery of Universal's songs actually goes down. This decline is quite pronounced. It's about a 10.6% decline. If we use the alternate control group, our findings are basically unchanged. So, essentially, what we find is, as a result of Universal's withdrawal from TikTok, the Spotify streams of Universal's songs declined, and it is likely a result of a reduction in discovery. So, these findings are important for platforms like Spotify. What it shows is, to some extent, consumption is contingent on discovery elsewhere. For individual artists and labels, it means that they should actually consider cross-platform spillover effects in their distribution and pricing decisions. And finally, for policymakers, it's important to consider these platforms not independently of each other. Thank you.

Hema Yoganarasimhan:

How do I make this work? Is there a computer here?

Speaker 41:

Can someone with experience point us to it?

Hema Yoganarasimhan:

Oh, okay. Okay, sorry.

So thank you, everyone. Thank you for the organizers again for taking the paper. [inaudible 07:24:22]. I have to stand by the mic. Okay. I'll not move from here. Thank you. Okay, so this is a very be interesting presentation because the question we are going to be looking at is exactly the same as what came in the talk before. So I have to thank Mimansa because she really set up the question and what they did and so on before. So let me actually go through what we did and I'll talk a little bit about what they have found. I should also tell you that there is one more paper, which is looking at the same question. Okay, so there are three papers all coming to different conclusions. So that part will be fun. But let me start by showing you what we did and talk a little bit about why what might be going on. I don't have all the answers.

Okay. So this is, I should say joined work with Maggie who is here and Elie Ofek who couldn't be here, both of them at Harvard. Okay, so again, I don't need to set up the question and all the reasons why we should study this because Mimansa already did it very well. So the things we are looking at is, as I said, very much the same thing, which is when UMG and TikTok had this big feud and they pulled all their library or silenced their library on TikTok. What happened to overall streaming demand on other platforms like Spotify and YouTube? And whether there are any heterogeneous effects across tracks. We also try to do something else, which is if we can quantify this a little bit and say something about the economic impact. It's going to be approximate because we don't have all the details of their licensing agreement.

Okay, so this is a beautiful picture that Maggie actually made. So this is what we are doing. And we actually look at this, use the same control group, which is Sony and Warner. We are then going to look at what happened to the streaming demand for UMG tracks on YouTube and Spotify. So here's the main results. So what we do is we use a logged demand model and we ... Can I? Okay, it's working. It's a simple diff and diff model. You can see that the overall effect on demand on both Spotify and YouTube is null.

This is true even if we do a bunch of other robustness checks, which I'll get to. So okay, that's the first result we find. So we don't actually find that there is any big impact on demand. The next thing we can do is we can look at availability of tracks on TikTok prior to this dispute. So what happens if they were actually here prior to the dispute? What you'll see is that so streams which were on TikTok prior to the dispute, generally tend to be more popular from well-known artists. If you're Taylor Swift, they've probably taken your songs and actually used it on TikTok. If you are a no-name person, then probably not. So there's quite a bit of heterogeneity. So what we see is that tracks which were used on TikTok prior to the dispute, actually, we do see a positive effect.

So we see that for these tracks, there is a substitution effect, which means that their availability on TikTok is actually harming their streaming demand elsewhere. Whereas tracks which were not on TikTok, prior to the dispute, actually see a negative effect, which means that they're actually being helped by some form of discovery on TikTok. And actually, I'm not going to show you the results in detail. But we find that much of this effect is actually driven, this negative effect, is driven by artists who have some kind of partial presence on TikTok. "I heard one song of yours here. Maybe I go and listen to another song and discover it." So we also then look at the economic impact and I'll tell you a little bit about what happened. So you can look at ... This stopped around May 1st. They reworked their licensing agreement and things are all happy. Now, you can listen to UMG's music on TikTok.

But if this had gone on for a year and these treatment effects were actually true, so we tried to do a little bit of economic analysis. So removal of UMG's music, what you can show is that because the baseline demand for the tracks on TikTok is actually much higher, you can show that overall, removal of this music led to about $900 million in revenue gain. Which means that by being on TikTok, they were losing about this amount of money. What eventually did happen was that on May 1st in 2024, they re-did this agreement and actually TikTok agreed to increase compensation, which is consistent with this overall positive effect, because there is no reason for TikTok to increase compensation if actually UMG was benefiting. Thank you. Is it three or five? Okay, thank you. It shows both, so I wasn't sure.

Okay, so this, when we first wrote this paper and Maggie and Elie presented it and so on, this was around May or so, and we had this version in June, we got an email from Jura and co. saying they were looking at the same question and, "Hey, we find positive effects." I was like, "Okay." And then we posted the paper in June. Then just I think a couple of weeks back, Anja and Mimansa say, "Hey, we looked at the same question and we find negative effects. What should we make of this?"

So apart from me being the queen of null effects, which nobody then wants to publish, why are we coming to different conclusions? It's like looking at the same question, why is this happening and what should we take away? So we spent some time thinking about why this might be happening, and I think this part is important. So there are some differences across the papers. So there is differences in data sources and length. So example, we have about two 40,000 tracks in our data. I think Mimansa and co. have about like 60,000 and all their data mostly comes from tracks. Okay, let me show this. Tracks on TikTok prior to the dispute, whereas we are going to look at ... Stand in front of the mic. Okay, thank you.

Okay, all the tracks are on TikTok. So this paper we must have presented, they're only looking at tracks on TikTok prior to the dispute, whereas our original analysis looks at all the tracks. But for tracks only on TikTok, we are actually finding a positive effect mind you. The other papers are using synthetic DiD. And what we actually find is that parallel trends are largely satisfied. They're not a hundred percent, but if there is a trend, it's about 0.3 to 1% of the treatment effects. So I don't know if I would try matching at that point. The level of aggregation is also different because if you want to match things, then you basically start aggregating the data and then you lose a lot of power in the process. I'm not sure if that's a good thing. One other thing which I think Mimansa will point out, is that maybe we are using the log of the dependent variable and that can cause some problems which has been shown, but we use levels.

So we do multiple things to stress test this, and there are various tests out there. So you can rescale the dependent variable, you can use levels. But nothing that we do makes changes, this positive effect to negative. So it probably is the length of the data. We have a longer panel as well as the size of the data. You can see that the effects are still positive. Or for YouTube, they're often null because YouTube data is much smaller than Spotify. YouTube data has more missing data in some ways. So this is the level of analysis by availability on TikTok as you see. Again, we find null effects. The only thing which makes me believe are positive effects a little bit more than the negative effects I would say, is that what happened eventually was that TikTok agreed to increase compensation, which is consistent with the fact.

Otherwise, There is no reason for them to increase competition. So finally, in terms of takeaway, what do we find? The details are all in the paper. You guys can go read all the three papers. There are three of them. So overall, we find a null effect. But for tracks on TikTok, we find a positive effect. Not on TikTok, we find a negative effect. But I think more broadly, for me at least, and my co-authors, the lessons, there were some bigger lessons in terms of null results are maybe okay. Also, there are many researcher degrees of freedom in empirical papers and we can all arrive at different conclusions. But it seems like having simpler and more robust models and stress testing them, seems to help have a little more confidence in the results. That is pretty much it. We are hoping to chat a little bit more and understand whether we can all arrive at one conclusion, instead of at least two to three different conclusions. But that's very much it. Thank you so much.

Youngeun Lee:

Okay. Well, first of all, thank you for having me here and I'll stay close to the talk. I'm Youngeun from Boston College. Okay. And today I'll talk about my work on shrinkflation. So first of all, what exactly is shrinkflation and why should we be concerned about it? So let me start with an example. In 2014, Chobani reduced its package size from six ounces to 5.3 ounces, while holding their item price fixed at a dollar and 29 cents, which in turn resulted in an increase in unit price by 10%. So product downsizing shrinks the package size without a corresponding change in prices to indirectly increase the unit price. And consequently, it's referred to as shrink inflation. Now why is this form of shrinkflation concerning? Well, it's mainly due to the lack of transparency in the changes that are made in these products. So if you look at the two product images on this slide, it's very difficult to see the difference among the two products.

So for a casual consumer, the item price would be identical, and also, the non-price attributes would also look very similar. So this may mislead consumers who are unaware or uninformed of these changes to purchase the downsized products, where in fact, they would be paying more per unit and receiving less for the same item price. So in this paper, I first systematically document this form of shrinkflation in consumer packaged goods for a 13-year time period, starting from 2006 to 2018. Then I move on to understanding how demand responded to this form of shrinkflation. So with the first goal in mind, I begin by documenting patterns of downsizing in products. And so this proceeds in the four steps you see here. Now, I'm not going to go into too much detail. But basically, the idea from step one to three is to identify products just like the Chobani example that we saw before, where there was a reduction in package size without other non-price attributes changing. So they remained unchanged, and systematically applied this approach across products that are available in Nielsen IQ data during that sample period.

And as a result, I identify products with downsized sizes across a wide range of product categories, including over 60 product groups and 300 product subgroups, product modules in the data. Now, you may think that firms would reduce the package size by a small amount so that consumers do not really recognize the change. However, the magnitude of these reductions are actually quite large. So if you look at the distribution of package size reductions, at the median, a product was reduced by 11% in its package size. I observed these patterns of replacing a product with a smaller version throughout the entire sample period, which suggests that this is neither a new nor a novel strategy that firms are using. Now so far, I've documented how package sizes change.

