16th Annual Microeconomics Conference Day 1

7:30-- 9

Hey everyone, welcome to the 16th FTC micro conference, we are excited you could make it. We're going to get things started with a speech from our bureau director welcoming everyone. Our director is Aviv Nevo, he’s the director of the FTC on leave from the University of Pennsylvania where he is the George A Weiss and Lydia Bravo Weiss PIK professor with appointments in the Wharton school and the economics department and it's been a pleasure to have him here.

[Nevo] Thank you. And good morning, everyone. My name is Aviv Nevo, and I’m the director of the Bureau of Economics at the FTC. I'd like to welcome you to the 16th annual microeconomics conference hosted by the FTC. Now for me personally, it's a true pleasure to be here. I was actually on a panel for the first one in 2008 that Chris Adams over there organized, and I was in a scientific committee. I think it was called, the Scientific Committee on the second, third, and 4th ones. It's really a true pleasure to see the conference go on and continue and to be successful as it's been. One of the reasons this conference has been successful is the fact that it occupies a very unique place in the space of conferences. The goal of this conference is to combine cutting edge academic research and discussions of real world policy problems.

Looking ahead, the program, I'm really hoping this will continue and be the theme that we go on. For those from outside the FTC, I want to say a few words about the Agency and the Bureau of Economics. As you probably know, the FTC is an independent agency that has two primary enforcing missions -- consumer protection and competition. The Bureau of Economics support these two missions. We have, in total, 120 FTEs with almost 90 of them being PHD economists. That makes us one of the larger -or maybe the largest- group of micro economic economists in the federal government and we do a lot for the agency.

We support directly both the competition and consumer protection enforcement missions. We provide economic analysis and support of investigations and litigation, and we apply in many cases cutting edge economic analysis, both theoretical and empirical to these cases. As you probably imagine from those who have followed the news or are here, this is a particularly exciting time at the FTC. I could probably spend half the day talking about the stuff going on, and I probably can't say anything about any of it, so I won't. So it won’t be a very short half day.

There's a lot of interesting stuff going on. And BE is right in the middle of the action. You might have heard rumors that economics no longer matters at the FTC. So let me just take that head on. That's BS. That’s not true. I can tell you and I can assure that's not the case. The commission looks to vigorously enforce the antitrust laws and economics plays an even greater role than at other times. For anyone who wants to take economics seriously, should understand that economics needs to lead the way. If you want to do good antitrust policy, it has to be based on good and solid economics. Similarly, on the consumer protection side, as the FTC has used more of its rule-making making authority, BE as an important role in conducting regulatory impact analysis, so we're playing a larger role than historically on both sides of the mission.

What's not well published is that to meet all these challenges. We have been growing quite significantly. You might have heard that I said that currently we have 120 FDEs. If you go back and look at what was said in previous conferences, people talk about how we're just a little over 100 or we have 80 PHD economists. They didn't get the numbers wrong and I didn't get them wrong. Both numbers are correct. What's happening is that we've actually been growing. We've hired 10 new PHDs, we hired 10 last year and 10 the year before that. In terms of PHD economists, we're the largest we've been in decades if ever, maybe ever. The BE used to be much larger, but we served a different function. Mostly data collection back in the 60s and 70s, but in terms of PHD economists, we're probably the largest we've been in a while.

And the good news is, we're not done. We're looking to hire again this year, so if you're on the market, or your colleagues are on the market, or you have students on the market or friends, relatives, anyone who is a good economist, please let them know that we had a vacancy announcement. You can go to USjobs.gov and look at that. I'm sure folks around here will be happy to tell you more about the job. So you can pass it on to your students. Let me just say my personal thing, and what I've always told my students – it's a great job. The pay is crappy, let's be honest. I wish we could pay more, but the job is exciting and we truly have the best clients that you could have, because at the end the of day we are working for the American economy and the American consumers. At the end of the day, it makes you feel good about what you do. Seriously, if you have students who are even remotely interested in real world policy, please send them over to us.

We've also had some changes in management. Just recently Ted Rosenbaum was appointed the deputy for Research and Management. He joins Devesh, who's a deputy for consumer protection, and Alison who's a deputy for competition and top of the bureau leadership team. Ted has organized this conference several times and has been heavily involved over the years and is somewhere here, there he is over there. His appointment I think truly reflects the agencies and the bureau's commitment to research and a strong connection to the academic community. So the idea is that this is the 16^{th} and I can already say we are looking forward to the 32^{nd}, 64^{th}, but the idea is that this is going to go on.

Let me move on to a few things. This conference would not be possible without a long list of people working behind the scenes. I'd like to thank Steve Berry and the Yale Tobin Center for cosponsoring the event. Thanks for the conference organizer Viola, Will and Stephanie, of the Bureau of Economics and the Scientific Committee, Mike Sinkinson from Yale, Heidi Williams of Dartmouth and Steve Tadelis of Berkeley. Special thanks to the admin team. We have a small, but very good and efficient admin team. That's Maria, Kevin, Constance, Pricilla and Tammy. To the research analysts and statisticians that are helping with the registration --you might have seen them on the way in-- that is Chris and Chris, Marylyn, Ken, Jen, Scott and Aiden. Last but not least the FTC media team and the event planning staff and the BE economists who help screen all the great submissions. That's Eddie, Tom and Matt.

This is the federal government, so of course, I have to review a few important administrative details. Please pay attention because there's a quiz at the end. Per government regulations, unless everyone in the room passes the quiz, we cannot hold the conference. If you leave Constitution Center building for any reason during the workshop, you will have to go back through security screening again. You saw the long lines there. Please bear this in mind and plan ahead, especially if you're participating on a panel, so that we can do our best to remain on schedule.

If you need food, there's some food in the back, but there's also a cafeteria around the corner. There's a restroom in the building. This all to say, 'why should you ever leave the building in the middle of the time because it's really hard to get back.' I'm not sure if this is true, but I've been asked to read, that most of you have received a plastic FTC event security badge. If you did, please return those. If you don't, we will hunt you down. Subject to FTC act section 19b-814, we can prosecute you or something, but anyway, please bring it back. I'm hoping I can keep you guys awake. It sounds like I'm doing a good job. If an emergency occurs or causes you to leave the conference center but remain in the building, follow the instructions provided over the PA system. If an emergency occurs that requires evacuation, an alarm will sound and everyone should leave the main entrance you came in, 7th Ave, turn left and cross East St to the FTC emergency assembly lot. Basically follow where everyone else is going.

Please be advised this event may be photographed, webcast or recorded. By participating in the event, you are agreeing that your image and anything you say or submit may be posted indefinitely at the FTC.gov or on one of the commission's publicly available media sites. Please do not bring food or drinks in the auditorium, only water is allowed. Now with that out of the way, I would like to turn it over to Eddie, who will keep kick us off with the first session shared by Steve Tadelis. Thank you very much.

[Eddie] We're really happy to have Diego here, from the Federal Reserve of Chicago. He's going to be presenting on data, privacy laws, firm production and evidence from GDPR. I should remind everybody we're going to do 30 minutes for the presenter and 10 for the discussion and five for questions after.

[Diego Hernández] Thank you so much. Let me just start by thanking the organizes for getting the paper into the conference. I'm really excited to be presenting this project. I'm Diego and I'm a Chicago federal employee. This is joined work with Mert Derimer from MITs loan, Dean Li who's a PHD student at MIT and Sida Peng who is an economist at Microsoft research.

The main motivation for this paper is that data plays an increasingly important role in the production of goods and services in the economy and because of this increasingly important role, many countries have enacted new privacy regulations to try to set some ground rules about how firms should go about collecting, storing and analyzing data.

The most important one of those is the European Data Protection Regulation which affected more than 20 million firms around the world, they target EU residents regardless of their size. I'll describe it in detail in the next few slides. But what I want to say is because of GDPR, firms need to take costly measures in order to comply with it. They need to expeditiously process customer delivery requests, they need to have very detailed record-keeping activities of what they're doing with data and many others, that entail both fixed and variable costs. All I want to say is that these costs oftentimes –some survey evidence suggests that they're on the millions of dollars even for small firms and on the tens of hundreds of millions of dollars for large firms.

What are we going to be doing? We're going to be taking a production approach to study GDPR. And why is that the case? Essentially, think of any regulation that tries to set up ground rules for a given input. When you think of that, that's going to increase the cost of value input. In this specific case, it's the cost of data and it's going to affect other input choices. Why? Because it generates a wedge between the marginal product of data and its marginal cost. And it is going to be affecting not only that input itself, but also potentially other ones. And while there's an important literature that tries to think about how GDPR has affected firm outcomes come, we know comparatively less about what happens with firm choices. Partly that's coming because we need a framework to analyze these types of firm choices. And that's what we are going to be doing and we're going to be studying two specific questions.

The first one is how do firms combine data and computation in production and after doing that, we're going to try to think about the cost of GDPR for firms and how they adjust this data and computation inputs in response to GDPR. To do so, we'll be using what I would say is very cool data, the confidential data from one of the largest cloud providers in the world, where we'll be observing monthly information about how much firms store in the cloud and how much computation they are doing in the cloud. I'll discuss computation in the next slides, but it's hundreds of thousands of firms coming from many different industries from retail to finance to software to manufacturing.

With this data. The first thing we'll be doing is an event study specification, where we compare firms in the European Union that are subject to GDPR to firms in the US that are not targeted by GDPR. And, of course, we're going to be studying data because it is directly targeted by the regulation. We're still be thinking about other markets like competition, because it might be affected because of substitution. After doing so and guided by the results that we find, we'll be building a production function where firms are going to combine data and computation in order to produce information through a constant elasticity of substitution production function, where GDPR is going to be a wedge between the marginal cost that firms have to pay the cloud provider and the marginal product of data.

What do we find? First of all, we find that through the event specification that GDPR changes significantly the data compute input mix where firms become less data intensive. Three years after GDPR computation goes down by 15 and the amount of stored data goes down by 26% in firms in the European Union when compared to similar firms in the US. Through the production approach we will find that data and computation are strong complements. What I mean by that is stronger complements than traditional inputs. In the last part of the paper, using the production function model, we'll see that firm behavior is going to be consistent with them facing a 20% tax on the cost of storing data on the cloud and there's going to be substantial difference on these wedges. They are twice as large for small firms than for larger ones. When we think about the cost of producing the same amount of information, now that we take into account these wedges, it's going to be around 4% more costly. You might be thinking 'hey, what about the other benefits about privacy laws? I myself think the same. This is not going to be a full welfare analysis. We will not able to talk about consumer surplus. We're only going to be thinking about compliance costs. In the interest of time, let me say what's slightly different on this paper. On the impact of GDPR on firms, there's important literature and we'll be studying one of the key margins that's regulated by these privacy laws -which is data- and we'll be trying to think about the firm choices rather than firm outcomes using this production approach. There's also an interesting literature where data is an input to production of goods and services. This literature has been mostly theoretical and what we're going to be putting forward is an empirical analysis about how firms use data and computation in production and providing estimates for the literature. There is also literature on the economics of privacy where we'll be studying the largest or one of the largest privacy regulations to date, and literature on misallocation where we'll be thinking of the GDPR as a wedge between the marginal cost of data and its marginal product.

Let me start by giving you information about the GDPR. This is a law passed in April 2016 and went into force may 2018 and it's actually quite comprehensive because it applies both to firms located in the European Union, but also firms located outside of it that target EU residents. The definition of personal data is very broad, which makes it quite comprehensive because it applies both to employee and customer data. So regardless of whether you're only dealing with businesses or with customers as well, you are going to have to comply with this law. Enforcement is at the country level -or the EU state- where upon request firms need to demonstrate that they're actually compliant. While many services have shown that not every firm was compliant at the time of implementation of May 2018 -only around 10%. But not only that, there's differences in the cost of compliance depending on exactly how your firm uses data.

What's interesting is that GDPR is a data protection regulation that sets a number of company obligations in addition to individual rights, so it's quite costly for firms to follow these costly regulations. We list a few of them over here. They need to provide very detailed privacy disclosures, assign data protection officers, handle expeditiously customer delete requests or transfer requests -so I think they have a month for that- and they need to keep detailed records of processing activities. If they want to add an algorithm, they need to do a data protection impact assessment. There's penalties in case of breaches. If there's a breach, they need to notify the authorities, but also the customers themselves -they have 48 hours. That's a big number of different things that firms now need to do in order to comply with GDPR. We were trying to think about whether to think of these as fixed costs or variable costs. There's no published statistics about whether these are one or the other. We did our best and in the paper we explained them.

Let me just highlight a couple of them. For instance, data protection officers, you might think of them as mostly a fixed cost. But many firms according to survey evidence hire more than one data protection officer and whether you hire multiple ones or not depends on whether you have a lot or very little data. Also, when thinking about customer delete requests, you might be thinking about the ones in large platforms. Most of the firms actually don't do that. 64% do these delete requests fully manually. Only 1% fully automate this type of delete requests. When thinking about liability insurance, many firms buy it to protect revenue from these large penalties. Some buy them specifically because they want to protect themselves against breaches and the breach insurance premium depends on the amount of data you have, so these are mostly variable costs.

Another way to see the same or something similar is that this is publicly available data on fines that have been enacted against firms in the European Union. On the right of the screen, you can see names that you might recognize and those are the ones that are appearing in newspapers, but if you see the mass distribution of fines, you can see most of the action is on the hundreds of euros, thousands of euros and they come from very different industries, like jewelry manufacturers, restaurant owners, police officers, salon owners. If you try to search for them, there's quite a few that pop up that show that it's not just the firms on the right that have to worry about this type of loss.

Let me switch gears slightly and talk to you about what we have on the side of cloud computing. Let me first define it. Cloud provides on demand access to scalable IT resources through the internet. Instead of having in the previous years where you needed to buy computers in order to run some services, you can just rent them from cloud providers, like physical hard drives Instead of having to use your operating system within your computer, you can use whatever they provide in virtual machines. It's a widely adopted technology, around 90% of the firms in the world use cloud computing in one way or the other.

Just to set some examples, these are two firms -I want to be super clear, this doesn't mean that these firms are on our data. If you search for firms that have adopted the cloud, some case examples that come up are Netflix and Carrefour -Walmart in Europe. Essentially, for Netflix, the videos themselves are going to be part of the data storage, but also the user information. Whenever you click on the screen, there's an algorithm that runs in the back end and they constantly monitor your bandwidth to make sure that the streaming quality is going well. There's analytics that they do and that's kind of the computation that goes around. In terms of Carrefour, in terms of data storage, such as inventory, payroll, sales, etc. In terms of computing, they have the point of sale systems and they do some supply chain optimization. When we we're thinking about information or producing goods and services with processes that combine data, storage and computation, these are the types of processes that we have on the back of our minds.

What we observe is data cloud usage from these providers that I was describing, where for storage we observed the amount of gigabytes that firms are storing at any point in time. For computation, what we observe is a measurement that's called core hours. Most computers have 4 cores. If you've rented a computer for 4 hours, that would be 4X4, which is 16 computation units. We observe him at the firm, service, server location, month. So, MIT, we know how much they requested for computing services and in the fictional east coast server in May 2018 and we observed both the lease price and also the paid price, that might include some quantity discounts. It's cool data, but it's not perfect. First of all, we only observe how much data they have, but we cannot distinguish between different types of data, valuable or non-valuable, personal data, etc.

Furthermore, there is more than one cloud provider. So many firms end up using multiple ones. If that's the case, we won't be able to observe what they do on other cloud providers. Firms have computers themselves and if they do many other things within their own systems, we won't be able to observe that either. We tried to limit that by using a second data set, which is called Aberdeen/Harte-Hanks, which is an establishment level technology adoption survey where they have some numbers at firms which they call establishments, they have a dummy variable, which tells you whether a given company is using a specific cloud provider or not. We know some information about the extensive margin adoption of other cloud providers and whether some of the firms that we have in our data are multi homing or not. It's been widely used in the literature -the older versions- to measure IT adoption. We also complemented our data with Duns & Bradstreet and Orbis's data to try and know a little bit about the firms themselves. We have in our data for the European Union firms information about the number of employees that they have. And we also have information about the industry specification for the firms we observed.

Some summary statistics for the top 8 industries. As you can see, they come from very different ones, services, software and manufacturing being the three largest. About half of the firms are going to be coming from the European Union, with some variation across industries. In terms of computation and storage, one interesting thing is that if you look at the proportions, they are roughly comparable to the number of firms we have. If you divide storage by computation, the average compute to storage ratio will be about the same within an industry, but there's going to be tremendous heterogeneity within the industry that I won't discuss, in the interest of time, but it's interesting.

Let me start by showing you what happens when GDPR came into force. To do that, we're going to be using a standard difference-in-difference approach, but of course, there's a challenge in the literature, which is the fact that you're studying privacy regulation, it affects firms outside the European Union, there's a lack of natural control group. What we did to circumvent this is to use the fact that we observe data centers, like the data centers where firms are using or getting services from. And we are going to be defining the treated firms as firms that are in the European Union that only store data within the European Union data centers and firms in the US that only store data within the United States. Naturally this will be eliminating firms that are in the European Union and are using both the US and EU data centers. Those are multinational firms. We think they are important, we just don't think this identification strategy would work for them, so maybe it's for another paper. And we'll be thinking about the sample of EU and US firms that will be continuously using the cloud for one year, two years before GDPR came into force. We're going to be using a standard diff-in-diff -like dynamic diff-in-diff- and allowing the time trends to vary by industry and pre-GDPR usage as well. So we have 100 different time trends on our diff-in-diff.

The first thing we noticed we run this is after the implication of GDPR, it increases the price of data storage and storage goes up, it's not surprising, but we do see a large magnitude. When we go about computation, we also find that computation goes down as well, but not as much as data. Which essentially means that firms become less data intensive after GDPR and also because the price of data storage goes up and computation goes down, we think this is also suggestive evidence that these are compliments. One important thing is this doesn't necessarily mean firms are deleting data, because we could have a trend that's growing on both, but after GDPR the trend starts going at a slower pace in Europe. Also there's limited evidence of anticipation as I was saying, this is consistent with survey evidence and also, there's a gradual decrease in usage after GDPR which we believe is coming from different points of time in terms of compliance. One last point that I wanted to make is maybe what firms are doing is eliminating data that wasn't useful. If that was the case, we wouldn't expect any effect on firm outcomes, but we think that's not the case because computation also goes down. And if it was the case that they were only deleting data, then we wouldn't expect another change, a change in computation.

So therefore, we think it's not it cannot explain everything.
I'm going to skip the industry part in interest of time, but I want to say that essentially, we find the same patterns across industries. Data goes down, computation goes down, data goes down by more than computation, so firms become less data intensive. There's many more analysis that we do in the paper. We show that these differences are not driven by changes in prices both in the European Union or the United States, that we think it's unlikely that substitution to other providers or in house ITs is what's explaining our results by running these samples using the fact that we can observe whether firms are using other club providers with this other data set.

One thing I wanted to say is we want to publicly thank Garret Jones and Sam Goldberg and other authors for giving us data. They constructed a measure at the county level about how strict the regulators are. And we try to run unique analysis to try understand whether firms that are in more strict countries within the European Union, as in whether we find larger effects. It is the case that our point estimates are larger for the firms that are in more strict countries, but we don't have any statistical significance. Take it with a grain of salt. That's what we have. Some takeaways that I want you to remember, essentially GDPR changed the firm and data and computational input mix by making firms less data intensive. This result suggests there's an existence of a wedge between the marginal product of storing data on the cloud its price. How big it is and whether it's important for producing information, we don't know. So that's why we need more structure. And that's why we set up a production framework to quantify the GDPR cost.

Let me quickly describe that, we are going to be thinking about firms that produce information, combining computation and data. There're two important parameters here. The 1st is this omega, which is unobserved, compute technology. Which essentially means that for the same amount of computed data, some firms produce more information, and some others produce less. This is not going to be observed to us and there's an elasticity of substitution parameter that determines how substitutable computational and data are. And it is going to be industry specific.

One interesting thing is after doing this, we can essentially say there's a production function that has capital labor and information. We don't need to take a stand about what this production function actually looks like. And in fact, we can accommodate many of the used cases that have been put forward in the literature. The things we do need to assume is that firms are choosing the amount of information to minimize the cost of producing information. In this specific case, there's cloud prices that firms take as given and they choose the optimal computation and data to minimize the cost of information every period. And because we know that firms can't adjust their decision flexibly, we think it's a reasonable assumption. How do we think about the GDPR? We think of it as an increase in the marginal cost of data storage.

Before the GDPR, the price of data is just the price they have to pay to the cloud provider and after GDPR, that's multiplied by 1 plus lambda, where lambda is the percentage increase in the price of storing data. I want to say that we do think there's fixed costs over here, we just don't monitor them because they won't be changing firm decisions. When you think about how this affects the 1st order conditions, when you take the 1st order condition and write down the log ratio of computed data against log prices, what you find is essentially this equation.

We know log prices are going to be important because when the price of data goes up, then firms are going to be consuming less data in the optimal isocline. But also, this is going to be tempered by the amount of elasticity of substitution that firms have or the estimated parameter. There's also going to be the impact of the compute technology, as I was saying, some firms are going to be able to produce more with the same amount of computed data. And this, we're going to be separating into firm specific components, into an industry time trend and time varying shock. This poses an important identification question because we cannot separate out of this equation. It shows there's still things within a subscript.

If you just run a fixed effect for a given firm, this won't be able to identify them separately. What we need are assumptions that allows us to identify them. And what we do is to do assumptions that are similar to the ones you would be doing in order to run a difference-in-difference, as we were doing in the first part of the paper, where we assume that these compute firms for specific components can be backed out from things that we observe before the GDPR came into force. The industry time trend, we assume it's the same between the US and the European Union. So absent the GDPR, firms in the European Union would have followed the same time trend -in the industry time trend- as the ones in the United States. Essentially what lambda is going to be capturing is all of the residual variation in the shift in computation to data ratio that's not explained by the shift in prices and the unobservables. And one thing that I also wanted to say is naturally the standard anagenetic concern that observables can be correlated with prices. We also have a shift-share design for that. I'm not going to go with that in the interest of time, but we explain it in the paper.

Let me show you some estimation results. This is what we find for elasticity of substitution for firms in the European Union, both before and after GDPR. We find the elasticity of substitutions are between .3, .4. These are larger than the estimated elasticity of substitutions that other people have found for, let's say, capital and labor, which are usually between .6, .8. Essentially, we think that both data and completion are more complements than traditional inputs. We also have some results that show that computation technology or elasticity of substitution changed before and after GDPR, but we don't think these are economically important, so we don't make too much out of them.

Let me use my last five minutes to think about the production cost of GDPR and let's start by thinking about these wedges, or the changes in the cost of data storage. Remember that we have a wedge for a given firm and each firm is also set to a given industry. What I'm showing here is the average wedge for a given industry across all firms within that industry. What we see is that GDPR is going to be equivalent or it's as if firms are having a 20% tax on the cost of storing data. Going to be larger for firms where information is more important. For instance, it's larger for software than manufacturing. If we actually plot the distribution, we would find there's large heterogeneity in these wedges. We tried to think about what explains this heterogeneity and we kind of show a couple of these correlations on the paper that show that wedges in this specific case are going to be larger for the smaller firms. We are not the first to document this to say that GDPR specifically has had a particularly negative impact on the smaller firms relative to the large ones. But we do think this kind of motivates where the changes are coming from, at least the thinking of these wedges.

Let's try to think about the cost of information in production. When you actually solve the cost minimization problem and substitute everything in terms of parameters, you get this, which essentially says for a given price, for a given computational ability and for a given wedge, what's the minimum amount of cost that you would have to pay in order produce a given level of information? Or in this specific case come up we normalize it to 1 unit of information. As you can see, it depends on the prices and the magnitude of the wedge, whether it's on the elasticity or substitution because of that, it's going to vary across firms. What we'll be doing with our model is to compute what happens when the cost of information under the status quo -what we observed in the wedges- and to think about what would have happened had we not had that wedge to begin with in allowing firms to reoptimize their choices.

What we find is that information increases, but only by around 3.7%, so remember that 20% tax translates into a 3.7% increase in the cost of information on average, so we kind of at this point are a little bit thinking about how to reconcile 20% with 3.7%. If you write down elasticity with cost of information with respect to the wedge, you can find the 1st order thing that happens is naturally, if you increase the price of a given input, you need to multiply that by expenditure share of that input in order to get roughly what would be the increasing cost. You would also have 2nd order terms about how firms reoptimize decisions. Over here what we find is that -and because we find that data in computation are very strong complements- we find that firms cannot substitute between the two. And therefore, most of the increase is just going to be coming from the smaller expenditures. Data is a cheaper input of the two. The last thing we can do is to try to think about increases in the cost of information, let's try to think about how the cost of information translates to increases in production costs. This is the first time when we need to assume a production function and in the paper we go about and explain how we did. We assume there's capital glass between capital labor and information. Under this admittedly strong assumption, we find the average increase for software is around 1/3 or 2/3 of a percent, and the cost of producing information and it's going to be slightly less for data intensive industries, so that's what we have. Let me conclude what we do is use a production approach to study GDPR. We find that firms after the GDPR became less data intensive and recovered the wedges that would rationalize these firm choices. 41:08
[Applause]

[Eddie] Thanks, Diego and we'll turn it over to Devesh, in house, to discuss the paper, and again you get 10 minutes.

[Devesh Raval] A lot of slides here. Here we go. I want to start with what we know about. So almost all the literature has looked at industrial production. This is a pretty beautiful cement plant. But this is maybe 95% of what economists have studied in production.

Now that's the old economy. If you think about the new economy, the digital economy, we very little about production in the digital economy. You can pretend it's the same. We've got computers as capital, these engineers are workers ad they produce money somehow. I don't think that's that satisfying.

There are three recent papers trying to look at this and they are actually all this year. The first model is Techies which are basically engineers. They have a model where engineers then produce productivity innovation, so they are driving productivity growth. The second that's forthcoming is by Lashkari has data on In-House IT, so basically software inputs and hardware inputs. Then the paper today is looking at the newest version of that, which is cloud completing using data from Microsoft Azure where they have data on data and computation.

I'm going to now put on my macroeconomist hat, maybe my macroeconomist dunce cap and talk about some of the foundational paper here. So the foundational paper here is from Hsieh and Klenow and the QJE. They're trying to explain why India is so much poorer than the US. They're going to start with these misallocation frictions -taxes on output and capital- then they will identify those with wedges in production first order conditions. The way I think about this is, the production function says the capital should be 30% if it's 20% you're facing a capital tax. If it's 40%, in actuality, you're facing a capital subsidy. They did a calibration exercise and what they show in the quarter factual is what happens if you give India US frictions, it's going to mean a 40 to 60% increase in TFP. And this is gigantic. In the 2010s India's TFP growth is 2% a year. This is 2 to 3 decades of productivity growth just from eliminating these frictions.

This paper has two big differences from Hsieh and Klenow. First, they're explaining where the friction is coming from and it's coming from GDPR and it's going to affect the identification because it's going to be this difference-in-difference -Europe versus the US, post GDPR- and 2nd, it's not a calibration exercise. They're going to be rigorously estimating the elasticity between data and compute in order to get at the frictions.

I have a couple quick questions. The 1st is what's the net distortion? We know these negative externalities of data from privacy and data security, GDPR may be directly affecting those, improving those. But you can also think of it as a tax on data, it's also something we want to do if you face negative externalities. It's beyond the paper scope to fully deal with this. But if you want to think about it. Directionally, are the distortions in the right space? If you think about meta as the archetypical privacy violator and the largest fine for GDPR. You do see larger distortions for web services, which is what you might expect, but there are also these larger distortions for small firms, which are maybe reflecting the first fixed cost component to GDPR. That might be a bad part of these privacy rules.