Now the next real question is, as a result, how did prices adjust, to really say this corresponds to shrinkflation? So when estimating the average changes in prices, I find that package size reductions lead to a significant increase in unit prices for majority of the categories, where the median is a 9.3% increase. So if we think about this magnitude, it's economically meaningful as it's half the size of an average promotion depth that I observe in the datasets. And also, this represents a permanent increase in unit prices, unlike temporal changes or temporal discounts that we see in data. So how did consumers respond to the significant adjustments made in products? So to measure consumer response, I measure the average changes in quantity sales in two different measures, where in the X axis you'll see the items sold, the number of products sold, and Y, the total volume sold.

And I see that the majority of these estimates lie on the fourth quadrant. So to put this into context, it means that the results suggest there's very little changes in consumers purchase behavior, where they're purchasing the same number of products as before or slightly more, but not enough to really compensate for the loss of volume per product, which leads to an overall decrease in the total volume purchase, or really no change in total volume purchase. Now focusing on the magnitude, this is very small in terms of demand response and also, this is supported by package size elasticities that I separately estimate, which are centered around one, indicating relatively inelastic package size elasticities in terms of package size changes. Now, we might expect consumers to actually respond more to greater and larger changes in package size. And now if that is true, the package size elasticity on the Y axis and the degree of change on the X axis would have a negative relationship. So we would see a downward slope.

Now, if we look into the data, that is actually not true, suggesting that consumer's response does not necessarily increase if there's greater and deeper reductions in package sizes. So these findings are surprising because in lab experiments that study shrinkflation or just the perception of package size changes, they say that consumers would definitely negatively respond to these changes in package sizes, especially if it leads to an indirect increase in unit prices. So one hypothesis may be that consumers are unaware or inattentive of these changes that occur in the package sizes and prices. So I study whether increased transparency in these attributes would lead to greater consumer response in the setting of a state-level mandatory unit pricing policy in the United States. And so with these price tags, we expect that the salience of unit prices increase and thus, it will facilitate unit price comparison.

So seeing this significant increase in unit prices, they may actually switch to a different product or exit the category entirely. Now, in terms of the results, I do see a more negative response directionally. But the magnitude of this effect is very small, indicating that there's a very small differential effect across these regulatory regimes that have a mandatory policy and not, which suggests that the disclosure of this unit price alone is not sufficient to generate a greater demand response to shrinkflation, as we might expect to happen. So I'm actually pretty fast. So given what I've talked about, I hope I'll leave you with this. So shrinkflation has been employed since the past where they leverage the package size as part of their pricing strategy. These adjustments are actually non-trivial, though we don't really see a great demand response to these changes. Thank you.

Avinash Collis:

Hi, everyone. I'm Avi Collis from Carnegie Mellon, and today I'm going to talk about new paper we recently wrote on measuring the disutility of ads on Facebook. This is joint work with Nils Wernerfelt, who is in the audience, who is now at Northwestern, Erik Brynjolfsson at Stanford, and several other collaborators from Meta. So basically, what we are trying to do in this paper is measuring the disutility of ads on Facebook. So do ads create utility or disutility? So on one hand, we already heard from Alessandro earlier today, ads could play an informational role, matching buyers and sellers, reduce search costs. But on the other hand, ads could be annoying, they could have a cognitive cost, and we heard Alessandro's talk about potentially seeing higher-priced, lower-quality products and so on. If we want to measure the consumer surplus from online ads, on a platform, what would the ideal experiment look like?

So the ideal experiment would randomly switch ads on and off to different users. It would be a field experiment and not a lab experiment to avoid concerns of hypothetical bias, experimental demand effect and so on. It also needs to be a long-term experiment because previous research, for example, a paper from Pandora shows that it can take up to couple of years for consumers to adjust their behavior if you experiment by changing ad loads. So we basically stumbled upon this ideal experiment which was happening at Facebook. We didn't design this experiment. This is something Facebook was doing for a different purpose. So basically, Facebook, to give you some background about this experiment, big tech platforms like Facebook, Google, Amazon, and so on, they run thousands of A/B tests every day. Many of these A/ B testing platforms have a holdout group. And the holdout group is basically a clean control group which doesn't have any intervention, which they're testing.

The reason for having this clean control group is to look at the cumulative impact of all the interventions which are being tested over the long-term. In the case of Facebook, the holdout group, which they have in their A/B testing platform, is basically a group which never sees any ads. They want to look at this clean control group and look at all the interventions they're running and look at the cumulative impact of all these interventions. We exploit this holdout group and see what happens when users are shown ads or not shown ads. So what happens in our experiment is in 2013, Facebook launched this experiment where 0.5% of all users on Facebook since 2013 are randomly assigned to a holdout group. If you are in that holdout group, you never see any ads. And this is a continuously running experiment. So it's not like it was done in 2013, but this is continuing even to this day. If you go now and create 200 Facebook accounts, one of them would end up in this control group and that account will not see any ads.

So we have 0.5% of all users who are in this control group of no ads and the rest of the population who see the usual amount of ads. So what we do in our paper is we recruit a representative sample of users in the control group and the treatment group. So representative sample of Facebook users. So around 54,000 Facebook users. Around 40,000 in the treatment group where they see the usual amount of ads, and around 14,000 users in the control group where they don't see any ads. Both of these groups are quite well-balanced across a bunch of demographics. And also, if you look at the amount of time an average user is on Facebook, both of these groups are very similar. So there doesn't seem to be any attrition across people in the ads group and the no ads group. We also have weights, which weights, which Meta has on these users. Which modeling like how likely they are to take part in our study and so on, using all the information Meta has. And applying these weights, we get representative samples of Facebook users in both the control group and the treatment group.

So what we do is we measure how much people value Facebook in the control group and the treatment group by running incentivized online choice experiments. So basically, measuring willingness to accept to give up Facebook if you're in the control group or treatment group. Difference in valuations of Facebook across the control group and the treatment group should basically measure the utility or disutility from seeing ads or not seeing ads. So basically at the top of the news feed, users were randomly shown a message to take part in our study. And the experiment, our study happens on the app itself. So we did our study in 2022. So we have people, some of whom are in the experiment for nine years, some of who recently joined, and a lot of variants in between.

The choice experiment basically looks like this. It's a simple take it or leave it offer, to keep access to Facebook or give it up for a certain cash amount, which we offer for one month. This is incentivized. So users are randomly selected and actually offered money, if and only if once we verify that they have complied with their condition. So real money was on the table. So we select around 400 users randomly, pay out over $20,000 in total. So this is closer to a real preference valuation of Facebook. So let's look at results. So this is how the demand curve looks like for users in the ad group and the no ads group. On the X axis, it's the price which we offer to people. On the Y axis, it's the percentage of people who choose to reject that cash and keep access to Facebook.

So it's a downward sloping demand curve. What is interesting to find is that both demand curves for ads group and the no ads group seem to basically overlap with each other. The median valuation across both of these groups is around $31. If you look at median valuations across the ads group and the no ads group, and look at the difference in median valuations, what we find is the differences are not significant. So the point estimate is around $1 of disutility, but it doesn't seem to be significantly different from zero. If you do a simple calculation to measure the minimum detectable difference that our experiment is power to detect, what we find is the disutility from ads is of the order of around $3 per month or less than that. So that would be the effect size, which if it's smaller than that, we wouldn't be able to detect. But if it was bigger, then we would have detected.

So that is around 10% of the value of Facebook, around $31 from Facebook, up to around $3.20 of disutility, but it could also be zero. So we cannot tell that. If you look at time spent, users in the ads group spend around 9.4% less time. It is possible that they're spending less time because they click on the ad and leave the platform and make a purchase. We don't know the mechanism, but they do spend less time. But they still value Facebook at a similar level. We look at heterogeneity along a number of different dimensions, looking at heterogeneity based on time spent on Facebook, splitting users into low, medium, and high-tercile of time spent. No significant and user valuations of Facebook based on time spent. We also look at heterogeneity based on how long they're on Facebook. Again, no significant difference based on if you recently joined Facebook or if you've been on the platform for almost a decade.

We also look at heterogeneity based on region. No significant differences between the ads group and the no ads group. If we take the point estimates seriously, even though they're not significantly different from zero, what we find is the point estimates around $4 of disutility in the US, around $1 of disutility in Europe, and then 15 cents of utility in Mexico.

But again, none of these numbers are significantly different from our field. So to conclude, there's a lot of policy debate right now on are ads good or bad for consumers. But there isn't a lot of research putting a dollar amount on utility or disutility from ads. So our study aims to contribute to this space. I want to mention that we are just looking at the consumer surplus from ads by just seeing the ad on Facebook or not. If you want to do a full welfare calculation, there are other things happening like users clicking on the ad and making a purchase, and Alessandro talks about it a bit in his paper. Nils and co-authors have also research on welfare, which goes to advertisers. And also of course, there is welfare going to the platform itself. So we are just focusing on consumer surplus from ads on the app itself. So yeah, that's it. Thank you.

Speaker 41:

Okay. Would the presenters come up to the stage so we can do a brief period of questions. Probably five minutes since we're running a little behind schedule. Of course, you can find somebody and ask them questions during the coffee break. Thanks.

Speaker 42:

I have a quick question for Avinash here. Just trying to understand the treatment group here. So the control group, I understand there was no advertising going on. But the difference between the treatment and control is only ad exposure or other things that could be built on top of that? Or were they held out from all updates? Can you talk a little bit about that? Because that'll affect the interpretation.

Avinash Collis:

Yeah, so both of the groups are exactly the same except you see ads or don't see ads. So the treatment group basically gets shown the usual amount of ad load and the control group has no ads. Otherwise, the app is exactly the same.