Next, I want to think about how large this distortion is. They find an increase of the price of data of 20%, price of information increase of 4%. David Byrne and co-authors look at the growth rate of prices in cloud computing, so they find for Amazon's database products swelling by 12% a year, Amazon's storage product is swelling by 17%. The data effect is maybe a year and a half of the expected price decline.

There's a difference paper in the data science literature by Wu and co-authors to look at AWS in general and they're finding a 20% price decline per year, so if you compare 4% increase the price of information and 20% price decline this is just two to three months of the trend. So to me, it's pretty small. Europe is 3 months behind the US. It's hard to square these huge costs of GDPR, there's huge literature on it. But if you look at these price effects, they're really pretty small relative to what the actual price declines are.
I've got 4 main points. The 1st is directly on price measurement. The way they are measuring prices in this paper is either dollar per TB for data or dollar per core hour for compute. They're not doing any hedonic

price adjustments unlike the previous papers I showed you, and this can matter a lot. In the Lashkari paper what they find is that if you account for hedonic price adjustments, such as increase in Ram and other improvements in technology, for normal IT that's going to double the price decline. This could very much change the elasticities in this paper. I have two suggestions. The 1st is, do the hedonic correction and show us whether it matters or not. The 2nd is, you need to provide more information on the price indices because the instrument he didn't talk about but is essentially variation in price shocks across data center locations. But we don't see anything about the overall trend in prices or how it varies across locations, which I think is important to understand the identifications.

Second, I think it's helpful to show how the distortion is identified. The way they identify the distortion is a fixed effect in the regression -post GDPR minus a fixed effect in the pre GDPR- and the assumption is the pre-regulation is productivity. There are two problems with this. First of all, fixed effects are measured over there and you're differencing in two fixed effects. Second of all they expect productivity to evolve over time and that's going to mean a lot of dispersion of this. They will be way overestimating dispersion for both of these reasons and dispersion is what matters when you think about these allocation frictions. I recommend following Hsieh and Klenow. Compute this both for the US and for the EU and then compare the EU distribution and the US distribution. If you see a big increase in the variance in the EU compared the US, I think that'd be evidence of greater misallocation frictions beyond just a level affect.

Third, they end with this counterfactual exercise. It's really missing in this paper. And I think it's something they can do. 80% of the EU firms are linked ORBIS and with ORBIS you can get data on revenue, capital labor and materials. Enough to estimate production functions and I've done so with the ORBIS data with France, Italy, Spain, etcetera. If they do that, they can look at counterfactuals. In particular what's the TFP effect of removing distortions. There's a level affect, this 20% increase in data on average, but there's also just the dispersion and distortions, which is motivating Hsieh and Klenow. To do so you'll have to estimate a production function and I think it would be natural if you were to assume CES across all inputs. You're there finding elasticity between data and computer .3 to .4. Whereas I found for capital labor about .3 to .5 and the Lashkari paper finds between it and non-it about .2. These are mostly in the same ballpark and as a starting point could assume CES across all inputs. Then let's do the Hsieh and Klenow exercise with the distortions.

Finally, I have a question if the macro production function different from the micro production function? If you think about GDPR, it is essentially a macro shock affecting everyone in the EU and there's been lots of interest in macro economists, like Chad Jones, thinking about the macro production function once you have data. It's been known since Houthakker 1955 the micro production function can be different than the macro production function. So Houthakker found if you have elasticity of 0 at the micro level, you can have an aggregate Cobbs-Douglas. What Oberfield and I found is that the aggregate elasticity is a convex combination between the micro elasticity of substitution and the elasticity of demand. The idea is there is some within firm substitution indexed by the micro elasticity and a cross firm substitution index by the elasticity of demand. Nakai, which is the perimeter between these two, is essentially a variance of capital shares. The more dispersion there is in shares, the more scope there is for cross firm substitution. For capital labor substitution refined .5 at the micro level turns into .7 at the macro level, but the Lashkari paper finds bigger differences with IT because there's more dispersion in IT than non-IT inputs across firms, so they find about .2 at the micro level to 1 at the macro level, so this is not that different than Houthakker. I think it'd be pretty interesting here because the paper shows there's lots of variation in data to compute across firms and we know from the Lashkari paper there's a lot of distribution from IT to non-IT across firms. And then finally, there's all this variation coming in from the GDPR distortions as well. I think that could be a big difference between the macro production function and micro production function. And I'm sure you'll get a lot more citations if the macroeconomist can calibrate the models to your paper as well.

I want to end with maybe taking off my maker economist hat and printing on my regulator hat. We're doing lots of work thinking about costs and benefits in this space. First of all, comment we are working on a big privacy data security regulation and 2nd, President Biden asked us this week to think about regulating the AI sector as well. What the economists have been in this room have to do is measure the cost and benefits of this. This paper will be helpful in order to think about those issues.
Thank you very much.
[Applause]

[Eddie] If we want to have our speaker come back up and we have time for a couple questions. if you want to raise your hand, we can run around and give you the microphone to ask your question.

[Q&A]
[Q] Thanks. Really interesting. A lot of people talk about this sort of stuff is the reason why the EU doesn't have big tech the same way the US does. Both the pros and cons of that. And I'm sort of wondering, it sort of seems like even though GDPR is sort of like a per se rule, but it's enforced -it's pretty discretionary. First of all, I'm wondering if enforcement, like this wedge, is nonlinear in practice, which affects your estimation, but then also, if it's somewhat -if you could run counterfactuals where it's discretionary. Where we know the production function for IT firms or Tech firms is different than other industries and maybe we're going to make the wedge smaller for those than for others or something like that?

[A: Diego Hernández] In the end, I think that given the production function, the way we recover the wedge estimates. In principle, they could vary differently across industries. We find a little bit of that. I would have to think a little bit more about whether the wedge being very nonlinear in practice can matter for regulation.

[Q] Hi. It sounded like you have price variation in cloud computing services that's running in the background here, so can you think about -maybe do this. And I didn't understand it in the presentation- You know the equivalent of a 20% -a reaction to a 20% increase in compute costs. Can that inform your estimate of the behavioral response to the GDPR? So, in the end, you have this exercise where you back out the wedge, the wedge is somehow like -if the price is just higher by 20%, right?

[Diego Hernández] Assuming this compute technology didn't change plus the time trend being the one in the US, the difference between the serve choices and what we would observe under the prices is ascribed to this wedge, so essentially counterfactually if we were to allow this computational technology to change relative prices by 20%, then in principle, you would expect the same thing.

[Q] And do you look at that?

[A: Diego Jiménez Hernández] We haven't. But we should definitely check it for –

[Eddie] Last one.

[Q] Great presentation. Just curious, do you have any areas of content distribution pertaining to streaming in the areas of unforeseens, as it applies to the wedges and the cost.? Is there any particular are there any particular areas that you see as unforeseen areas?

[A: Diego Hernández] As unforeseen areas? No, so given the type of analysis that we have, we cannot distinguish between let's say, content distribution specifically, but yeah, that would be interesting. And to the rest. Let me say one thing, thank you so much for the comments. I don't know where Devesh-- thank you so much for the comments, the only thing I wanted to say is one of the things you mentioned, was to think of whether we could estimate the existence of a wedge in the US and in the EU that changes after GDPR, that requires stronger assumptions than the ones that we are making. We did that in the first part of the paper and we found very similar distortions to ones that I'm showing you in the presentation, but 0s on the… Sorry, let me rephrase that. We find wedges in the European Union similar to the ones that I just showed you. And zeroes if we were to estimate the same in the US. We think it's reassuring because it's as if the method essentially estimates a 0.
[Applause]
[End of Q&A]

[Eddie] Alright. Thanks everyone. comment. Now we have Nils from Kellog School of Management presenting on estimating the value of outside data to advertisers, evidence from meta.

[Nils Wernerfelt] Good morning, everyone. It's a pleasure to be here. This is my academic job market paper. And this is joined with my wonderful co-authors, Anna Tuchman who's at Kellogg, Brad Shapiro who's at Booth, and Rob Moakler who's an employee at Meta.

What's the motivation for this question? When you think about advertising today on major digital ad platforms -- Facebook, Instagram, TikTok, Snapchat, Twitter, Pinterest -- many people think what's powerful about these platforms is the user generated content that people created on them. So for example, you can deliver ads based on user profiles on Facebook or what kinds of photos people are liking on Pinterest.

It turns out that's only a fraction of the overall data platforms used to target and deliver ads, and all these platforms have technologies that enable advertisers to track and monitor user behavior outside of the walled gardens of the platforms. Web pixels for example, is 1 technology that enables advertisers from these platforms to do that. And this data, this is why, for example you might see an ad on Snapchat based on websites you visited in Safari or actions you took in the Old Navy ad might affect as you see on Facebook. And this data is actually viewed as some of the most important data across the board for digital advertising today. Intuitively if you want to think about delivering ads based on-- Nils like soccer and he likes it on his Facebook page, that's different than knowing Nils is looking at websites for concerts or soccer tickets or looking at what's for sale in the Ticketmaster app. So, this data is viewed as very powerful, but it's also now policy relevant, in particular in the number of countries around the world today, there's pending regulation that might limit the ability of advertisers to use this data. And there's also looming product changes from Google and Apple that might further restrict this ecosystem. And many of these changes are motivated by an appeal to consumer privacy. So consumers don't like being tracked across these applications, they don't like this data of being shared. Much of this is motivated by an appeal to consumer -- The consumer privacy motivation, but for any holistic policy evaluation, we need to think about what are the effects on other parties. Right now there's no logged evidence on how valuable that data is for advertisers.

That's where this paper comes in. We want large scale experimental evidence on how valuable this data is for ad effectiveness. You might think this is an abstract question and to put some numbers on the magnitude we're talking about here, those of you with iPhones likely have seen this prompt that appears, 'ask app not to track', and that was introduced in April 2021 by Apple, iOS 14.5. If you opt out of tracking, that prevents the focal app from collecting data about you from other sources. Therefore, you can't use the data in Safari to deliver ads on Snapchat, for example. And that one change is estimated to cost Facebook for $10 billion in revenue alone and afterwards announced the evaluations of major tech companies fell between $100 and $200 billion, so the stakes here are large. This is an active area of policy development and it's important to understand the trade-offs here.

What specifically do we do in this paper?
Advertisers use offsite data in a number of ways. We do not attempt to do a holistic program evaluation of all different methods. What we do is focus on one arguably primary use case and estimated effectiveness with and without such data. So, in particular what we do is we take a large sample of advertisers on Meta who are using this offsite data according to this use case today. And we do two things.

First, we use a novel experimentation platform to randomly hold some users out from seeing the ads. What this allows us to do is generate baseline estimates of advertising effectiveness, so we can say you have an advertiser today who uses offsite data, 'how effective are your ads'?

The second thing we do is take those campaigns and modify a small fraction of their delivery to reflect a loss of offsite data. We then hold a small fraction of users out from seeing those modified campaigns and that lets us see if the same campaigns today, how effective are they with offsite data and how effective are they without offsite data. This sets up a large-scale meta-analysis across our sample where it can make pretty general effect statements about both the baseline effectiveness and the change in effectiveness from using this data.

When you think back to this paper, I want you to think about two main takeaways. The 1st is that baseline effectiveness, in particular, because we're able to experiment on live campaigns, we have minimal concerns about file problems for example. A lot of past advertising meta-analyses had relied on who's been using measurement tools, and it turns out only a very small fraction of advertisers do that. So here we are able to make pretty generalized unselected statements of advertising effectiveness. In addition, past ad studies, it's hard to know is the advertiser optimizing for direct response or for brand. So, for any outcome they're looking at, is this actually the one that the advertiser cares about. In a feature of our data is we actually know, the advertiser tells us what they are optimizing for, so we can make statements directly about the outcomes that advertisers are caring about. Finally, our sample is sufficiently large, we can make pretty tight estimates of advertising effectiveness and not just the average or the median, but we estimate the entire distribution of advertising effectiveness. That is the first contribution here, is understanding the baseline effectiveness of online advertising on most major platforms.

The second contribution is the policy relevant one – that I mentioned earlier- so the change in effectiveness. For this larger preventative sample, how much less effective are they if they lose access to this offsite data. So, thinking about this paper, these are the two things I want you guys to have in mind. Now to go through the requisites and caveats to this analysis. This is the partially equivalent analysis, so policy changes, like the Apple one I mentioned, when that happens prices change and you have endogenous entry and exit of advertisers. We don't get into any of that, we just do a large-scale experiment, we hold prices fixed and we sort of stick to what we can closely identifying this experimental setting. The second thing is similar to the 1st paper, we don't try to make any statements about social welfare. We don’t get into estimating consumers evaluations of privacy. That's an important and thorny area. Here, we simply focus on the advertiser side. So what is the effect on ad effectiveness. Finally, there are, of course, other platforms and other ways that advertisers use offsite data. Our sample in Meta is economically meaningful and the use case, we chose as I'll describe shortly is offered on many different platforms. So, there's some shredded external validity here, but we of course can only experiment on Meta.

That’s sort of the high-level motivation here and I want to -before diving into what we do- I want to provide some high-level background on digital advertising and our specific use case that we focus on.
These are screenshots from Facebook ads manager, if you want to buy ads today on Meta, this is one of the interfaces through which you can go and purchase ads. And if I can use my –oh look at this- you can see that there's some fields here which users are familiar with generally. You can see we ask how much you want to spend, what's your target audience, but we also ask this question that sometimes people are surprised about, which we ask what is your objective? You, as an advertiser, what do you care about? People are sometimes confused, why does Facebook care what my objective is? If I give them money and say, ‘here’s my target audience, can't you just spend the money on my target audience’? Here's the intuition on why we ask for that. Suppose that what you really care about with your objective is driving likes on your Facebook page. You could tell Meta, ‘Here’s my target audience, here’s my budget’ and Facebook could just spend your money uniformly amongst your target audience. But notice that, as Facebook is showing ads to people, they can actually see who is going and liking your Facebook page. Suppose when they are doing this that no women are liking your Facebook page.

Facebook could keep on spending your budget uniformly across the population. But it's probably not good for you as an advertiser, because women are not going and liking your Facebook page, so you're spending your budget on users who are not doing your objective action. It’s probably not good for users to come up because these women are seeing ads that are less relevant for them, they’re not actually engaging with them. A way to potentially improve outcomes here for both the advertiser and consumer is to train a model to predict the probability that a user will go and like your Facebook page and use that to dynamically inform the target audience.

In this case, Facebook would start spending money on men and women and the model would realize that women are not engaging with the page and it would shift the budget to only spend on men. Arguably, under some assumptions, this is good for both users and the advertiser themselves. This is the high-level intuition behind what’s known as delivery optimization that Facebook and all these other major ad platforms do. Of course, if you're an advertiser, not many advertisers actually care about likes on their Facebook page, they care about other outcomes. For example, purchases. Those of you who have clicked on ads -and there are more than you might think- on Facebook and these other platforms, you know that when you click on the ad you are taken offsite. So you’re taken to Safari, Chrome, to a mobile application. When you leave the walled garden of these apps, they have no way of knowing in their own world what actually happens offsite. So, when you go into the Old Navy app, they don't know if you’re making a purchase or if you're just browsing. That’s where web pixels enter. An advertiser can put a pixel in their website and if a user goes and makes the purchase the pixel can fire and tell Facebook that this user just went and purchased something.

Now, in the previous example, we had this left-hand side variable and the prediction problem of page likes, so now with pixels you can actually get this other left hand side variable that advertisers care about more, purchases, so now you can actually optimize delivery for purchases. Facebook shows the ad for people, see who makes a purchase and dynamically changes the target audience to show it to people who are predicted to go purchase the product.

Three comments on this. So, one; think about how powerful this actually could be for businesses, so you may not know who your targeted audience is and this is a way to actually provide data driven selection of your target audience. This is very hard to do over TV, radio or other channels. It is arguably a major upside of digital advertising today. Second thing I wanted to mention is that this is not retargeting. This is a much more fundamental aspect of ad delivery. This is not ‘I'm going to show the ads to users who visited my website’. This is “I'm going to train the model and inform the entire target audience based on who's taking the advertiser specified objectives”.

Finally, one thing you may have noticed in everything I just said is there’s no discussion of incrementality. So, one of the criticisms of the optimization is that you don’t actually know are these sales driving incremental returns, are these ads driving incremental returns? Maybe you're just being really good at showing ads to people who are already going to go purchase the focal product. That’s the point we take seriously in our experimental design. And as you’ll see, we actually constructed our experiment to estimate incrementality. And again, we sort of experiment on Meta, but I wanted to highlight that we chose this use case, this optimizing for offsite purchases because this is a product that nearly every major ad platform offers today. So a lot of digital advertising today relies on the sharing of data across applications and optimization for outcomes that occur outside of the focal application of interest.

To recap, I said advertisers use offsite data in many different ways. The main use case we focus on here is off site conversion optimization, so you're optimizing for these purchase events. Intuitively, the counter facts we consider is, if advertisers lose the ability to optimize for this offsite action, this purchase event that happens outside of Facebook for example, what will they optimize for instead? The counter facts we consider is click optimization. So here, intuitively clicks on the ads are the last thing that the platform observes as the user exits the platform. So, you click on the ad and you leave the walled garden. And so the counter facts consider is optimizing for clicks. Here Facebook shows the ads to people, see who clicks on the ad and optimizes delivery to show them to people they think are likely to click. This is, in practice, another popular optimization objective and what we focus on for our counterfactual.
Here is the whole experiment in four rectangles. We take a large sample of advertisers who are optimizing for purchases and the first thing we do is take 5% of their target audience, 5% of their budget and we say on this 5% of users, deliver the ads according to purchase optimization. Within this user segment you are going to show the ads to a bunch of people, see who purchases and dynamically find this target audience.

Initially in that target audience we're only going to hold 10% of users out, so these are these two upper rectangles here. And what this does is allow us to estimate the baseline effectiveness, so for advertisers today who are using offsite conversion optimization, how effective are their ads? The second thing we do is take another 5% of their target audience and another 5% of their budget and we say on this 5% of your target audience, we're going to deliver the ads according to click optimization, so we're going to show the ads to a bunch of people, see who clicks on them and then it randomly holds out 10% of those users from seeing the ads. This means for the same campaigns; we see how effective they are with offsite conversion optimization and click optimization. There's a lot of engineering work that goes behind these four rectangles. This might be deceptively simple, but this was many months of blood, sweat and tears to set up. This is the experiment in a nutshell, so you can see within the same campaign how effective it is optimizing with offsite data and optimizing without.

All right, what is our sample? I mentioned a few times that we have this large sample of advertisers. To run this experiment, we sent the notice out in Facebook ads Manager and this is a mock-up of what the notice would appear like on the mobile version of the app and we described the experiment and gave advertisers the opportunity to opt out if anyone wanted. We sent the notice to about 4 million advertisers and then the people who were included in the experiment, the people who logged in and saw the notice, did not opt out, and then run ads a few weeks later when we actually ran the experiment. In practice, only a small fraction of advertisers opted out. When we compare the demographics for our final sample versus the relevant population, they are very similar across demographics. We feel pretty good about the representativeness of our sample. It remains a point to call out, that Meta publishes best practices for how to use offset conversion optimization and intuitively, if you have a prediction problem and your left-hand side variable is very sparse, you might not be able to train the model as well as if there's more

signal. So, Meta recommends that you hit a minimum of about 50 conversion events per week to use offset conversion optimization and it turns out a large share of advertisers aren’t hitting that.
I’m going to talk about today is the upside of the offsite data and restricted advertisers who are actually hitting that minimum. Who are advertising according to the best practices. In the appendix we reran the analysis on the full sample and you find results of similar magnitude, but today I'm focusing on this one subset of our data. When you do all that comment you're left with 70,000 advertisers, so that's our final sample here.

And I know in these text slides it's easy to lose track of what we're actually talking about, but here's the map of where our sample came from and it came from almost all of the world. We have broad representation from verticals and sub verticals that advertise on Meta. If you want to think about the typical advertiser experiment, you should think about an American E-commerce company. But again, this is a large global sample of advertisers.

Okay. Now, what are the results? What do things look like for these 70,000 advertisers? Here's the empirical distribution of treatment effects, so across 70,000 advertisers, this is that first rectangle.
This is looking at baseline today. You optimize for offsite conversions. How effective are your ads? This is empirical distribution of the 70,000 treatment effects and what you can see consistent with past advertising analyses, there’s a skewed distribution, so there's this long tail of advertising effectiveness. What you don’t see here are the standard errors on these estimates, because we experiment on a small fraction of traffic, these treatment effects have massive standard errors. In particular notice that there’s this mass below 0 here, which suggests that for each dollar advertisers are spending, they’re actually generating negative customers. When we do our estimation procedures you see that much of the mass actually disappears. This is the raw empirical distribution treatment effects. Again, the outcome variable is cost per incremental customer, so for every dollar you spend, how many incremental customers are you generating.

That’s the baseline estimate and now this is the empirical distribution of treatment effects for the change in effectiveness, so this is bottom rectangles minus the top rectangles. This is how many incremental customers you’re getting from click optimization minus the number you’re getting from purchase optimization. You can see how many fewer customers you’re getting when you optimize for one versus the other. And you can see here that sort of consistent with this data being valuable, there's this mass below 0. A lot of the mass here suggests that advertisers are getting fewer incremental customers when they optimize for clicks versus the offsite outcome.

These are just the raw distributions and now there's an exercise to actually say what’s the latent unobserved true distribution of effects? We do a meta-analysis. We use an empirical based method to estimate the entire latent distribution of effects. For time reasons I'm not going to go into that today. But when you do that model fitting procedure this is the results you see. The top row here, the baseline estimates, for every $1000 you spent, how many incremental customers are you getting. In the bottom row is the change in effectiveness, so for every $1000 you spend, how many fewer incremental customers are you getting when you optimize for clicks versus purchases.

Some things to call out here. One, if you look at the median estimates, the median advertiser for every $1000 they spend generates about 23, 24 incremental customers. The median loss, the median change in effectiveness is about six fewer. You can convert this in terms of dollars for incremental customer. And when you do, you see the estimates, the median cost for incremental customers about $42.00 and the median loss increases about a $56, 35% increase in cost and you have to optimize without the offsite data.

This is sort of the first punchline from the paper, which is that this offsite data is quite valuable in this experiment and when we lose the ability to use this offsite data, this estimate here suggests that cost increases quite substantially. Some other things to point out, we have estimated these entire distributions so we should get our money's worth from this. First of all, note that there's this wide range, a wide dispersion of ad effectiveness in the baseline case, but you don't see that in the change of effectiveness. The 2nd row here is actually quite narrow and what this suggests is that the effect of using this data is closer to homogeneous shock, so you have this huge distribution of advertising effectiveness, but the distribution of harm is pretty concentrated, so it's not like there's a few guys that are really getting hit by this and everyone else is unaffected. It's pretty close to a uniform shock across people. I'll also call out that the baseline estimates of effectiveness are actually larger than a lot of the ones that are in the literature. And by larger, I mean, you have to spend more for each incremental customer and we think this reflects a larger unselected sample here, so a lot of estimates in the literature are either based on non-incremental correlations or based on highly selected samples and advertisers that are actually using the measurement tools. I'll mention that about 10% of advertisers actually generate more incremental customers per dollar under click optimization versus purchase optimization, and we think that's speaks to the non-incremental nature of optimization. So what's happening in that case is that using the offsite data, using this extra data, you're more likely to find people who are more likely to purchase the product versus when they use click optimization, they're casting a wider net, but still for 90% of advertisers in our sample, their ads are less effective without the offsite data.

That's the main results from the paper and now if you can indulge me in some additional results here and I want to call up two specific ones. One is comparing small versus large scale advertisers and the other is looking at long-term effects. The motivation for the first one is that we know small advertisers disproportionately rely on online advertising and if we see large effects this might be potentially relevant from a competition standpoint. Indeed, after Apple launched this iOS 14.5, there were anecdotes of small businesses where their ads were no longer effective, and they actually had to go out of business. So, there's a question here, if we look at small versus large businesses, what's the distribution of effects?

On this next slide, I repeat our estimation procedure where we do a median split on ad spend, so we look at advertisers who spent below the median amount of the sample versus above. There's a lot of lines on this graph, but the blue ones are small scale advertisers, and these are the guys who are below median spend and the dashed line is the baseline effectiveness and the dotted is the change in effectiveness. The main thing to see here is that the two red distributions are shifted inwards, so smaller advertisers, their ads are more effective. But they're also hurt more by losing offsite data. And so if you look at the medians for the loss distributions, you see that the median for the small guys is about five times as large as that for the large guys. What that means is in an absolute sense, small advertisers are hurt more by losing offsite data than the larger advertisers.

The next result that I want to call out is looking at long term effects. And there are two motivations for this. One is that advertisers care about sales, but also about customer lifetime value. You could be in a world where offsite data finds a bunch of impulsive people, people who buy products, but then don't actually engage in repeated purchases, or they make purchases that they regret. On the user side. And this is more speculative, but there are sort of claims that this surveillance capitalism actually sort of tricks users into buying products that they don't want. And so, if we look in the future that consumers are still purchasing these products, it suggested that they actually were not tricked, that they actually found products they liked that they're selecting to keep on buying. To do this, we have the user IDs and who has consigned to treatment and control for each advertiser, and we can look six months into the future of their purchasing behavior and see if this user randomly decided to click optimization, do they hold out, what are they purchasing in the future?

When you do that, you see the ads that are delivered with offsite data in terms of substantially more long term per dollar than ads delivered without offsite data. This is consistent with the story where the offsite data is helping to improve the match quality between products and users. And it's helping users find products that they otherwise wouldn't know about and that they then self-select to keep on purchasing. Those were the two additional analyses to call out. I want to now get out my armchair and talk about some potential applications of this for different parties here.

Looking on the advertiser's side, this can help directly inform understanding of willingness to pay for offsite conversion optimization, so you have advertisers thinking about optimizing for purchases or clicks, this can help you think about your willingness to pay there. I think a deeper and more interesting point to call out, two people had this great discussion of this paper this over the summer, and she called out the potential for data markets here. So, in some sense if we think that users may have a certain evaluation of their privacy here, but if what we are seeing here is that this data is actually quite valuable for advertisers, then maybe there are gains for trade here and we should set up potentially markets that allow advertisers to compensate users for sharing and creating this data. I have to give credit to JP Dube for is that we actually do this already in a number of different settings. Think about every time you go to CVS with your customer care card, what you're doing there is they're giving data from you and you're getting discounts, so it's a similar value transaction there.

We haven't seen much of that in digital space. But I think that is one potential direction going forward. Another application of this work is to understand the value of this data for platforms. In the aftermath of iOS 14.5, a number of platforms for example, TikTok, have launched products where users can purchase products within the site itself and TikTok has marketed this as a seamless purchasing experience for users, when in fact, this is a way to get the conversion data's first party. Now they actually see purchases on TikTok and can optimize for the outcome. I think we'll see in the future more platforms expanding the walled garden to bring its purchase data from offsite to onsite.

Another implication here is investing in these privacy enhancing technologies, technologies that potentially allow us to obtain some of the upsides of this personalized data while protecting consumer privacy. I personally have not seen a report on these technologies, but I think this is a way that we can potentially get the best of both worlds. Finally, the effect size here is sufficiently large there might be potential competitive implications that's beyond the scope of our data. But you might hear later today about some of this last bullet.

With one minute left, that's a wrap. Thank you all. Again, the higher level thing here in this large scale experiment under estimated ad effectiveness without and without this offsite data and we found evidence of large effects, that this data is quite valuable for advertisers today, in particular for small advertisers, and I look forward to any questions, feedback and whatnot, thank you so much.
[Applause]

[Eddie] And now we've got Samuel Goldberg from Stanford for the discussion.

[Samuel Goldberg] Thank you all so much for having me. I'm excited to be here and appreciate the invitation. I'm excited to discuss this excellent paper by Nils estimating the value of offsite data to advertisers: evidence from meta.