Speaker 42:

And the ads, the click data is not used for any other services on Facebook? Because that could have some consequences.

Avinash Collis:

Right. So from what I understand, everything else is the same. Clicking the ad or not shouldn't impact the organic stuff you see on your news feed and so on.

Speaker 43:

Also for Avinash, I'm wondering if this might be something specific to Facebook. As far as I know, Facebook doesn't have a subscription you can pay to avoid ads, unlike say YouTube or many streaming services that have introduced an ad tier. So it might just be the fact that Facebook ads are less disruptive and less annoying to consumers than a lot of other advertising.

Avinash Collis:

Yeah, I think that could very well be true. On YouTube, I suspect the disutility would be way higher people. Clearly, a lot of people do pay for YouTube Premium. So you're right. So this is only looking at Facebook.

Speaker 44:

I got a question for Youngeun, I forgot what the sample was, but for shrinkflation, let's say yogurt, I was concerned about calories and, "Oh, they make it a little smaller. Happy to pay for it." So is there any difference between food items or non-food items? I don't know if you looked at that. Also, within food items like hedonic foods and non-hedonic foods.

Youngeun Lee:

So in terms of the heterogeneity across categories for demand response, I consistently see that regardless of whether they're food or non-food, there is small demand response to these changes. So I do have products like detergent or shampoo included in the sample. But there's not a clear pattern whether the categories determine whether consumers would be more or less responsive to these changes, at least in the sample that I have.

Speaker 45:

Hi. Yeah, I have a question for Youngeun as well. So I noticed sometimes, the consumer packaged good market sometimes make the package big. And so I wonder if you have a look at that to how prevalent it is. I won't be surprised if the result is different.

Youngeun Lee:

Yes, that's a great point. So the approach that I take does identify products with bigger sizes. But since I focus on shrinkflation, I just disregard that entire sample. But I do have a follow-up paper that looks more into the general trends in how these package sizes are changing. We would think that everything is becoming smaller because of all of the media press. But actually the package size does not necessarily decrease over time. There are categories where we see bigger sizes happening. So you're correct, your insights are correct. Yes, in terms of promotion, that also might be true because you would see marketing labels saying you have some percentage more for the cereal than the original one. Yes.

Speaker 46:

Thanks, I appreciated your talk and I have another follow up for Youngeun. So I'm wondering if the brands or the products that are engaging in shrinkflation are firms that have maybe more market power.

Are they doing it differentially for some of their products that are more the dominant products in the space?

Speaker 47:

... some of their products that are more the dominant products in this space. How strategic is this decision? Who are the actors that are choosing to use shrinkflation, which ones are not? Is it just because they know they have a very dedicated consumer following? Is that partly explaining the minimal demand response?

Youngeun Lee:

So, that's a great question that I wish I had an answer to. I could give you two examples that show it's not straightforward where the bigger brands are doing it or if everyone is following and reducing its sizes. For example, the Chobani example. All the yogurt now for a cup of yogurt is reduced to 5.3 for the majority of the brands, whereas for ice cream, pint ice cream, Ben & Jerry's still has a pint, whereas Häagen-Daz has a smaller size.

So to really understand what is happening in this industry, is I think we have to take a more case study approach when we focus on a specific category and know what the costs are for manufacturing these products, and then understand how competition plays a role there.

Speaker 48:

So, I had two questions. One for Avinash and one for the Hema-Mimansa team. First for Avinash, in your last slide you said that you might be underestimating the value of these ads on Facebook. The question being asked of the consumers in the experiment was, how much are you willing to pay for Facebook? Presumably that should include not only the exposure to the ads, but also what utility those ads are giving them, right?

Avinash Collis:

Yeah, you're right. I guess it's a question of how much. So, they may click on a product and then make a purchase and they get consumer surplus from the purchase. It's a question of if they are internalizing that into the Facebook experience or not. So I think for example, you may click on an ad and make a big purchase and get thousands of dollars of surplus. I don't know if users internalize that. So that's possible, but yeah.

Speaker 48:

My question for the other presentation was, there might be a big difference between well-known artists like Taylor Swift, or some little known artist in terms of discovery. So it's not surprising that in the case of Taylor Swift, probably there's no much effect because people probably know their tracks already. They don't have to be exposed on TikTok to know that.

The other side of the equation is even Spotify is a big entity. Hardly anybody is not a member of Spotify. So in that sense, people may already be multi-platforming between TikTok and Spotify. In that sense, there may be less effect on Spotify because of this change at TikTok.

Hema Yoganarasimhan:

This one don't seem to have good output.

Speaker 48:

Hold it, hold it.

Hema Yoganarasimhan:

Okay. So, I think it is true that tracks already on TikTok tend to be from bigger artists or at least bigger songs like more well-known songs. Those are the ones, they're actually truly worried about the cannibalization effect like that. Okay. It's not like I'm discovering Taylor Swift through TikTok. I already know her and if I listen to her songs enough on TikTok, then I'm not going to go consume it on TikTok. At least, that's what we find.

That tracks on TikTok already tend to be more popular, comes from more well-known artists in more advanced career stages, whereas tracks not on TikTok prior to this dispute, tend to be from smaller and less well-known artist. As you said, they tend to actually have more of a discovery effect. For them, the loss of TikTok actually has a negative impact on Spotify.

I think they actually, they're only looking at tracks on TikTok and they're actually finding a negative effect. So, I don't know. I'll let you speak.

Mimansa Bairathi:

We looked at effects based on also popularity more generally, and we find consistent results there as well. Of course, yes, I agree with you that on TikTok songs which are already present are likely to be songs which are already popular, but we looked at heterogeneity across songs more generally as well. So for instance, songs that were released more recently, which would be more in need of discovery presumably. Even for those songs, we actually find that the results are much stronger. So, they actually witness a larger decline in streams on Spotify as a result.

So regardless of whether it's based on popularity, whether it's based on recency, we do find consistent results throughout. So, that's not necessarily something that is actually explaining the differences.

Speaker 49:

So I guess this question is mainly for Youngeun, but thank you all for the papers. So jumping off of Elise's question about what type of firm strategy would rationalize the types of effects you're seeing. Apart from a case-by-case basis, I'm just curious, have you thought about what would rationalize in the specific case you are looking at? Why would Chobani shrink its packages if it can anticipate that demand would go down? What is that story?

Youngeun Lee:

That's a great question. So what I do observe throughout, I think I include this example in the paper. Usually firms when I talk to them, use this as a way of dealing with cost shocks or cost increases. So, they have to figure out whether they're going to increase the price or reduce the package size to deal with this cost. That was one of the reasons that they told me would rationalize them reducing the package size instead of increasing prices directly.

Speaker 50:

Okay. Well, thank you very much. We're running a little late, so we'll resume in five minutes. You have a chance to leave the room and do urgent things.

Ginger Jin:

Thank you. We have the last session, Online Platforms. Our first paper will be presented by Hangcheng Zhao from University of Pennsylvania. Thank you.

Hangcheng Zhao:

Hello everyone. Thank you so much for including our paper in today's program. My name is Hangcheng. I'm from the Wharton School. Today, I am very happy to share with you our work. This is a joint work with Ron Berman who is also from Wharton and also here today. The title of today's paper is Algorithmic Collusion of Pricing and Advertising in E-commerce Platforms.

The setting that we consider is the e-coms platform setting. Of course, there are many, many different examples. One of the most prominent examples that we are going to focus on today is Amazon. Suppose that we are the Amazon customer, and we would like to buy some kind of products on Amazon. So what we usually do is that we go on the Amazon, and we search the keyword that we'd like to buy. Then, we will see the search result page which contains both the sponsored products on the top and organic products.

So on this e-coms platform, the sellers make a two-dimensional decisions: what is the price for their products, and how much they would like to bid for those sponsored listings. Now in the market, there are a lot of third-party Amazon pricing or bidding software providers telling the seller that they are using AI technique to help avoid price wall to increase the seller's profit.

So a very natural question here is that, is that good or bad? Should the sellers be using these algorithms? More importantly, is that beneficial or harmful for consumers? Because traditionally, we would think that firm would have the incentive to increase the price, but higher prices hurt consumers and raises antitrust concerns. For example, recently FTC says that price fixing by algorithms is still price fixing. DOJ recently sued RealPage, which is the algorithm pricing software that allows the landlord to charge higher prices compared with the competitive prices.

So in this paper, we asked the following two research questions: How do reinforcement learning algorithms influence how much the seller would like to bid for those sponsored listing, the prices for their products, the seller's profit, the platform's profit because this different seller are selling through the platform and more importantly consumer surplus. Initially, I'm going to focus on the reinforcement learning algorithms. Later, I'm going to generalize to different types of algorithms that are able to change the price and determine their bid in real time at high frequency.

For the second research question, by now there is already a very famous result in literature saying that if the seller is using algorithms to charge the price, they can learn to charge higher prices compared with the competitive prices and that hurts consumers. We actually wonder whether that is always true. Can the algorithm pricing be beneficial for consumers or not, and is there any empirical evidence supporting that?

The plan for today. First, I'm going to talk about the contribution of our paper, the intuition, how these different algorithms operate and why the algorithm pricing can be beneficial for consumers. For the second part, we're going to use a combination of both reinforcement learning simulations, and we are going to compare it with a benchmark scenario with the algorithms to pin down the condition when this algorithm can be beneficial for consumers. For the third part, which is the empirical part, we are going to present empirical evidence, actually supporting that this different algorithms can generate beneficial effect for consumers. Finally, we're going to conclude today's talk and talk about the managerial implications.