I want to start here by taking a step back and talking a little bit about the context in which this paper is happening. I think Nils did a good job of going over this. There are two big themes going on in ad tech today. The first are these large, complex privacy regulations. We heard a little bit about that from the last presentation These include the GDPR, California Privacy Protection Act (CPPA) and more recent sort of competition authority, like the digital markets act and various other different regulations, both in states and across the world. The second major theme here is that there's this big innovation towards more privacy centric advertising, so this includes Google's Privacy Sandbox, Apple's Tracking and Transparency Framework, as well as a variety of other proposed measures. In general, the industry thinks this is the death of the cookie, death of a lot of this online tracking that's happening and the technologies that are used to do that right now.

What does this cumulatively mean? There's a big push towards limiting the use of third-party data online. And that is where this paper will come in. Most of us are economists and we recognize there are probably tradeoffs to limiting these types of technologies. On the good side, there's likely benefits to privacy, so this will limit tracking across platforms. And that's something that consumers keep demanding. On the bad side, there's going to be significant cost to measurement. That includes both firms --it will be more difficult for firms to learn if their ads are working if their ads are incremental. And it's also going to be much more difficult for them to get those ads to the consumers that they care about. It will be much harder for them to target their advertising, and there's also potential implications for competition in these different markets as well, and we'll talk a little bit about that at the end.

What is this paper going to do? They're going to strive to answer this question of how much value does third party data deliver to advertisers. This is a tough question to answer for three big reasons. The 1st is that we really care about incremental value and that’s an extremely noisy thing. It’s very difficult to estimate incremental value of advertising. There’s a whole literature on why and in what situations that is the most difficult and easy to do. There’s a lot of heterogeneity here and so you’re really going to want to have scale. Different firms have very different returns for advertising, different industries have different structures and we will want to see this for a large swath of firms. And then finally comment nonexperimental methods just generally don’t work in this setting. There's also a large literature that shows that if you’re not running an experiment, you’ll never get incremental measures.They’re going to tackle each of these challenges in turn and that’s part of what makes this paper so unique. The first thing they’re going to do is they’re going to partner with Facebook to run an exceptionally large experiment here. Which means they’re going to have this scale to show us how this heterogeneity and treatment effects is across advertising. They have 70,000 advertisers using third party data on Facebook's platform.

The second is they will run an experiment and this experiment is pretty simple. It has a differences-in-differences intuition. There’s going to be a group where we will estimate the incrementality under the status quo. That is for the firms that are using third party data. That’s this purchase optimized targeting strategy. And then we’re also going to estimate the incrementality under the counterfactual here, which is if we’re going to target based on 1st party data or the clicks that occur on the Facebook’s platform. Then you’re going to difference those two things to get that treatment effect that we care about, which is the value of third-party data for these advertisers. Finally, they’ll introduce this meta-analysis to deal with the noisiness of these measures that Nils touched on.

This is an amazing paper. There's so much good going on. But what I want to do is show 1 plot that Nils didn’t show in his presentation. This is in the appendix, and this plot says everything that needs to be taken from this paper. This is a really amazing plot. On the X axis, we have the incremental return to the purchase optimized, that’s in the status quo setting. This is how much advertising is worth on the X axis Then on the Y axis, we have the counterfactuals. That’s how much advertising is worth in the counterfactual. We are just using first party data. What you can see is the density of the treatment effects is under that 45-degree line. Which is saying you are getting more, you're getting higher returns when we use purchased third party data.

Just to emphasize what is so great about this paper. Scale makes it generalizable, so we have these 70,000 firms. Even if you don't think that we can take this outside of the context of Facebook, Facebook is interesting. It accounts for 10% of digital ads, 20% of mobile ads and is an enormous platform for these firms to target and acquire customers. I think this experiment is simple, it’s intuitive and it’s careful. It lets us separate out that incrementality and measurement problem, which is a very novel thing.

I can’t just say nice things, so I’ll give a couple of suggestions, really quick. The first thing I think this paper could have improved on a little bit is to lean more on the model free evidence. I think what is nicest about this is that you can just plot these distributions of treatment effects and read as you will. You could do a bunch of technical work to make that meta-analysis a little bit more robust, but really the model free evidence tells the full story So leaning into that by showing more heterogeneity in these treatment effects cut by different types of advertisers or which target audiences they’re going after, would really be helpful for us to understand which types of firms are going to be affected the most, and what that means for policy in this setting. I want to be careful in interpreting their small versus large firm results here, so I do think they suggested that small firms are hurt more, but you'll note that the cost per incremental conversion. The levels here are very different for large and small firms. Which does suggest that they are probably targeting different types of audiences and so it's a little tricky to think about what that means for who is competing with who and how we want to interpret these in that context.

If you could break this up by target audience or browsers, that would be super helpful for robustness here. Finally, I think also in terms of the long run effects, long run attribution is extremely difficult. It's a very difficult problem in advertising effectiveness literature. And so if these results are true, they are really impressive and important, but I would like to again see more robustness and exploration here. It seems really easy to break this out by durable versus non-durable goods and we should expect very different patterns in long run attribution and thinking more about heterogeneity. Which types of firms are driving these differences.

With that, I'll conclude a little early here and say this is a really nice paper and it provides scale and generalizability to estimate a really important policy question. I'm really happy to discuss it here. Thank you.
[Applause]

[Eddie] Great. Thanks. Now we have a few a have minutes for a couple questions.

[Q&A]
[Nils Wernerfelt] Thank you Sam, that was a great discussion. Your suggestions, I agree on all of them. Thank you. It's a rare thing that that happens.

[Q: Sam] I'm Sam from the Federal Reserve Bank in Minneapolis. I have two related questions. One is, what if this is all completely 0 sum? The ability of one person to advertise slightly better just reallocates output? Would I then conclude that that was valuable? The second related question is about scale and generalizability to estimate these incremental values, you have to run this experiment where some people have access to this technology. If I expanded that to everybody, would I think anything would change?

[A: Nils Wernerfelt] The second one, could you say that again?

[Q: Sam] If I allow everybody to skip the line, then it doesn't matter whether I've allowed anybody to skip the line. You are estimating these incremental effects off of allowing some people to have access to some technology, like a third-party information. Do you think that those estimates of value would scale up if just everybody did the same thing? It relates to the first question as well, of whether we think this is valuable or if it's entirely zero sum.

[A: Nils Wernerfelt] Perhaps some unsatisfactory thoughts and response to that. If all advertisers are indifferent between Google and Facebook, Facebook ads get nerfed and they just substitute to Google and there's no less well off, it's a very different interpretation than everyone is effective and the advertising technology degrades. We can talk about those as outside the experiment, as more GE style effects, and so I think I will say that because so many of these advertisers or these advertising platforms use offsite data, I think it's more likely to think of this as a shock to the industry. Maybe this advertising technology is just getting worse or better. And if everyone kind of moves up or down, it just effects spend and maybe all that's happening here is that Facebook gets more money. I think to answer the question you want to look at, what effects the ATT policy actually had on firm outcomes. Is this actually driving the changes the number of businesses that are being created and wages and employment and I think we have some work cooking to look at, and this is an empirical question.

[Q: PHD Student] I'm a PHD student at the University of Maryland. Maybe something that would be very easy to check that I had in the back of my mind is consumer spending. What happened when that iOS update was launched? If it didn't really change, maybe that would be some evidence that this is more 0 sum.

[A: Wernerfelt] I might bug you for data on consumer spending, that sounds good.

[Q: Georgetown] Hi, I'm from Georgetown University. My question is: can we evaluate the value of this data that we are studying by looking at the decrease in predictive performance of predictive morals for retargeting or for other purposes, in terms of the (*inaudible*) or an F1 measure or something?

[A: Nils Wernerfelt] I had not thought about that. We could do that. In practice. I worry not all that data might be logged for a pre/post comparison after the iOS change, but there could be other experimental ways to get at that and to just- Correct me if I'm wrong, what you're saying is can you see how the variance of the expected click through rates and expected conversion rates would vary, so you lose this data and do the estimates become less accurate and how much less accurate?

[Q: Georgetown] You can also use shapely value decomposition to see if you have a larger set of variables.

[A: Nils Wernerfelt] That's great. We've not looked at that. And I will find you after and talk more about this.

[Q: Peter] And then I'm from Columbia University. Are these external data traded in a market and if they are used as a compliment to internal data, would we be worried that these very big companies build a barrier to entry into the advertising market by being able to use them very effectively and other companies wouldn't be able to do so?

[Q: Nils Wernerfelt] Potentially. So, you think about if some of these incumbents have lots of data on who's purchasing, they can use that to train better models and you're start up, you can't do that that. They can potentially form barriers to entry. Yeah.

[Nils Wernerfelt] Thanks so much.
[Applause]
[End of Q&A]

[Speaker] We are going to take a short break and we will reconvene at 11:15.

[Eddie] We're going to get started again in a minute or two. So, if you could start making your way back to your seats that would be great.

All right, everyone. Well, we are going to get back with our first keynote. We have Steve Tadelis here, who's at Berkley. His research is primarily now about Ecommerce, economics of the internet. Things you're familiar with from the papers he selected today. Steve is going to be presenting on targeted digital advertising: challenges and promises. Steve, take it away.

[Steve Tadelis] Thank you very much, and I notice that today I am feeling, especially old, and I wondered why, and it's not because my hair is so much whiter than Aviv's, we were classmates -or at least what's left- but it seems to me like yesterday that I interviewed Devesh and Ted when I was working at eBay and now they are running half of the FTC. That is really weird for me. So anyway, with that, I'm going to do a little trip around different things that I have worked on and that many of you are familiar with the space, and hopefully it will be interesting and offer some questions to think about.
So, one thing.
I like this quote that I'm sure many of you have heard in the past. I like calling it Wannamaker's challenge. John Wannamaker was an important retail person at the turn of the 20th century. People claim that he gets credit for inventing the department store and he has a famous quote where he says 'half the money I spend on advertising is wasted. The trouble is I don't know which half. Target digital advertising has been often sold as the technology that solves Wannamaker's challenge. If we think of display ads, those are kind of the new version of billboard or newspaper ads, but there is better targeting based on the interests of people who are browsing online. If we think of video ads, they are another form of display ads that replace the new TV ads. Better targeted towards interests. Then we have a completely new quote unquote at 20 years old, but a new way of creating advertising. Which is paid search, where people give a signal of intent and then together with some demographic data and possible interests, we can get better targeting. Last but not least, there are going to be some things that are going to be very similar to what Nils has covered and some things that he alluded to that I'm actually going to show you. Social networks are a lot more demographic information and on possible interests that help to target.

The questions that I think are interesting here or at least interesting to me, one; to what extent is Wannamaker's challenge solved by these digital ads? Second, do businesses and especially small and medium sized businesses benefit from this kind of targeting? And finally, what about that tension between privacy and targeting that we've already heard about today, what does that mean for welfare? And when I

get to that, my answer like the answers before are going to be very partial, because, we are getting more and more evidence on the cost of these privacy regulations. Every time we talk about the benefits, it's like people want privacy. Well, do they? So we will talk a little bit about that.
First, I will just touch very briefly on aid search advertising, this is a paper that together with Chris Nosko and Tom Blake, we published about 7/8 years ago -wow, more than eight years ago, another reason to feel old- in Econometrica and it's the easiest paper, the simplest paper ever published in Econometrica, I'm not kidding. I'm dead serious. And what we found is they are using a very large-scale experiment on eBay is that, for eBay, the average returns on advertising spent for their paid search expenditure which at the time was on the order of $250 million a year, they have a negative 80% ROI. Not something to be proud of, and this paper almost never saw the light of day. But we were lucky that we took advantage of a day where the president was very upset at Google, and that's why he approved the publication of this paper.

But the reason that there was something interesting on this paper from an economics perspective is that we were able to use proxies for how informed people are about eBay. In particular we use their past behavior vis-a-vis eBay. We knew who never bought on eBay before, during the experimental period, who bought only once or twice in the past, so they are not very familiar. Versus people who come to eBay all the time. And what we were able to show is that consumers who we call less informed about eBay, those who never had an eBay account or are relatively new or haven't purchased a lot. We saw that those people are heavily influenced by the ads versus people who would come to eBay with a reasonable frequency, they are not influenced really influenced by ads. So, I would think, for example, if I ask people in this room, 'how many of you purchase regularly on Amazon,' I don't need you to raise your hands, I know that it's everyone. So, what's the point for that company to advertise to you if you're going to go there anyway. What would happen there is that even though some users -this is back in 2012- there was an experiment done where eBay had more prominence among people, less and less so as time goes by, that most of the advertising clicks were happening by people who are going to eBay anyway, and therefore, that was wasted money.

In eBay's case, very much wasted money. Now of course, this is likely to be different for small and medium businesses because they don't have that brand recognition that a company like eBay has. The problem is that the smaller you are, the harder it is to measure the effectiveness of advertising. So, I was asked three years ago -there was a little piece on Freakonomics radio where this was made was made into a nice, digestible piece for a lay-audience. I remember getting a few emails from business owners 'oh, could you help me?' And the answer was, 'well I could talk with you, if you want.' In one case, I actually recruited my older son to help them do the data analysis on a similar experiment. This was a medium sized business, ad revenues of a couple 100,000 a month, maybe, so that's not a tidy business, and the signal is just lost in the noise. So even though this may be effective for small businesses, it would be very hard to measure how effective this is.

I don't know if this is going to work because now I'm moving on to a different kind of advertising that has targeting. This is video ads. This is a video, but I have a feeling it's not embedded in this, which is a shame, because they wanted to show you that video.

(You know what I'm going to do something unusual I'm going to depart from protocol. Hold on. I'm not that old, but I'm getting there. So, bear with me for a second.
OK, you won't see this very well. Obviously but I do want you to hear it.

(Video captions) "Donald Trump is urging all Americans to receive the COVID 19 vaccine'. [Trump] 'I would, I would recommend it. I would recommend it to a lot of people who don't want to get it, and a lot of those people voted for me frankly.'

[Steve Tadelis] It actually continues a little bit more, and the reason I wanted you to hear it was because, in a second, I will show you the results and exactly what this experiment was. But you noticed, in the video you can't see, you see this news anchor, you hear Trump in the background as part of this interview that he had, and of course, all of that dramatic music.

This is what's called creative work in advertising, we had a professional videographer, so I will show you some results from the ad, of course, I won't be able to tell you the success was because of targeting, maybe the success is because of how good the creative is. I don't know how to do that analysis, so I just want to be frank about that. So, what did we do in this paper? And by the way, I just put the paper there, it was published this past July in Science Advances. For those who are not familiar with that part of the world, what AJ is to AR science advances is to science as a way to think about it. The title of this article counter stereotypical messaging and partisan cues, this comes from an idea in political science, people from my generation might be familiar with the term 'it takes a Nixon to go to China.' Meaning if Biden would turn to people who are, say Republican and say hey, you could /should get the vaccine, and say, 'no, we know that you're corrupt and you're on that side.' But when someone like Trump says you should take the vaccines 'like oh, that's surprising. We never thought he'd say that maybe there's something about these vaccines that we should take seriously.'

So that's the idea of a counter stereotypical messenger. And what we did here, is we selected about 2000 counties. This is joint work with Brad Larson that I know many of you know and four political scientists. We selected about 2000 counties that back in the summer of 2021 had relatively low vaccination rates, so this was several months after the vaccines were introduced, so counties with less than 50% vaccination rates, less than 1,000,000 residents, meaning rural and definitely leaning Republican. Just to give you an idea, the median county of these 2000 counties just by this election had about a 70% Trump vote share in 2016. Then, we basically split them randomly 50/50, half of them were eligible to see this ad on any YouTube channel, and the other half were not, we spent close to $100,000 on these YouTube ads, and that gave us close to 12 million ads, costing about 850 per melee (CPN) per one thousand impressions. About 6 million unique viewers, some of them saw it more than once, and it played on what's called mini YouTube channels. So, you could go to the YouTube Fox channel, you could go to the YouTube MSNBC channel, and there are hundreds of thousands of YouTube channels that these were potentially shown on.

Our primary targeting -again, the theme here is what are we getting from targeting- was to target these counties based on the geography but the geography was selected based on the vaccination rates.

Importantly, about 1/3 to halfway through the experiment as we are bleeding money, we suddenly get this realization, that young people are not getting vaccinated that much. They are not seeming to get influenced, so let's cut them out so that our dollars are spent a little bit more effectively. Then we can even better target when we take information like that into account.

Now, what did we find? The average treatment effect was quite strong, we had an increase of about 102 vaccines on average per county. There were a little over 1000 counties treated which means that our just shy of $100,000 generated a little more than 100,000 vaccines, or it was under a dollar a vaccine, which is very cost effective, and much cheaper than some other studies that try to use incentives and lotteries and stuff like that. That targeting -again, I cannot tell you how important the quality of the ad was- it could be that if we just had a simple ad that Trump says 'get a vaccine,' maybe that would have worked. I doubt they would be as effective, so, the creativity part -again, I can't say anything intelligent about that.
One thing that we did want to explore a little bit further to try to better understand how these ads might be working. As opposed to something like paid search where the customer is first giving an intent, such as fishing rods. 'I'm probably interested in fishing rods, but maybe I'm at the beginning of the purchasing funnel or I'm just doing exploration or maybe I know that I want to buy it.' That's that issue with intent. With YouTube, nobody goes to YouTube with the intent to buy or to do something. If they're doing something, they're probably persuading you to do something that you weren't on your way to doing anyway.

One of the things that we wanted to look at is, the treatment intensity. If you are 'treated' more, how do you respond in terms of uptake of getting a vaccine. We have a possible endogeneity here. We can't just replace whether you were treated or not at the county level with the number of ads, because the number of ads may be endogenous to Google's or YouTube's algorithm. They might be showing more ads in areas where they think there might be more uptake. You recall from Nils's presentation that they run a bunch of ML algorithms to make sure that the targeted audience is more responsive. So, we follow standard procedure and used an IV approach to estimate the effect of more treatment intensity on this parameter. We're basically running a regression where we're instrumenting for ads post, which are the effective ads seen in the counties after the implementation of the ads. We're instrumenting that with tree post and that's going to give us the average causal response of seeing more ads.

What I want to point out here, is this relationship with different types of engagement, or different types of treatment intensity. What you are seeing here is the average causal response of a 1 standard deviation increase in several different ways of looking at engagement. In column one, we see engagement rate. That is a fraction of users that watch at least 10 seconds of the ad. Now, I'm sure those of you who have seen stuff on YouTube, which I assume is most of you, you have these running ads. It says skip in 5, 4, 3, 2, 1. I don't know about you, but whenever I'm there, I'm usually looking at the skip not at the ad, waiting with my cursor on it. If people spend 10 seconds, probably they engage with it a little more. And what we see is that, a one standard deviation increase in that kind of engagement increases the rate of vaccines by about 8 to 25 vaccines per county. The view rate is taking things a little bit further. These are fractions of people who watch the whole ad. Of course, I don't know if they started the video, turned around to make a cup of coffee and then came back. But, if they did that, we wouldn't see any effect, the fact here that one standard deviation on this measure gives us an even stronger response of 12 almost 12 1/2 vaccines.

Then, importantly, in that ad, there was a little URL that people could click on that would take them to the Fox News Channel. These would be people that would be watching and say, ‘I don’t think Trump would ever say that. Let me go and check that.’ They would go to the Fox News Channel see the whole segment. These are people that really engaged in the ad and here we see a humongous effect of one standard deviation in that measure. So, it seems that when people are more engaged, they are more impacted, which means that these ads are creating some response that seems to be based on persuasion that the better you target the better you are able to persuade.

So. that was what I played around with in that sandbox of targeting, and now I am going to the sandbox that Nils talked about. I have been fortunate in the past year plus to be playing in that sandbox together with Nils and a few other folks at Meta. The first project that we are playing with -the working paper exists but there's still a lot of work to be done- is related to how these social media ads help businesses. Especially with a focus on small and medium businesses. We have to start by asking -we have this technology there, ‘how well are these businesses using the technology, how well do they understand them?’ As economists we write down models. You may notice that in this talk, that there are no models. You are not going to get any in case you're waiting for one. In our models, we assume profit maximization we assume that companies understand the technologies that they are working in. I love this quote from Chad Syverson in a wonderful survey on this literature that he has been a main contributor with regards to measuring productivity and productivity differences across firms. He says, ‘some producers seem to have figured out their business or at least are on their way while others are woefully lacking.’ When we take this seriously, take our approach that firms optimize and we estimate everything across the 1st order conditions, there is tension here. The question is, can variation in how well firms are using marketing effectiveness be part of the story of how well they are succeeding as a business? Because at the of the end day- Say if Aviv and I are two manufacturers, and we might have some differences in how productive we are, above and beyond the productivity, we now have to sell our stuff. Which means we need to get consumers aware of them, interested in them, possibly even sales people, etc. That will be an important driver in how well we’re able to convert production into sales. That’s one question.

Then, a second question is, will we be able to learn something about the sophistication and learning of these advertisers and how that impacts the success they’ll have with ads? And as I hope to convince you in a second. The answers are yes and yes. Yes, there is a lot of variation in marketing effectiveness that would have an immediate implication on returns on advertising and therefore profitability. Second, there seems to be some interesting patterns with the way these advertisers seem to experiment and engage with the platform and their success.

So let me tell you a little bit about this. You saw most of what’s on the slide, everything to the right of little picture there. What you see here are kind of the four parts of social media advertising, creative, objective, targets, and budget. What do I mean by that? On the very right side, you see the photo of an iPhone with a Facebook ad just sitting there as part of your feed. That’s what a typical ad might look like, here there’s no video it's just a picture, but you can click on it to get information on signing up for a nursing program to learn how to be a nurse.

So that's something an advertiser has to start with. What's the creative that I'm going to show? And again, I'm not going to say much about one creative being better than another. This is for creative people and I'm an economist. But then what you see here on the right in those three boxes, the blue, the black and the green, these are the ways in which advertisers engage with the Meta platform to decide on how to target their ads and how much budget to spend. You saw that in one of Nils's slides. Everything you see here in the green is the targeting, geography, age, etc. Then, I just want to point again to what you see there in the blue at the very bottom where that gray highlighted area. It says sales, so this is for advertisers who are interested in more sales, that's their stated objective. And again, I will mention the pixel in a second and I'm going to be able to do that quickly thanks to Nils's presentation. Then in the black boxes is where they set the budget and everything. Just to talk about the budget and how these ads are allocated. Meta uses a generalized Vickrey Clarke Groves mechanism, so this is actually used in reality. The way it basically works is every ad is going to get a score or a value that is a combination of the bid multiplied by an action rate which -as Nils explained- is how likely does meta believe you are going to engage with this ad and then some quality score that has to do with the ad, and then they take all of these values and basically plug them into a VCG mechanism and create a ranking from 1st to 2nd to 3rd and so on. So that whenever you open your app and you click 'open app,' that auction is run those ads are then placed, and what they pay is exactly their externality value on the VCG.

So that's how the auction works. Just so you know, very few bidders today actually set bids, they usually say 'here's my money. This is what I'm willing to spend, do your magic.' So that's basically how it works. Which already suggests that maybe these people are very sophisticated and they understand everything but they're just a little too busy or it's like 'yeah, no. I don't understand this, but here's what I'm willing to spend.' I don't know about your prior, I have a very strong posterior after spending quite a while in the industry. And then finally, and this is what's important, is this idea of the pixels. We will have the revenue data from these third parties thanks to these pixels, Nils mentioned this, but I will explain it more in a second.

So again, the idea of this experiment is exactly the type of experiment that Nils leveraged. Here we didn't have that extra step and that's what's really cool about Nils's paper is they had a bunch of engineers design a different algorithm based on clicks instead of sales and they implemented it. Here we're taking the standard we're doing incrementality test. And again, that's what Nils did as part of their analysis. Here's how to think about it, imagine an advertiser has a daily budget of about 100 bucks, and their targeting means there are 100,000 users that are eligible for this particular ad. What would happen every time one of these users open the app, there's going to be an auction and this advertiser will be part of the bidding. They will show up wherever they show up and everything is done.

Now what we then did is take \$5 from this 100, 5% of the budget let's take 5000 people, 5% of the target audience and now this is our experiment. We are going to take 90% of those 5000 people and allocate 100% of that \$5 budget to them and then we have 500 people that are holdout. They will not see the ads.

For example, if I am the user who is in the hold out, what does that mean? I click on my app, it opens, the auction runs. Let's say that Mike's company is 1 in which I am the holdout, so, the auction runs, Mike is #2, he would be the second ad in my feed but because I'm in holdout, we take that out. We move everyone up by 1. The holdout group is the true counterfactual of what would happen if Steve chooses not to advertise on Facebook anymore. That's the idea, hence the incrementality test of advertising versus not advertising, and what does the pixel do? Every time someone goes to my website and purchases something, that data goes back to Facebook. And it says this person, meaning this cookie purchased this much money, right now here's the data. Now when it goes back to Facebook, Meta, you look at that and you say, 'Okay, who are you that made this purchase?' You were either treatment or control or I don't know who you are. All of that is being measured -the I don't know who you are means I can't track the cookie back- but I could know if you're in control or treatment, and then the idea would be to compare the purchases by control versus treatment to get the incrementality of the ads. What this then means is that we're going to have a lower bound on measuring the returns to ads precisely because compliance is perfect on holdout, but not on treated. On holdout you're not going to see my ad. On treated, you see my ad because it's #3. But I actually don't know that you reached it, because you may have opened the app, I won the auction, and now you got a call from your kid, 'I need you to pick me up,' you shut down the ad, you didn't actually see the ad. But it was registered as if you were in treatment and that you had gotten the ad.

Very quickly, huge sample size, we had about 210,000 advertisers, about 700,000 ad campaigns, about 4 billion user ad pairs. What we did is collect data on a whole bunch of conversions, but were really focusing on purchases, and using that, this is just the benchmark result for every dollar spent on average. So, you're not seeing that distribution that Nils showed you, but on average, you are getting $3.31 in incremental revenue. Now, this is incremental revenue not incremental profits, so you would need a profit margin of just around 30% for this to be profitable and this has to be a profit margin over variable costs. In retail, that would actually be a relatively low profit margin. So, on average it seems that these ads are working quite well.

But what's more interesting is to actually see that there's a relationship between the ad spend that you had historically, or we've also done this with how old you are as an advertiser, and there seems to be a cross experience. So, what you see here is the five quintiles of historically least to most ads spend. The picture looks very similar historically. How long have you been advertising on Meta? What this suggests is, that advertisers who have more experience are getting a higher return on investment. Of course, that could be pure selection. Historically, her ads worked really well, mine didn't come. So, she's spending a lot of money now and getting a lot of returns and I'm hardly doing anything, because it's not worth it to me. Those are the two extremes, so it might not be causal in any way that experiences driving returns. So, what we then did, is pull an idea out of Arrow's famous 'learning by doing' paper, so quoting from that paper, 'learning can only take place through the attempt to solve a problem and therefore, only takes place during activity. The idea that we had here is, I could launch my ad, but then if all I did is launch once and do nothing. Well if you are a hardcore Chicago economist, you think 'well, of course. That's what you would do because you are optimizing, and your first ad must be your optimal ad.' I teach at Berkeley.

The idea is, if you just launch and forget you're probably not very sophisticated and you're not trying to figure out how things are working. Versus someone who comes and tweaks the budget, or tweaks the target, or maybe even tweaks the creative. So, imagine, I'm going to look at people who never make changes versus people who make some changes to their targeting and let's see about this pattern of experience and returns to advertising.