For the contribution of this paper, by now there is already a very famous result in literature saying that if these sellers are using reinforcement learning algorithms to set only their prices. These algorithms can learn to charge higher prices compared with the competitive prices without algorithms. That hurts consumers and raises antitrust concerns. This is a figure from the very famous result in literature which tells us how these algorithms learn to collude.

The X-axis is the time while the Y-axis is the price in the market. As we can see, even if we start both of the firms from the very low competitive price because these algorithms always have some probability of randomly exploring some of the other prices within the strategy space. Gradually, these algorithms will realize that coordinates on higher prices will give both of the firm higher profit, but that hurts consumers.

So the contribution of our paper is that by adding another very important aspect, which is the advertising aspect, and we found that these algorithms still learn to collude, which means that it benefit the seller. However, what we found here is that the algorithm pricing can actually be beneficial for consumers and the condition that we need is that consumer consideration set is limited.

Think about how these algorithms learn to collude. What the previous literature tells us is that they want to maximize the revenue and they learn to charge higher prices. However, when we also consider the cost side which is the advertising part, they can also learn to collude on lower bid and that is a lower cost. Ultimately, this lower cost can pass on lower price and benefit consumers. So we actually need to empirically estimate consumer behavior and we use Amazon dataset, and we try to answer two empirical question. The first is that, is there any empirical evidence supporting that this algorithm pricing can be beneficial for consumers? The answer is yes, and we indeed find enacting interaction in fact between consumer consideration size and algorithm usage on prices. For the second part, we try to predict how popular this phenomenon is and it turns out a lot. For more than half of the electronic market, these algorithms can learn to price lower compared with the world without algorithms.

The first part is the model setup. The reason why we need a model is that we first need to understand the scenario what will happen without algorithms. As we discussed before, this is an e-com platform setting three different types of players, the platforms, the sellers, and consumers. For the platform, they randomly, they display the products into two types of position, sponsor position and organic position.

For the sponsor position, the platform is going to look at the bids submitted by different sellers and display the sellers who submit higher bids, who win these auctions. For the sellers, they make a two-dimensional decisions, what is the price for their products and how much they would like to bid for those consumer price? Finally, consumer make the purchase decisions and consumers are heterogeneous. For some consumers they consider less of the products, while for others they might consider more the products. They might stop anywhere within the search page.

For the next part, we're going to let the seller using these algorithms to determine their price and bid. Some background knowledge about what the reinforcement learning algorithms do. So the reinforcement learning algorithm is an exploration/exploitation types of algorithms. So this algorithm is going to randomly explore some of the pricing and bidding strategy and try to find the combination that will maximize the profit.

This is the result: In this figure, we are going to present the prices on the different scenario as a function of the consumer who only consider the top few products. This is a result from a duopoly setting. There is one sponsor listing, two organic listings. So there's two different seller, are going to bid for one sponsor slots. The consumers who only consider the top is on the X-axis, while the Y axis is the price.

So the black line is the competitive price is from the world without algorithms. So the sellers, we're going to assume that they have the complete knowledge about the market. They're going to set the price, and also determine their bid. Again, theoretical format without algorithms, while the blue line is the outcome from the algorithm world, is the outcome from the simulation.

As we can see, the previous research is the red dot. What the red dot tells us is that when consumer can see everything, these algorithms can learn to charge higher prices compared with the competitive prices, and that hurts consumers and raises antitrust concern. However, what we find is actually the blue region. If we compare the blue line which is the algorithmic price with the black line, which is the competitive price, we can actually see that the algorithm price end up being lower competitive with the competitive price. This is important because the takeaway here is that consumers can end up paying less, and this reinforcement learning algorithms can generate a beneficial effect for the algorithms.

One concern might be, it might depend on the algorithm details about the parameter that we would put in the algorithms. Actually in the paper, a very cool result is actually we show that. It actually does not depends on the exact details about the algorithms, it only depends on the market parameter. The reason is following.

Now we consider another scenario which is a full collusive case. By full collusive, the two different seller is going to sit in a room and they're going to decide their price and pay, in again theoretical format to maximize the adjunct profit. If that is the scenario, those two different sellers will decide to bid zero because this is a zero cost, and then they will decide on their price. The price in that scenario is the red line, is the full collusive price.

As we can see, if consumers turns out to only care about the top products, and then the monopoly would have the incentive to lower the price a little bit to increase the demand and increase the total profit. Then, these algorithms will always facilitate some kind of collusive outcome, but the collusive outcome in the right region is beneficial for consumers.

Then, if we look at the point where the full competitive price interacts with the blue line which is the algorithmic price, is actually the same point where the black line interacts with the red line, which is the full inclusive price. This means that the crossing point that we need actually depends later on the algorithm details. It only depends on the market parameter that we can empirically estimate.

Finally, if we look at the region where consumer consider everything, these algorithms increase the price. Whereas consumers only focus on top products, these algorithms lower the price. It means that these algorithms tend to impact different market differently, and this is actually interaction. In fact, later we are going to empirically test whether this interaction is negative or not in the data.

So as I mentioned before, we are going to test the interaction. In fact, the first part is the consumer behavior is the distribution of consumer consideration set. We're going to use the variation in terms of product position, the daily sales, to pin down the distribution of consumer consideration set. That is from the demand side. The key challenge here is that we actually don't know whether a seller is using algorithm or not. Then, we are going to infer from the price that is charged by different seller. What we do is that we are going to generate an algorithm usage index.

The interaction is here. So, we're going to look at the pricing pattern across different sellers. If the price among different products are highly correlated and the price changes are highly correlated at high frequency over a period that we are going to look at. Then we can infer that the sellers might be using algorithms to chart the price, because for human they might not be able to change the price at that high frequency based on the knowledge about the market average price. Then we separate the market where consumer consider more of the products, less of the products, and the market where the algorithm adoption rate is high versus low.

As we can see, the interaction in fact is indeed negative in the data. In the market where consumer consider more of the products, algorithms increase the price. Where in the market where consumer only focused on top products, algorithms lower the price. That means that the interaction in fact is indeed negative in the data. It tells us this is a really nice piece of evidence that the price in collusion and bid in collusion might be both happening in the real world. This is also what we predict using our theory model.

So to summarize what we have today, basically three main findings. The first finding is that these algorithms can learn to generate a beneficial effect for consumers. The cause and condition that we need is that consumer consideration set is relatively limited. So we actually show that the cause and condition, and the main result actually depends very little on the algorithm details because it is pinned down by a full theoretical result. As long as we can pin down the market parameter this result, we can already use it to draw the prediction. We try to provide some empirical evidence, and we indeed fund the enacting interaction, in fact between consumer consideration set and algorithm usage on prices.

Finally, we also consider the effect on the platform. The main result is that the platform can also benefit if the platform's profit is decreased, and then we show that adjusting the commission rate might be a useful tool to response while adjusting the reserve price in the auction will lead this algorithm to learn to collude on even lower base. That is not a very useful tool for the platforms. Thank you so much for today. Happy to answer any question.

Kanishka Misra:

Thank you very much for ... Sorry, no. It'd be funny if I start presenting this. Oh, no. All right. Thank you so much for having me present this paper and really fun to read. It's an area that I'm very passionate about, and I'm glad we're getting a lot more research in this area. Before I talk about that paper, let me quickly talk about what we've known this area. I think just a framing of how this paper would fit into the larger literature.

The literature we're looking at is a literature which talks about firms setting prices using algorithms. So if you think of AI and marketing, AI and pricing, there are lots of different uses of AI. This is AI or pricing for price discovery. So the supplier, the firm does not know the right price. They're using an algorithm to figure out what the right price is. The reason why these words are non-trivial is because the kind of algorithms we tend to use are written in a computer science word that is written for a stationary environment. The moment I'm using an algorithm, you're using an algorithm, we've got a non-stationary environment. In fact, it's endogenously non-stationary. So, these environments tend to be very hard to study.

What we have in the literature so far is a number of papers, including some of my own work, that shows that if multiple firms use AI type algorithms and are competing in AI type algorithms, we can get to market outcomes that are super competitive. So, these are defined by much like the presentation side, market outcomes where prices are higher than Nash. The question then is, what should we do with this? How should we interpret this? What's the general outcome of this? What is the so what? What's the implication for policy makers for this?

I would say one thing in this literature so far is we actually have limited theory. We have results for some algorithms can reach collusive outcomes. In fact in our paper, we have some words where firms can use bandit algorithms and reach Nash outcomes. So I would say so far what we have is understanding that collusion can happen or super competitive prices can-

Kanishka Misra:

.... understanding that collusion can happen or super competitive prices can happen. And I would classify this more as a... Showing this could be a concern rather than a broader mechanism for exactly what's causing the concern. And again, people are looking at mechanism. Hopefully, we get a better answer here. What I mean by mechanism here is these firms are learning these different outcomes.

Exactly why the firm is picking this outcome versus another outcome or why the algorithm is picking this algorithm versus the other is something we as a literature don't know yet.

All right. What does this paper do? This paper takes this literature and says the literature start talking about a world where there are two firms selling products against each other, but let's take a more realistic view of the world where these two firms are selling these products in a platform. The platform is showing the prices but also doing a second thing, which is running an auction to say which position should these different products be in. And buyers also are not fully aware, like the rest of the literature is saying. Buyers are more realistic, saying some buyers have search costs, some buyers don't. And there's a discussion there.