What you see on the left panel over here are the people who never update their campaigns, and you don't see any relationship between experience -measured by age or measured even by spending, it doesn't matter what we look at- and returns. Whereas for those who do make these changes and these updates, we do see a strong relationship between experience and returns that is monotonic, that is quite suggestive that people are learning. Those who don't learn, don't succeed. On that note quoting an even more famous scholar, 'the more you read, the more things you will know, the more you learn the more places you'll go.' Doctor Seuss 1978. So, what we are looking at here is we can see what kind of data these people are engaging with. I mentioned this idea of pixels. Well, revenue is coming from the sales pixel, but some advertisers choose to measure add to cart and register and visit this page or that page. They are learning more about consumer behavior. Some advertisers are using higher quality data -this CAPI integration- and one thing that is interesting, we see a strong correlation between people we call more data sophisticated and people who are learning. They seem to be associated. So, for those in academia, think about your grad students. Some of them are more diligent, they try more, are more creative. Others have less of those correlated skills. What we find here is that the monotonicity is really strong in those who exhibit data sophistication, and pretty much nonexistent for those who don't.

The last thing that I want to talk about is touching exactly on the topic that Nils presented. This is the idea of privacy considerations. So, when testing ad effectiveness -as Nils explained- does is it requires you to know who's who. Who's treated, who's not treated. In order to do that with that offsite data, you need what's called these third-party app data sharing. There's something called the IDFA, the identifier for advertisers, this is a random ID on Apple. There's something similar to IDFA -the Google advertising ID for Android. What Apple did in 2021 was implement this ATT -this app tracking transparency and that's what you see here on the right. When you download Facebook; do you want to allow Facebook to track your activity across other companies? And notice the convenient talk default is 'no' and once you click on 'no' that takes away the ability to track those sales that we talked about, and that Nils talked about earlier. So, what will ATT do for advertisers? What Nils already showed is that advertising effectiveness is going to go down. If advertising effectiveness goes down, what does that mean about the advertisers' behavior? What we did here -and this is a work in progress. Well, the previous thing was a work in progress too, but at least there's a working paper here. You won't even find that yet- so if you look at the top right graph, you will see on the left, the US, and on the one before the last on the right, you see Germany (DE). This is the opt out rates on Facebook and Instagram as of August 2023. You see the opt out rates are much higher in the US than they are in Germany. You might say, 'well, why is that?' Well, the answer is simple. In the US, 60% of people, roughly, have iPhones and in Germany only about 30% of people have iPhones. If you're an Android user, this is irrelevant for you as of yet. Google is planning to release this. They say in 2024.

Then we took this observation that this is randomly because of the number of people who have iPhones versus Androids, and within Facebook -I'm out of time, so I'm going to steal another two minutes- the advertisers are classified into industries. What you see here in the pink versus the green, we can look at industries that have more iPhone and therefore are going to be more impacted, that's what you see on the left in pink, so, verticals like retail, like consumer-packaged goods, they have about a 60% opt out rate. On the right, you see the verticals like gaming, energy, natural resources and utilities, they only have a 41.44% opt out rate. We are going to take advantage of that difference to see how spending by advertisers on Facebook, across these impacted industries look like. And this is the diff-in-diff from that simple analysis, is that industries -the five most impacted industries- have roughly a 43% higher exit rate of advertisers from Facebook, compared to the least impacted industries.

Now there were some questions, it could be 0 sum, or maybe they're just going to Google. If you believe they're doing something optimally and now they have to respond, well clearly, they are negatively impacted. The question is, are there stronger pieces for that? And this is very much hot off the press, we looked at data from the census, the business formation statistics. Looking at new business registrations, in every one of these quarters pre and post this ATT by the diff-in-diff with the industry, so similar to what we did on the spending on Meta, here is new business registration. It turns out there's an almost 15% drop in registration in those industries that seem to be more impacted. With some preliminary analysis, a decrease in employment by 3%, decrease in wages by about 2%, decrease in total output of about 6%. And now we are getting data on exit rates to see if we could establish this on exit rates, which would be a pretty important metric.

To conclude, ads seem to work well when they are targeted well, and targeting does require savviness and measurement by the advertiser. So, it's not only gathering the information, but presenting it to advertisers in ways that they can use that some of these platforms are doing. And now we have all of this talk on privacy concerns. I know that they are real because my wife uses DuckDuckGo. I know one other person that uses DuckDuckGo. So, some people really care about it. But how do we measure those returns to privacy? I don't know because we know we have the privacy paradox, that most people, you give them a Snickers bar and they'll give you their social security number, their mother's maiden name and all of that. I'm not convinced that the benefits are high and that we are seeing more and more evidence that they are imposing significant costs, and we have so much more to learn about this. Here I just pulled this from a recent Fortune magazine article, 'What the well-meaning critics of online advertising are missing and how they could hurt the communities they are trying to protect.' This was written by a trans owner and founder. She was one of the founders of the Huffington post I believe. I'm talking about how all of these very niche businesses that are trying to target different groups, they really need that targeting. And these are typically small businesses that a lot of people on the progressive left are saying 'we need to support this.' Well, you take away tracking and these same people are saying 'privacy, privacy, privacy,' you would be really hurting these businesses.

And with that, I am done.

[Applause]

Is there a plan for Q&A? But I won't be offended if there aren't any questions.

[Eddie]I think we have a couple minutes for Q&A.

[Q&A]
[Q: Dave Benson] Hi, Dave Benson, Federal Reserve Board. Is it possible that the tradeoff is less about the value of privacy, and more about the use of the information by the businesses. For example, to price discriminate? I don't know if there's any evidence in your data on that.

[A: Steve Tadelis] I don't have evidence on that. As we know price discrimination could be pro or anti-competitive depending on whether it creates value or just takes- It rarely reduces total surplus, but it could sometimes reduce consumer surplus, and sometimes it could increase it. I haven't seen any robust evidence on that. So, it's a fair concern.

[Q: Matt Leisten] Matt Leisten from the FTC. So far today we have seen a talk saying that data and the ability to process that data are complements in production. We have also seen studies saying that maybe the smallest, least sophisticated firms have the most to gain from data, or at least the most to lose from having it withheld. I wonder if you have some sort of grand theory of being able to hold those two things together in our head at once, how do those two things comport with each other? Do you have anything to say about that?

[A: Steve Tadelis] So, that data and computer complements make sense to me. You know I spent a year playing in the AWS sandbox when I worked part time at Amazon. I am very convinced that it's the small and medium business that benefit the most from being sophisticated. At the same time, they are the least sophisticated. This is where the platform kind of analytics come to play by creating these dashboards and trying to basically help these advertisers get more from their investment.

This is something that I hope to pursue as part of that project that was the penultimate one that we looked at that learning and sophistication to see if there are ways that you know maybe future research, we could play with, how much information we are pushing to these advertisers to improve their performance. I know that not everybody will agree with this potential theory. I think that the platforms that are focused on the long term. Let's take Amazon, they always talk about how it's all about the consumer, long term, because that's how we gain our long-term profitability is by making people happy. I think that a lot of these advertising platforms if the advertisers are not going to be happy in the long run, they are going to pull their advertising dollars, so I think that the incentives of the platform and of the advertisers, are not misaligned. Therefore, I think that the platforms can help these small and medium businesses through these kind of dashboards or even more paternalistic ways of taking over these –'we think you should do more of this and we are going to do it unless you don't want us to'- type of approaches. I hope that answers your question a bit.

[Q] So on that last point, something that, such as when you look at the Facebook platform for advertisers, they let you choose to maximize sales, but not incremental sales. Right? So, one tool that they can give advertisers is, maximize what you really need to maximize instead of maximizing the wrong thing. Any musings on that?

[Steve Tadelis] In the results that Nils showed, that I showed, everything is very statistically significant because I'm aggregating experiments from 700,000 campaigns. Take any one campaign on its own, the signal to noise ratio is going to make it very difficult to say things with statistical confidence. So, the question then is, could you really measure this at the advertiser level? In a consistent way that would take incrementality into account above and beyond just total sales? I think that's a high order. I don't know how- You seem skeptical. No, they do learn. That's something that we just started playing around with. Who is learning? How much data do they have to learn? We haven't yet explored that. It could be that all they're doing is looking at point estimates. If point estimates are in the right direction, then that's fine. Then we can help them do that, but that means that we would need to give them the kind of incrementality tools that are quite sophisticated. So, will they know how to do it? I'm not sure, it's an open question, but I think it's a great one too. Okay. I think we are done.
[Applause]
[End of Q&A]

[Aviv Nevo] Thanks, everyone for a great session this morning we are going to break now for lunch, we have sandwiches outside, and we need to reconvene at 1:00 PM, I just want to ask everybody to hang on to their badges or their name tags, even for tomorrow too, because that's going to be an easy way to get back in. If you do leave the building, just be aware that you have to pass through security again. So, enjoy lunch.

[Viola Chen] Welcome back, I hope you are all well fed and ready to go for the next session. My name is Viola Chen. I am a staff economist here at the FTC and one of the co-organizers. I wanted to take a few moments to reiterate our thanks to all of the folks who make this conference possible. It takes a lot of people doing this behind the scenes. I wanted to thank all the folks in BE who helped us out. As well as our two points of contact from the FTC event planners, Panar Gezgech and Annabel Bendz. I also want to call out Bruce Jennings, who was the main point of contact for our production team. Specifically, I want to call out Stephanie Aaron, who is here in the back waving her arm, she is the one writing all of the emails to all of you. So, I am sure all of you have seen her name on emails. She is the true superhero who helped to put this conference together. She had an amazing attention to detail with all aspects of this conference from the call to papers to checking all of the slides. She put them all together, so thank you Stephanie.

With that said, let's get back to more research. The 1st paper in this next session is going to be presented by Michael Lipsitz, a staff economist here at the FTC, who is going to be presenting the paper Innovation in the Enforceability of Non-compete Agreements. It's co-authored with Matthew Johnson and Allison Pei of Duke University. It will be discussed by Liyan Shi from the Tepper School of Business at Carnegie Mellon.
So, with that, I give you Michael Lipsitz.

[Michael Lipsitz] All right, hello everyone, it is really wonderful to be here and to be presenting for all of you. As Viola said, I'm going to be presenting on Innovation and the Enforceability of Noncompete Agreements today. Before I jump into any background, I just want to make sure you all know what noncompete agreements are because obviously, I will be talking quite a bit about them. A noncompete agreement -and in particular labor noncompete agreement- which is what I'll be discussing, is when a worker signs a contract with their employer that says they cannot go work for a competitor after they leave. Sometimes it is constrained by time, like for a two-year period. Sometimes it is constrained by geography or other things like that.

With that said and with that in mind, there is a bit of tension in terms of labor and mobility and innovation. That tension is basically that labor mobility tends to facilitate innovation. Someone brought this up the other day when I was giving a talk on this paper. It's like the water cooler effect where people bump into each other and if you were seeing new people all of the time, that can facilitate innovation. On the other hand, that's not necessarily always a great thing for firms because if their workers are leaving and they're bringing some of the IP that they have to their competitors that leakage can be problematic for them. So, what do the firms do? Sometimes they will use noncompete agreements and that way they can hold on to the IP and keep it in-house. About 20% of workers across the workforce have noncompete, at least as of 2015. And something like 35% have them in tech industry. So, it is not surprisingly pretty prevalent there where IP is important.

The question I want to discuss today is, how does the legal enforceability of noncompete agreements impact innovation? If firms use noncompetes to protect their IP, and if that prevents workers from going and creating this cross pollination of ideas, what is the overall impact on innovation? This has been discussed in the literature to some extent. Even going back as far as 1994 and before, there's this idea that Silicon Valley rose to prominence because there is no enforceability of noncompetes in California. You can't use noncompetes in California. Therefore, in theory, inventors can run all over the place and share their ideas, and that sparked these tech clusters. Contrast that with Route 128 outside of Boston, which used to be a tech hub and has fallen off in comparison to Silicon Valley. Noncompete enforceability has been posited as one reason for the discrepancy. The flip side of this is that there is a holdup problem. If a firm wants to invest in their worker, such as spending R&D dollars, maybe training them, the existence of a noncompete can actually enhance that investment incentive. The firm will be more willing to do so when they know that IP will stay in-house. In theory, that can also increase innovation. So ultimately what we have is an empirical question about which of these effects dominates and that's what we will discuss today.

There is obviously policy relevance to this. The FTC has proposed a rule in January to ban noncompete clauses, so understanding what the impact is going to be on innovation is an important component to understanding what the impact of that rule will be. While this has been discussed in the past to some extent -I mentioned the paper in 1994 which was more qualitative- there is very little direct comprehensive evidence on this question that really covers a large set as externally valid. In particular, the mechanisms that underlie it. What exactly am I going to do today? I'm going to discuss data on state level noncompete law changes that form the independent variable in this project and we are going to measure innovation using patent quantity and quality, a whole slew of measures since it is somewhat difficult to measure innovation. And what we ultimately find is that making noncompetes easier to enforce -so I will call that higher enforceability as the shorthand- when non competes are easier to enforce for firms, that leads to a large and persistent drop in state level patenting. We interpret this, along with robustness checks, as evidence that when noncompetes are more easily enforced in courts and when firms have more of an incentive to use them, that harms innovation. In other words, this inventor cross-pollination theory may dominate. We will look at that more granularly as we go.

Patents don't necessarily reflect innovation, and we take that idea really seriously here. We do some checks to make sure that these results are not driven by useless or strategic patents, and also that noncompete laws aren't just reallocating inventor talent across state lines. It could be that either firms or workers are going to different places to do their innovative stuff, and that's actually causing a reallocation. So, when we look at state level analyses that could be problematic, but I will show you that it actually isn't.

Then finally, I will turn back to this theoretical tension and try to reconcile those contrasting theoretical predictions, by thinking about worker mobility, entrepreneurship, and job creation. How those things are affected by noncompete enforceability and also, looking where we can, at publicly traded firms and showing that the holdup problem is real. There is sort of an action out there along that dimension, that firms do increase investment to some extent when noncompetes are more enforceable -- and yet nevertheless, we still see lower rates of innovation.