The main results here are very nice saying all these results that we've got in literature that Q-learning leads to higher prices is something that can be reversed. They said when you have auctions and you have... When you have search costs are high enough, you actually can flip those results and get lower prices from Q-learning. And I would say what the main, my takeaway from this is showing some of the previous results that we have in the literature are knife edge. The moment you add some extra reality to the results, results can potentially flip.

My big question for this, and this is a general question, when thinking about these kind of Q-learning mechanisms and applying them to more complex environments, one of the big downside of Q-learning is it's very slow. So learning is very, very slow in Q-learning versus any other forms of algorithms. If you extend this model to include advertising, you're now creating an extra dimension that the firm has to experiment in. And this is going to take a very, very, very long time to experiment. So in a simulation we can do it. My guess that's going to be 100s of millions of observations. If we try and apply this to a real setting, we have to believe that demand and supply is going to be stable for all these 100 millions of price experiments. And if we do believe that, then perhaps this is a little bit more believable. And I think that's something you can show.

I do have some other small suggestions for the authors. I think one of the suggestions is when you think about consumers, you've got a very mental cost view of consumers that consumers cannot think of more than product, more than two products, perhaps thinking of it more like a choice friction or search cost where it's endogenous of how many products I think of depending on what comes out in the ranking, that could be helpful. And the Amazon data, I think your main thing is searching for search costs. It's very hard to do that with Amazon data, but it's much easier if you can get click-through data. So if there is any option to get any form of click-through data, I think that would make a strong addition to your paper. But overall, I think this is an interesting addition to a growing literature. We need more papers which add a little bit more reality to this literature. Thank you very much.

Speaker 51:

So I guess I was wondering if you think about Amazon, there are a couple of ways they can monetize. One is by auctioning off the slots through advertising. The other one is seller fees. And so there is a lot of collusion. They would have the opportunity to move from one type of funding to another type of funding. So I'm wondering if you were worried about that in the long run if there is a lot of collusion that goes on as a way for it to avoid that collusion.

Speaker 52:

Yes. So here what we consider is the impact on the platforms. On one hand, these algorithms can learn to price colluding in the pricing part. On the other hand, they can also collude in the advertising part. So for the platform they have two revenue channel. One is the commission. One is the advertising. So our main result is actually... So if the seller is collude on lower bid, this is a loss from the ad revenue part, but at the same time the price is lower, so the total demand is increased. So for Amazon itself, actually their effect is unclear. So if the price is low enough compared with the world without algorithms, so the total demand is enlarged, the total commission is increased. So for the platforms, it also can benefit from it. And indeed the platform can also suffer. So as you discussed before, they can also adjust the commission rate, but if the platform is going to increase the commission rate, the consumer will suffer. That is a limit, boundary condition of the result.

Speaker 53:

Question about the limited consumer choice. So if the platform is... If the products that the algorithm is looking at are limited and the consumer choice is actually limited to a certain subset of products in the market, is one reason why there could be lower prices is because there's still other products in the market that are not being taken into account with that algorithm? So the algorithm is trading information with other... Is seeing other prices in just a subset of the market and they're still within that relevant market, the larger relevant market. There's still products where there's no price discovery. So you'd have lower prices maybe just as an endogenous effect of the fact that search is expensive, it's expensive to look at a bunch of products at once. And is that an effect that you're looking at or seeing there?

Speaker 52:

Indeed. So it's because the consumer consideration set is limited. It's because for the sellers being on the top is very important. So in a world without algorithms, the sellers will compete very strong in terms of the ad [inaudible 08:36:51] part. So for these algorithms, they will collude on lower bid. So because the bid lower, it is lower much the cost. So ultimately the sellers don't need to price that high for their products. So indeed it is because the consumer consideration set is limited, consumer search cost is limited. The competition to being on the top for the sellers is very important.

Speaker 53:

Thank you. Thank you.

Speaker 52:

Thank you.

Shrabastee Banerjee:

Hi. Okay. Sorry. Big green button. Yes. Perfect. All right. Hi, everyone. And thanks again to the organizers for putting together this great conference and also including our paper. This is joint work with Anita, who's in the audience and Giorgos in Boston.

So the motivation for this project came from the fact that travel platforms often need to decide on what kinds of estimates they would display when searches are conducted without entering specifics. And in the context of a travel platform, the specific is usually a date. So in general we see variations of this in practice. For example, if you look at this screenshot here from Google, you would see that if I just enter a location without entering any dates, then Google is going to impute some dates on its own and then it'll give you the prices for those dates which it thinks would be helpful for your search.

On the other hand, Airbnb takes a different approach where if you search for just a location without dates, it's going to give you a few different date ranges with different prices. So it can show you both versions to give you an estimate of prices for those two different ranges. So in general, something we were interested in is trying to figure out whether it matters how high or low these estimates are relative to the prices you would eventually have to pay. And in our particular setting the context is from prices or floor prices on an online travel marketplace.

So why do from prices matter? So in general, these sorts of prices are used quite widely in online settings. For example, even for search engine optimization, if you do a Google search for hotels in Boston, it would often give you these links with the from price information in there. And then similarly, many travel marketplaces also are often not price setters, but they have full control over how to display different kinds of price information. And here you would see an example where Hotel.com has displayed these two from prices for hotels which consumers search for without entering specific dates. So one interesting thing about from prices is that platforms do have a lot of control over how to select the window over which this minimum price is displayed. So in some cases it could be, say, a 30-day window or a six-month window, but in some cases if it doesn't have access to that granular information, it might even go back in time and try to estimate what the from price is.

So what we do in particular is we conduct two large field experiments in collaboration with the travel platform where we are able to raise the displayed floor price that consumers see by 10%. So the first thing to try and think about is what are the possible impacts if we do raise this from price. So firstly, it is possible that it has no impact. So especially if you're a user who knows your travel dates, you go on the platform, you entered them, you never see from prices or they don't even matter to you. So in that case, you wouldn't expect anything to happen to consumer decision making. However, it could end up having a positive effect as well if your from price or floor price ends up being closer to the price you eventually have to pay. So in other words, if that price that you see up front is more realistic and it helps you set more realistic price expectations, then maybe it has positive effects on engagement. However, it could also have a negative effect if users see high prices and eventually they drop out as a result of that.

And in general, Diego also talked about this a little bit in his presentation, essentially you have two different camps which would predict these two opposite effects. So the high floor prices good camp would be predicted by literature around sticker shock and anchoring, which would say that if this floor price displayed is higher, then there's lesser discrepancy between that price and the downstream price that you see, and then that would lead to positive outcomes and the final price would rise less, which would be perceived as more fair. However, on the other hand, the high floor prices bad camp would be predicted by the search and salience literature. And Sarah also has a paper in this space where they're looking at upfront versus backend fees and they also find that if you give people a high upfront price, they tend to drop out and they would engage less. So basically in this sort of a situation, even though a higher floor price would be more realistic, users would react adversely, which would eventually generate these negative outcomes.

And what we find is that we find support for the search and salience hypothesis. So essentially what we see is that when users are exposed to these higher floor prices, they reduce their engagement and we measure engagement by looking at outbound clicks on the listings, looking at the number of searches, how much time they spend on the website. So all of those tend to go down. We also look at the propensity to book, as well as the total revenue. These are directionally negative, but they are more noisy due to sparsity. And then finally we see that these effects are all amplified by trip urgency. So if it's a trip that's closer in time to your search date, then you are affected more by seeing these high floor prices upfront.

All right. So the experiment itself, basically it was conducted on Holidu.com over two periods of time. So the first one was March to April, 2019, and the second one was January to March of 2020. And this was pre-registered. And then as I mentioned, the treatment was to raise the displayed floor price across the whole website by 10%. And the randomization happened at the user level. So every single user, as long as they did not clear out their cookies, we could track them over the duration of the experiment. And this sort of a design avoids listing levels spillovers. We also considered randomizing at the listing level, but eventually if it's at the user level, we have more granularity. And here in this figure, this is just a manipulation check to see whether raising floor prices actually ended up doing what we thought it would. So in green, you have the floor price distribution for treatment, and in pink, you have the control distribution and you see that the treatment distribution is shifted out to the right, which means prices that people see in that condition are indeed higher.

And this is what the search interface on Holidu looks like. So if you go in and you enter a destination without entering any check-in or check-out dates and you hit Search, that's when it would trigger our experiment. And you would see something like these screens here. Note that there are some other differences, for example, in the button colors, et cetera, in the two screenshots as well, but those were all orthogonal to the main experiment we were running. So they don't make a difference to our results. But essentially, the screenshot on the left is the baseline and the screenshot on the right is giving us the 10% higher floor price condition. So eventually, based on this experiment, we have data on about 6.3 million users and searches are conducted quite widely across a large number of cities and countries and from a lot of different domains across the world. So in total we have about 32 million search instances, which are all aggregated at the user level, because as I mentioned, the randomization was at the user level. And the outcomes that we're interested in are the number of searches, the number of outbound clicks, the number of searches with dates entered, and also booking-related variables such as the total amount of bookings, the amount spent on those bookings, as well as the time on site.

So the first thing that we noticed, even without putting this into a regression just based on raw numbers, is that higher floor prices tend to lead to lower engagement across the funnel. So if you look at the first column that simply tells us that users were split almost equally into treatment and control. But in that situation you see that in the second column, we can see that basically after seeing those from prices, the number of searches conducted by treated users are lower and the number of dated searches conducted by those users are also lower, and notably, even after users know the dated prices, which do not change for our experiment, those remain constant, and even after users have been exposed to dates, they still end up conducting fewer searches in treatment. And even when we look at the session-level metrics, those results remain the same. So then we basically look at the same variables in a regression and we look at it for the outcome variables that I just mentioned. And what we find is that in the treated condition, which is the higher floor price condition, all of those metrics are negatively affected. As I mentioned, the booked and booking value outcome are directionally negative, they are not significant, but outbound dated searches, time on site, et cetera, are all negative. Then finally we delve into the effect of trip urgency and we try to see whether that might have some heterogeneous impact on our main effects. So if a user is searching closer to the travel date, they might be less price sensitive, so you might expect from prices to not affect them as much, but on the other hand, we know that if you're searching closer to your time of travel, you're under more time pressure and that might end up triggering these sorts of heuristics that you rely on. So floor prices could have a stronger effect as well.

So for this analysis, we look at the set of users who have conducted at least one search by entering a date. And doing so, what we find is that urgency actually amplifies the impact of floor prices. So if you search way in advance, then floor prices wouldn't affect your decisions as much, but the closer you get to travel date, the negative impacts are amplified even more.

So in summary, we investigate the effects of displayed floor prices using field experiments and real browsing behavior. In terms of implications for platforms, we try to look at these different kinds of interactions to show that even from prices which may even not exist very frequently on the calendar, they could still have these sorts of impact downstream on consumer behavior. So the kinds of decisions that platforms make in terms of selecting this price window or how wide to make this window could all have implications for behavior. And in terms of regulation, we show that consumer inattention and possible negative effects on engagement could happen even when platforms are trying to provide a more realistic price estimate. So there is an unintended sort of an effect that kicks in.

So the open questions here of course are then what could be an optimal level of transparency? If people have not entered their travel dates, what is the best data platform can do? How can it make sure that it's providing a realistic estimate while not affecting engagement? Those are all open questions that this project made us think about. And maybe one possible mitigation strategy which can deal with at least the consumer inattention aspect is to provide all prices on the calendar. So you assume consumers can just look at the calendar and then they have full information when they're trying to decide, which is similar to what Google Flights does for example. All right. Thank you for your attention.

Navdeep Sahni:

Okay. Should we have some back-ups now [inaudible 08:52:28]. Okay. Very happy to be here. Thanks to the organizers for sending this paper to me and inviting me to discuss it. My name is Navdeep Sahni. And I'm going to use my five minutes that we have for this discussion. And I just want to make two main points in my discussion. First of all, running and successfully conducting these kinds of large-scale studies is not easy. Having done a lot of them and having seen a lot of them in the literature, there are many ways in which these experiments get corrupted, get messed up. And there's a lot of evidence in this paper that the experiment was done correctly and there's lots of checks that Shrabastee didn't show that show that the randomization was implemented correctly, which is not trivial, especially if it's a nonstandard experiment that the platform runs.

So the next point I want to make is about once the experiment is done, how do we interpret? What do we take away? What is our big picture takeaway from the data? And that's what I want to spend some time talking about. So the overall result is clear. I'm not going to summarize it. When we increase the floor prices, people reduce their search, bookings go down. So when I was reading the paper, when I looked at the interpretations, the paper uses some plausible mechanisms to think about the data. One of them is that when people use the floor price, the from price, that has been talked about, they use that as a signal which that price becomes salient. Maybe people are using that as a reference. Maybe people are using that as an expectation. So these are possibilities.

When I was thinking about this, I was taking a different perspective on how people might be interpreting these floor prices, and here is a quick version of the model that I had in my mind. So let's imagine a consumer who goes to this travel website. They are at the homepage. They see this from price, so they get a signal. Okay. So the decision they got to make is, "Should I continue to search and explore this website more? Or am I done? Should I exit and choose a competitor?" So let's imagine, so that's the stage where the consumer is. So they have to make a simple trade-off. If they continue searching, there is a probability they have a belief based on all that they see on the homepage whether they're going to find an acceptable price because their travel time is presumably stable, so whether they're going to be able to find a price which is low enough for them. If they do find that price, then they're going to get a value, v, from searching on the website.

They're going to continue searching if this expected value, which is v times pi* over here, which is the probability of finding an acceptable price given the signal, which is the floor price that I see over there if it is greater than c or not. So that is basically the decision that the individual is making. And the pi* how to translate the signal into an expectation that is formed in equilibrium based on all the experience that people have, based on the conduct that is all across in this industry. Okay. So now given this model, the proportion of people that are going to continue to search is 1 - F(pi*), and F is just the heterogeneity in c divided by v, which is... Everybody is going to have their own value and based on the CDF of that, we're going to have a proportion. Okay. So that is basically the setup that I was thinking of.

So now given this setup, how do we think about the treatment effect? So the treatment effect, the main effects in the paper are just F(pi c*) which is the belief that I have about me finding the right price in the control group minus finding the right price in the treatment group. Okay. That is what it is. And in the treatment group the signal goes up by 10% and we're seeing a negative effect. So this tells me that people are updating their beliefs about the likelihood of finding the right price correctly. So should they be... If the floor price increases, the likelihood of finding the right price should go down and that is what we are seeing in the data. So the effect's in the right direction.

The next question we could ask is, are people correctly updating it quantitatively? To comment on that, one thing the paper could do is to actually look empirically at the distribution of from price and the actual prices people end up paying. And that can be something that can be studied. So that'll be one side of the equation. So just a simple correlation of two properties, one with 10% higher price, another one with lower. So are people actually paying higher prices? And by what magnitude? On the other side, we can check if beliefs are updating by 10% or by whatever magnitude. That'll be interesting to evaluate the from prices, and that'll depend on the shape of the F, which will have to be estimated, and that's not obvious. So that's one result.

The second result is about conversions, about actual bookings. Over here, this model would predict that if people's beliefs are reasonable, we should expect that the percentage change in bookings should be higher than percentage change in search. Why is that? Because when there's a 10% increase in the floor price, if the beliefs are reasonable, the people who exit searching are those who are more affected by the change in floor price. So we should see that the percentage change actually increases. So there is some suggestive evidence of that, although that's not very precise, but there was a smaller decrease in search and a larger decrease at least based on the point estimate on sale. So that was also suggestive of this mechanism. So overall, just to conclude, it's a very well done experiment. I commend the authors for all the effort in doing so. I think taking it a little forward, pushing on the mechanisms and maybe providing one model for the audience to take away will help it further. Thank you.

Shrabastee Banerjee:

And thank you so much, Navdeep. That was super helpful.

Speaker 52:

I'm just curious. The experiment is raising the floor price by 10%. Why the company choose to raise rather than lower it? And do you think the effect will be symmetric or asymmetric if you lower it?

Shrabastee Banerjee:

That's a good question. So initially our thought was also to lower it. That's what we had proposed them, but then they did not want to do it that way. They said, "Yeah. We don't want to explicitly make it even more deceptive. So if we raise it, then you can treat the baseline as the lower from price and the other one as the higher from price." But yeah, that is an interesting question. I think here we can only test this one level, so it's just a 10% increase, but it would be interesting to see if we make it very unrealistic, then will it still continue to hold? So we don't have the whole demand curve of course, but yeah, that's an interesting question.

Speaker 54:

Hello.

Shrabastee Banerjee:

Hey.

Speaker 54:

Hi. So given the way the from value is calculated, it seems like there's enough information there to calculate something like the average. Is there any advantage to using the floor versus, say, the average?

Shrabastee Banerjee:

Yeah. That's a good point. So I think there are some platforms that might even use the average and some, like Google or Airbnb, they actually impute some random dates to give you these estimates. And I think, yeah, somehow floor prices have been used more than averages. I think just because they would tend to be more attractive upfront. But of course it is very sensitive to the time window as I was saying. So if you select, say, a two-year calendar and you say, "I have a $30 night somewhere after one and a half years," that's not helpful for the consumer, but somehow we still see this practice happening. So yeah, that's a good point.

Speaker 51:

I was wondering what the optimal seller strategy would be here, because you would think if was a seller, you would want to have one week in the year where prices are really, really low and then you're sort of setting the reference price really low, you'll get a lot more demand. So I guess what keeps that from going to a price of zero? And do you see sellers that seem to be doing that?

Shrabastee Banerjee:

Right. Yeah. So we don't... That is a good point. We haven't looked at how unrealistic it gets. It might be interesting to see if there are indeed some sellers who really push this by just selecting on very specific dates where a low price might be available. But, yeah, on the whole, we were just trying to figure out whether these signals on average whether making them more realistic would have this effect. But at the same time, because the platform has so much leeway in selecting these windows, yeah, it is interesting to think about how much they can push this. So I think that was a little bit the motivation where we were seeing that we have these price signals which often are not very commonly present in the bookable dates. So in that case, if platforms are trying to make it more realistic, does it actually have this backlash effect on engagement? Yeah.

Speaker 55:

Yeah. Thank you. I'm wondering about if you could say a little bit more about effect sizes, because I'm thinking about an alternative story that's... It seemed, if I saw your results, that you see less search activity but the magnitude isn't very large and that you don't see a significant change in the bookings or booking revenue. So the people who are more price sensitive drop out, but they wouldn't have bought anyway. And so in some ways the firm can say, "I'm being more truthful," and it doesn't actually have a significant impact on the revenue.

Shrabastee Banerjee:

Yeah. That's a good point. So essentially, we do see these effects on outbound clicks and on searches, which are, yeah, around 1 to 2% in magnitude. And in terms of bookings, magnitudes are larger, but it is more noisy. So yeah, I think there is this tension between firms being able to say, "If bookings aren't really affected, I can continue to be more deceptive just because I'm bringing more people in." But I guess that's the tension we are trying to navigate where we say, "Okay. Is that really what you should-

Speaker 56:

We are trying to navigate where we say, "Okay. Is that really what you should be doing, or should we just be more transparent about the prices? Even if it doesn't have this immediate impact on revenue, do we still try to provide more realistic estimates up front?"

Thank you.

Chuan Yu:

Great. Thanks everyone for bearing with me for the last paper of the day, and I think let's give a round of applause to the organizers and FTC for putting together such a great conference. Today, I'm very happy to share with you my paper on The Welfare Effects of Sponsored Product Advertising. To start, let me show you some search results you may get when you look for something on Amazon, Walmart, Instacart and Uber Eats. Across all these pages, you can see these sponsored products, circled in red, at the top of the search results. Sellers are ranked in real-time ad auctions and pay the platforms for each click on the sponsor listing. Below these sponsored products, our organic results determined by each platform's ranking algorithm. For all these products, when consumer makes their purchase, the seller pays platform a commission fee, usually a percentage of the price.

In recent years, we've seen a huge expansion in this market. Here are the ad revenues for these platforms I showed you in 2023. Remarkably, Amazon's ad revenues surged from $2 billion in 2016 to $47 billion last year. For Instacart, although it is a grocery delivery platform, nearly 30% of its revenues came from selling ads. However, this trend has received mixed reactions. On one hand, some consumers seems to be frustrated about search results getting worse, and sellers on these platforms complain about higher selling costs. On the other hand, proponents argue that advertising can help sellers reach new consumers and increase the visibility for less-established products. Moreover, this trend substantially changes the ways that many platforms used to make money, which used to be primarily commissions. As we know, the FTC launched a lawsuit against Amazon last year, and one argument for its market power is the expanding advertising business. In this paper, I'm going to develop a framework to quantify the welfare effects of sponsored product advertising on consumers and sellers using empirical focus on Amazon. So introducing and advertising on these platforms may affect welfare through different channels. First, organic rankings in search results may not always reflect consumer preferences, especially for newer products. So, advertising could allow some high quality or low-cost products to reach more consumers and enhance efficiency. However, if organic rankings are pretty accurate, some low-quality sponsored results may displace high-quality alternatives. Meanwhile, sellers can pass on their advertising expenses to consumers through higher prices. Last, as I alluded before, the platform usually has access to multiple revenue streams. If advertising became unavailable, the platform may rely more on other instruments to raise revenues.

To quantify these different trade-offs, I've collected sponsored and organic results from over 1 million searches on Amazon. I also obtained information for all these products and information also on the ad auctions. I then use the data to estimate a model that incorporates interactions among consumers, sellers, and the platform. Using model estimates, I can simulate counterfactual scenarios with different sponsor positions and quantify the welfare applications.

Now, let me tell you a bit more about the data I use in my paper. I first identify several thousand keywords based of the highest search volumes on Amazon. I then scrape their search results six times a day over a two-months period in 2022. This gives me over 1 million searches in total. For each search, I can observe all the sponsored and organic results on the first page. There are usually 60 products on the first page, and 12 are sponsored ones, located at the top or in the middle of the search results. For all these products, I obtain their daily prices, sales quantities, consumer reviews, and listing time on Amazon. Compared to top organic products in the same search, top sponsored ones usually have a shorter listing time on Amazon and substantially fewer reviews. They also tend to have a higher prices.

I also obtain the commission rate set by Amazon on the platform. For most markets, in my analysis, this commission rate is 15 or 17%. Last, Amazon reports the median, lowest, and the highest bid among recent auction winners for each keyword as a guidance to the sellers. I'll use these statistics in my estimation on the auction model. To quantify the welfare implications, I need a model to predict consumer demand, seller profits, and welfare, and there are some counterfactual set of sponsor positions in search results. The agents in my model are consumers, sellers, and the platform. The primitives are consumers' price sensitivity and search frictions, sellers' organic ranks, quality, and cost, and the platform's objective function. A market refers to a single keyword.

The model has three stages. At the beginning, a platform selects the average commission rate to balance its short-term revenues and some longterm considerations. Then, sellers compete in prices and bids to maximize expected profits. Last, consumers purchase products and their search frictions. Let me briefly describe the model backward. On the demand side, consumers utility follows a simple logic specification. The search results on the platform contain N positions indexed by small n. Importantly, consumers in my model will consider only the top n products in search result, where n differs across consumers. This can affect the consumer search friction in practice, which make them only consider a subset of products among all available options. Therefore, if a seller is located lower in the search results, it'll be considered by fewer consumers. A seller's market share is a natural extension of the standard logic model, but now, it also depends on the ranks of the products in search results and the share of each type of consumers.

On a seller's side, a seller chooses the price and bid to maximize expected profits in a week. They account from uncertainties in both organic ranks and auction outcomes. Organic ranks on Amazon are determined by a realtime algorithm and can vary across searches. Assume the distribution of organic ranks is fixed for a keyword in a week, which may or may not align these actual consumer preferences. Auction outcomes on Amazon are also stochastic due to manufacturers, like seller's budget constraints or the platform experimentation. I introduced randomness in auction outcomes in my model. So, sellers bid for probabilities of winning each sponsor position. In equilibrium, sellers cannot increase expected profits by setting different price or submitting a different bid. On the platform side, the platform chooses the average commission rate to maximize a linear combination of commission revenues, ad revenues, consumer surplus, and seller profits. This parameter meal can capture Amazon incentive to invest in user acquisition or retention just to maximize longterm profits. I want to point out that...

Sorry. So I want to point out that this commission rate in my model can be a representation of a broader set of monetization instruments available to the platform. These instruments can act as a substitute to advertising revenues. For example, a higher commission rate will lower sellers' margins and reduce their willingness to pay for the sponsor positions. This will suppress their bids in auction and decrease the platform's advertising revenues. The same logic applies to other types of fees as well. This substitution across different instruments is key in my model.

Now, let me briefly discuss some estimation results. This figure shows the fraction of consumers who consider each position in search results across different markets. We see there's a gradual decay in the number of consumers that consider each product. For media consumer, they consider 28 products in search results, which should be interpreted as the upper limit of the page position that the consumer ever considers. Here, I'm comparing the estimated utility for products in each position. Products with higher organic ranks tend to have a higher average utility, and this is largely monotonic. On the other hand, sponsor results on average deliver low utility to consumers. This figure compares estimated marginal cost for sponsored and organic products. Sponsor results on average have lower marginal cost, however, the ad payment required for sponsor placement eliminates and reverses this cost advantage. Last, to determine the platform's weight on the consumer and sellers, I assume the observed choice of commission rate is optimal.

Finally, I can tell you some counterfactual results. In the first exercise, I replaced all sponsored positions in search results with organic listings while holding the commission rate constant. Here, the blue bars told the outcomes in the status quo based of sponsor positions, and the red bars told the counterfactual scenario with only organic listings. While holding the commission rate constant, removed advertising benefits consumers and sellers. This is consistent with the model estimates. Sponsor results, on average, deliver low utility to consumers. Although they have lower marginal cost, the ad payment reverses this cost advantage. When the platform can adjust the commission rate in response, its objective is maximize at higher commission rate than the current one. Having sponsor positions will incentivize the platform to maintain a low commission rate, which can bring more advertising revenues to the platform. When this force is removed, my model predicts a substantial increase in the commission rate.

In this figure, the blue and red bars replicate the previous figures, and the new, green bars told the counterfactual scenario based on the optimally-adjusted commission rate. In this case, compared to the status quo, both consumer surplus and the selling profits decrease. This reversal in the welfare comparison highlights the importance to account for the platform's response. If advertising were removed, Amazon would have a strong incentive to raise more revenues through other channels. It does not have to be the commission fee, but can be any instrument, the X and the substitute to advertising revenues. In my model, compared to reason of the commission rate, introducing advertising, not only raises more revenues for the platform but has a smaller negative impact on both consumers and sellers. This is because the auction mechanism allows the platform to price-discriminate against different types of sellers. You can think of two types of sellers, high-quality, high-cost organic products, and low-quality, low-cost sponsored ones.

This advertising, the platform can lower the commission rate on the first type, but charge a higher effective commission rate on the second type. This brings the platform more revenues and could also enhance efficiency or create less distortion that can be shared by other market participants. I also consider alternative numbers of sponsor positions in search results. Here, I vary the number of sponsor positions from 0 to 20. As we see here, both the consumer-optimal and a seller-optimal number of sponsor positions are lower than the [inaudible 09:17:55] number and lower than the platform-optimal one. Therefore, a limit on total sponsor positions could benefit consumers and sellers.

Of course, in practice, policymakers don't want to tell Amazon how many sponsor positions they can have on the platform. To the last exercise, I consider a simulation or I vary the weight that platform places on consumers and sellers. This can be seen at the exercise to vary the level of platform computation. An increase in computation could force the platform to share more surplus with consumers and sellers. As the MNO place more weights on consumers and sellers, it will choose fewer sponsor positions and also lower the commission rate, which can benefit consumers and sellers.

Now, let me wrap up. My paper highlights the importance of counting from a platform response when policymakers try to regulate platforms with access to multiple revenue streams. Advertising in Amazon could benefit consumers and sellers by incentivizing a lower commission rate. A limit on total sponsor positions or more competition among platforms could benefit consumers and sellers.

Thank you. That's all I want to share with you today.

Leon Musolff:

All right. Big, green button in the middle, I've learned. Excellent. Okay, thank you so much to the organizers having me discuss this paper. It's always nice when you're asked to discuss a paper that you've actually already really carefully read because you really like this topic. So, that was nice. I know it's been a long day, so I thought I'd start us out with a joke. A policeman sees a drunk man searching for something under a streetlight and asks what the drunk has lost. He says he lost his keys, and they both look under the streetlight together. After a few minutes, the policeman asks if he is sure he lost them here under the streetlight, and the drunk replies, "No. Why do you ask? I actually lost them in the park." It's just that this is where the lights are. Right?

I think that describes a lot of digital economy IO. And not to say too much about the other papers today, I thought there was some excellent work here, but I do think we saw a little bit on display that there are some areas that are much easier to investigate because firms are willing to run experiments with you. These areas, somewhat unsurprisingly, don't overlap a whole lot with worries of antitrust. Oh, I'm sorry. All right, I have to work on my glasses here. These areas don't overlap so much with antitrust. So this paper had shown you, to its immense credit, refuses to search where the streetlight is. And instead, it goes where I think we actually lost the keys. And this is advertising.

If you know anything about the digital economy in these past few years, you know that Amazon has been developing as an advertising behemoth, and this came up recently for me as I'm doing some work on the dominance of Google. It struck me when I was reading that DOJ case against Google, that actually, Amazon is now an important market share level competitor of Google depending on how you define the market. But I totally didn't even have that in my brain, and that has been a development that has happened very rapidly over these recent few years.

So I think it's a really great thing that we have a paper that searches where we lost the keys. Now I do have to say, searching in the dark is hard, and that happens to be where we lost the keys. Chuan didn't have a lot of chance to explain the details of the identification to you, so I thought I'd just point out one thing, which is that what's absolutely crucial in the paper is for us to know the impact of an organic rank on your sales. Because to the extent that the organic rank doesn't matter to your sales, so there is no incentive to advertise, and this whole problem won't emerge. So, we want to know this limited extent of consideration. However, your organic rank is mostly an average of your past sales, and you can see how we're going to run into endogeneity challenges here.

And to be clear, Chuan is very aware of this and he addresses this in the paper using a variety of techniques, and I definitely think that they get us most of the way there. But I do worry a bit that sales are going to be very autocorrelated on Amazon, and in particular, the core identification strategy. And the paper relies on this idea that sales are autocorrelated but ranks are going to be even more autocorrelated than sales. And that sounds plausible. But of course, if I understand correctly, you don't actually have sales. Instead, what you have is you have these other ranks which are also a smooth version of past sales, and then it becomes a lot harder to understand whether that strategy will work. Now having said that, I only have this figure in black and white, I didn't get it in color. But I do think it's one of the key figures that tells us that around 75% of consumers don't make it to the second page according to these estimates. I did some quick Googling and that seems pretty consistent with numbers that are reported in the industry.

And maybe I would just emphasize, maybe you can get some other numbers from industry and just get some sense from ClickStream data, how consistent your estimates are with those, because that would really help us build confidence for your results. So really quickly, I think one core part in the paper is the accuracy of these organic rankings. And it took me a little bit to understand how you can claim to have better accuracy of consumer welfare than Amazon, because Amazon's organic rankings, they are estimates of how good each product is. Right? But of course, what you have is when we advertise this product, you know how much do sales increase. And my understanding, and you should correct me if I'm wrong, is that that gives you this additional boost. That if we just had organic rankings, we wouldn't have this additional variation.

So, we couldn't identify that. And on that note, just a final comment here, is I kept thinking, "Sure, Amazon may be able to extract some additional information here using these auctions. But could they also do something else?" And the thing is, if advertising is so good, why don't we let consumers choose whether they want to see the ads or not? And then we could just have a toggle, and consumers that think, "Oh, the ads are bad for me," can toggle them off. And then this whole empirical trade-off that you expertly estimate but it doesn't even emerge, and that makes it maybe easier for regulation.

And with that, I'm going to conclude this paper is really great. You should all read it. But it also highlights that digital platforms have us at a data disadvantage. And hint, hint, we are here with a bunch of regulators. The European Union is starting to consider what can be done about this data disadvantage that we have relative to platforms as academic researchers, and I think that should be a topic that we think more about in the future.

Thank you.

Julia:

Thank you. So I've been waiting for this paper for a good amount of years now, thinking about how do we compete for shelf space in the digital world? As we said, nobody goes to page number two. And obviously, in the digital world, it happens in a very different way than what happens in the brick-and-mortar. Right? There, you get into your slotting, vendor fees, whatever else it is.

So, I was curious. And there again, those fees were pretty big part of the profits, which is the same thing we're seeing in the digital world. So thinking more from a theoretical perspective, whether this new way of, let's call it selling space because that's what we're doing, is very different than the old way of getting the lump sum fee and getting a product in.

Chuan Yu:

Yeah, that's a great observation. I think actually in my paper, I discussed that I think from theoretical perspective, it's pretty close to the previous guest debates or literature, these slotting fees in the physical retail industry. I think there's argument for both anti-competitive effects and pro-competitive effects. I think here, it's more like I could, by scraping the search results, I get to see the exact product placement. And also from some data, I can see some payments for these products, which could be better data compared to the old physical, I guess, retailers.

But I agree with you. I think at its core, it's not that different from a theoretical perspective.

Speaker 57:

So this probably doesn't matter for your theoretical results, qualitatively at least, but I wonder whether you've considered that if a consumer considers the first end products, there's a chance that they might disregard some of the advertised products.

Chuan Yu:

Yeah, great. I think that's a great point. I think my baseline model assumes that consumers will consider this top-end products. My paper considers extension where I allow a fraction of consumers to just escape all sponsor results and never consider them. I think my results will remain quality robust as long as that fraction doesn't exceed 50%. So if you have 50% of people escaping sponsor results, that could change my result qualitatively. But as long as it's not that big, I think my results remain robust to that.

Quantitatively, yes. So I think when you do that, you have two channels that can play. First is that the estimates of the quality difference between sponsor and organic will shrink. So right now, I'm using estimated quality based on review preferences. If actually a large part of the consumer, they never considers them, it would imply that the difference is actually not that large. So, that's one effect. Another effect is I think as you mentioned, it would just make more consumers search more and they will incur welfare loss when you only consider some top-end products, where top-end content is mainly advertising which you don't consider. I think these two forces offset each other in the counterfactual.

So that's why when the fraction is relatively small, it doesn't affect the result. But if you have half of the consumers, that will change the results.

K. Sudhir:

Yes, we need to get to dinner, so let me be brief. Click. Yes, let me make sure that my name is there. Okay, good. So when Julia asked that last question about how does this relate to slotting allowances, as somebody who worked on it, I said, "Well, wouldn't it be wonderful if I could pay slotting allowances for each individual exactly as they walk into the store?" That's the difference, I guess. Compared to slotting allowances in the retail store, you get to customize your slotting allowance exactly for a position at any given point of time. But that said, what a wonderful day. We've had such great talks, great discussions. And we could not have imagined as Jason and I started thinking about, "Let's revive this conference after eight years. We missed the COVID years." We couldn't have imagined such a wonderful conference and a great audience here.

So, let's start giving a round of applause to everyone who spoke today and to the audience. I was particularly delighted to see the session that Devesh organized with the folks of the FTC. I think it was a great learning experience for a lot of the academics as well to understand what goes on here at the FTC, and I think for the folks in the academic audience to learn in these interactions. But I think, I hope it creates more interactions across all of us so that when we come here back in 2026, in March, hopefully there'll be fruits of those interactions that we've had showing up as well.

So again, I look forward to even more successful conference. This has, I think, had much bigger audience than the one that we held first time. And so hopefully the third time, it'll be even bigger and it just grows the community. So that said, let me thank... We had 134 submissions this year, so it's going to be hard to beat that but I imagine we will. And I want to thank everyone who supported us here at the FTC, starting from Aviv, Devesh, Nelly, and especially Catherine, who has been amazing at keeping us all... Thank you. I'm going to leave the floor again for Catherine to come over and tell us a little bit more housekeeping tasks, but look forward to seeing you all at dinner at Zaytinya and continuing the conversation there.

Thanks.

Catherine:

Hello, everyone. Okay, we are running super behind, but housekeeping, if you brought in any papers or there are any papers or water bottles near you or trash or pieces of anything, please pick them up with you to take them on the way out. Thank you. That would help the custodians and everyone a lot to help clean things. If you have any comments about this conference itself, like how it's run, the food, the presentations, the format, anything, please send us an email at marketingconf@FTC.gov. Like we said, this is the second time we've held this conference, and the last time we held it was like eight years ago. So, all and any comments are appreciated. I will read every single one of them. Yes, conference organizers is me, so I will read all of them and all comments are greatly appreciated. And yeah, I hope to see you all again in 18 months for our third marketing and public policy conference.

Again, dinner. If you are going to dinner and you're not sure if you signed up, please check with me. I will be sitting by the registration table and I can confirm. If you still want to sign up, you can. There still are some spots available. We greatly appreciate. It's a great way to hang out. Zaytinya's really great. It's great food. And finally, we want to just let you know, badges and name tags, we're reusing name tags so you can keep it if you want. But if not, there's a bin for you to drop that off into. Same thing for the FTC visitor name tags, there's another bin for you to drop that off into by the registration desk. Please do that. Do not walk out with your FTC visitor name tag. That would be really, really bad.

Yeah, so please drop it off. That's it. Thank you for coming.