I want to talk a little about the data, since it's quite important here. Noncompete enforceability, so how likely a court is to uphold a noncompete and enforce it. It is set by state law, broadly speaking. About 10-15 years ago, legal experts started to think about how to quantify this and there are roughly 7 dimensions of enforceability that they have identified as being the most important. They did it in 1991 and 2009. What we did in a prior project is we extended this database to every year between 1991 and 2014 and found every single law changed that happened. What that resulted in, was 73 law changes over that sample period. 91% of these, the vast majority, were because of judicial decisions. So, a case went to court, the judge made a decision and that affected precedent. Then the remaining ones were through statutory changes, which were relatively uncommon. The mean law change- Just to give you a sense of this, and we are scaling it between 0 and 1, so 0 is noncompete which absolutely will not be enforced. One is the most generous enforceability that we see in our sample, which turns out to be Florida. That doesn't mean that all noncompetes get enforced, it means that a noncompete that is reasonable and meets certain guidelines will be enforced, so a mean law change shifts the noncompete score by about point .08. We're not talking about spanning that entire space from 0 to 1. You can see that a lot of the standard deviation in the sample, though not the majority of it, is due to within state standard deviation. I bring that up because that's the identifying variation that I will use here to look at the impact of noncompete law changes.

Turning over to the dependent variable side, we are using patent data here. So, we are going to use the universe of patent applications -these are utility patents. We have the data from 1976 to the present, we are mostly using that sample window that I discussed and then we are going to assign patents to states based on the inventor's residential addresses. The operative location is the state where the worker is working, and that's what law will govern their contract. So, we are matching to where those inventors are. Of course, there are complications with that and we have to split patents across state lines when there are multiple inventors, but broadly what we're doing is matching on those residential addresses. We can also assign each patent to a technology class, which has been done previously. CPC is the measure we are using of that. So, you will hear me refer to it a couple of times. The number one primary measure I will use, is the patent count weighted by forward citations. A forward citation is not how many papers you cite in your works cited, but rather the number next to the paper in Google Scholar. It's how many times your patent has been cited. We also do truncation to make sure there are no mechanical issues or differences across fields, across time to make sure that these are comparable -which are pretty standard things in the literature. The idea is that these forward citations will reflect the patent's quality or significance. It is not definitely true that this is the absolute best measure or that it reflects all of the variation in true innovation associated with a patent. I will also show you some figures with raw patent counts as an alternative measure.

Moving on to the counter metric -the estimation strategy- we have several problems. Noncompete law changes happen in different states, in different years, it's a continuous measure, it's not binary. They can go up or down. Sometimes they go up and then down. We have all sorts of problems that all of you empirical folks out there have read about in the literature in diff-in-diff over the past few years, and it makes doing diff-in-diff in this context very difficult. So, we face a tradeoff. What we can do is use the universe of changes and expose ourselves to all these econometric issues or we can try to walk away from these econometric issues by using more of a simplified design that extracts some of the variation that we have in our sample, but not all of it and uses that variation in a way that is much more econometrically sound.

We implement a stacked design. A modern innovation in the past five years that has been used in the minimum wage literature and now there is recent econometric literature about its econometric properties. Basically what we are doing, is creating these sub-experiments which is defined by one treatment. A noncompete law change and we have parameters around when we count something as a treatment or not. Basically, there shouldn't be a countervailing law change happening right after it. Then we find some controls and the controls we use will be these really clean controls which are states that don't have law changes over our sample period. There is some sacrifice. We're sacrificing some of the variation in our sample in order to get, what we consider, a much cleaner sample.

Here's the specification. It ultimately ends up looking a lot like a two-way fixed effect, but you'll notice that B floating around as a subscript. That is the block or the sub-experiment. That interaction with the sub-experiment is what turns this into a stacked design. We create these sub-experiments, aggregate them in the data and then run this with that interaction with the block. The sub-experiment. We will also use Poisson regression to account for the fact that these are count based dependent variables. I'll also note that in some specifications we will do this at the CPC level. We will look at this by the different technology classes, as opposed to aggregate it to the state level. One last issue here, is that when California and

Washington had a state law change, they outpace all other states by a long shot or at least our control states in terms of their innovation. So, if we are worried about non-parallel trends, you can't have California as a treatment state in the sample. And Washington, who had a law change around the .com boom. We admit these to be conservative. Our estimates are way bigger when we include them, but we will also get a pre trend. So, that's why that is. We take them out to make sure that we have an in-support control group.

Turning to some results on the effects of enforceability on patenting. This our main specification here. What you are looking at is these normalized forward citation weighted patent counts. This one is at the state by CPC level. What you see is that, when those law changes occur, we get a persistent decline in the rate of patenting. Or at least the rate of forward citation weighted patenting and we interpret this. These law changes -this is normalized to a positive law change- so what this means is that, as enforceability increases, as noncompetes become easier to enforce by the court, what we see is a big persistent drop in innovation. Again, supporting the water cooler theory -this inventor interaction theory. You can see that the impact of a mean score change in our sample is almost 20%. This is a pretty sizable effect. At least in the specification. In some of these other verifications, the numbers in the top right corner might be small to read. They're a bit smaller, but all are in the double digits. This is just a couple of different specifications to show that the broad pattern tends to hold across those different specifications, whether we are looking at raw patent counts. Whether we look at the state, or CPC level, it's the same kind of broad pattern.

To interpret these results numerically, the average increase of the noncompete enforceability score in our sample, which is about .08, reduces state level citation weighted patenting within technology classes by about 18%. If we look at the same thing and we don't weight our patents by importance, we get 11%. Some weak evidence there that the quality matters in addition to quantity. Just to give a benchmark - because when I first estimated the numbers, they looked a little bit big. So, I wanted to see in the innovation literature, what other changes -legal changes or different things like that- cause a decrease in patenting of about the same magnitude. For example, if an inventor moves from a tech cluster that is at medium sized to the 75th percentile, that gives a comparable magnitude. A 10% decrease in the tax price of R&D. A 1% standard increase in the firm's exposure to Chinese imports. We are not talking about crazy changes in these three things, but they are meaningful. So, we interpret that as a sense that our estimates are reasonable, but meaningful.

I want to now turn to a 'kicking the tires thing'. The issue is that, it is possible that patenting -even the citation weight patenting which in theory helps us get more towards the quality of innovation as opposed to just the number of patents- we want to make sure that this reflects a true loss in innovation. There are a couple of things that could preclude that. Number one is that, if the averted patents -the patents that we lose due to increase in noncompete enforceability- are either useless or strategic, there are a lot of games that firms can play surrounding patents. And we want to make sure that we are not just picking up changes in how those games are played. It could also be that these are for ideas that eventually get discovered in other states. If a firm decides 'hey. Noncompete enforceability shot up in New York, we're going to reallocate our R&D arm to New York. We would have done that same innovation in Texas anyway, but we will reallocate it to New York.' That's a reallocation. At a national level, that's not going to lead to an impact on innovation. It's just going to lead to a reallocation but will ultimately be zero sum. We want to investigate that and see which of those two things is true.

Turning to this first issue about strategic patents, what we will do is use some metrics that have been developed in the last several years to try and 'kick the tires' on this and see if there is truly a quality effect or if it is just a quantity effect. The first thing we will look at is whether the number of citations that a patent ultimately receives are in sort of one of the top percentiles inside their cohort, which is defined by a technology class and year. The idea here is that there may be a long tail of patents that are not really that important for true innovation, or the type of innovation that sparks economic growth. It may be just those really important patents at the top that are important for true innovation. We will also use a recently developed textual measure, which basically compares the texts of patents to previously filed patents, and it assesses whether or not it is a breakthrough. It asks if the patent is breaking new ground based on the textual similarity embedded in those patents. Those are the two measures we will use.

Then we will do one more thing, which is to guard against the idea that noncompetes in patents might be substitutable ways to protect ideas. So, if I am a firm and I know that there is some potential for my IP to walk out the door, maybe I have two different options. I can use a noncompete and keep it as a trade secret to make sure that the inventor doesn't walk out the door -or whoever else knows about it. Or I could patent it, with the understanding that that innovation becomes public knowledge, but I have protection over it for a few years. In some ways those are substitutes and we want to check to make sure that whatever we are picking up in the state level analyses, is just the substitution pattern. So, what we are going to do, is look at a sector where reverse engineering is ubiquitous and the ability to reverse engineer is so strong that patenting is very, very common. The idea is, that you cannot use trade secret protection and noncompetes, if people can figure out what you are doing just from reverse engineering. You have to use a patent to lock up that innovation.

We will look at those sectors to see if we find similar effects. I will present all of these to you all at once. -I should really put the primary specification measure on here as well. You see that we do get wide confidence intervals on some of these measures, but by and large, what we are seeing here is a pattern that's very comparable to what we're getting in our main estimate. We get impacts on the really high-quality patents and not just the useless, or strategic, patents where we see the action. We get impacts on these breakthrough patents, which is the textual similarity metric. We also get impacts on non-breakthrough patents. The impact on breakthrough patents is bigger, maybe not significantly so. But either way there is an impact on both types of patents. Finally, while not statistically significant, we find an impact in the pharmaceutical and medical equipment field as well.

Moving on to that second question, the second 'kicking the tires' question, we want to know if noncompete enforceability truly reduces innovation or if it reallocates it across state lines. So, turning back to this 128 for Silicon Valley idea. There are stories about entrepreneurs who decided that they will leave Massachusetts and pick up and go to California. If that's the case, maybe they would have discovered the same things in Massachusetts, but because of noncompete enforceability they decided to take those innovations to California instead. But at a national level, at an economy wide level, the impacts are going to be identical regardless of where they are coming up with that innovation. This is analogous to the migration response to taxation where you move to the places where taxes are most friendly to your firm. If reallocation is perfect, we expect little to no effect on aggregate innovation. What we are going to do to test for this is we are going to flip the design on its head and do what we think of as an exposure design. The basic idea here, is that we are going to imagine that there is of sort of baseline spread of technology classes, in some given baseline year. And that is exogenous to subsequent changes in noncompete enforceability. What we are going to do is take that baseline spread and look at all of these changes that occur in different states and we are going to calculate the effective change in noncompete enforceability for given technology classes. Then we will see if the technology classes where the exposure to noncompete law changes has been the most positive. Meaning the biggest increases versus the most negative meaning the noncompetes become more difficult to enforce. We will see if there's a difference in innovation in those different technology classes.

So, the graphs presented for unweighted patent counts as well as the citation weighted patent count. What you see is exactly a story that tells us that there is a true reduction in innovation, that this is not just due to reallocation across state lines. What we see is at the higher end of the X axes, where we are thinking about technology classes exposed to increases in enforceability, it is easier to enforce noncompete. That is where we see lower percent changes in patent counts or citation weighted patent counts. Consistent with the story that there is a true reduction going on here, not just a reallocation across state lines.

Last thing I want to do to is to go back to the Gilson story of the 128 versus Silicon Valley versus a Grossman Hart Williamson story about a hold up problem and the solution the to hold up problem potentially increasing investment by firms. What we want to do here, is to think about -and I kind of jumped the gun a bit- but there are other reasons we might think that 128 vs. Silicon Valley, that kind of argument will hold weight here. Could be that there is reduced entrepreneurship because the entrepreneurs are worried about innovating for whatever reason. It could be that noncompetes make it harder for startups to grow. And again, the job-hopping idea that I motivated it all with. On the other hand, we have the holdup problem that a noncompete can incentivize firm investment because they know that the IP or their dollars are not going to walk out the door to a competitor. Our baseline estimates suggest that this kind of Silicon Valley versus 128 story dominates. That doesn't necessarily mean that there is no investment hold up problem. It just means that empirically, in the reduced form, what we are seeing is a bigger impact of the water cooler Silicon Valley story rather than the solution to this hold up problem. We're going to dig in a little bit more here and investigate the constituent components that I am discussing.

A couple of things to look at. Number one, job mobility. Not going to spend a lot of time on that because it has been covered quite a bit in the noncompete literature. Not to mention that if you artificially lock up employees it's not surprising to see reductions in job mobility. But I will take a look at that. We're also going to look at entrepreneurship. Similarly, if you write a contract with an employee that says you cannot be an entrepreneur in this field, it's not surprising that we see reductions in entrepreneurship. But I'll show you that as well. We will also look at job creation through the same lens, which is maybe a little bit trickier to understand because there could be competing impacts. Another thing we did here is we took the CrunchBase database, which is sort of this website looking at startups and things like that, and we matched them to the patents so that we can tell which patents are coming from startups. We can look at what's going on with startup patents, independent of the rest of the universe of patents. For all of these outcomes, I will just point out –since these are population-wide outcomes- we're going to focus on innovative industries. We use an NSF definition for that.

Here's the mobility result. I won't go into details about why we're presenting all these, in the interest of time, but there are technical issues with it. By and large, what we see is that there are reductions in mobility where noncompetes are more enforceable.

Turning to establishment entry rates and job creation rates. We see patterns consistent with what you might expect there. We see decreases in new establishment entry and decreases in job creation, associated with new establishment entry. Then column 7 and 8 show us patenting differences for startups versus other firms. And we do see that there's a bigger effect for startups versus other firms.
Finally, I want to turn back to this holdup problem. The other side of this trade off. For publicly traded firms, we can actually look at investment data. This is because some Duke researchers have matched patent data to Compustat data. So, we can use the Compustat investment metrics in order to test whether firms subject to different enforceability regimes are patenting at different rates and are investing as well. These 1st two columns look at investments. And column one is the R&D investment, intangible investments, and what we find is evidence consistent with the holdup story. We find that firms are indeed investing more when they are able to use more enforceable noncompetes. We don't find a similar response to capital investment, but in intangible investment we do. Turning to columns 3, 4, and 5, we find that there is still a large reduction in patenting even among these large firms.

This tells us that that results we saw with startups, it's not just the startups that are driving the negative results that we see. I will caution that because there are a lot of multi state firms, we need to use a different specification. It's not ideal, to be quite honest. I wish there was a different way to do it, but we cannot omit states the way we do in the main specification. So, these numbers might not be quite apples to apples with the main ones. But we nevertheless see these large reductions in patent count and citation-weighted patents. And then the final metric, which is actually something we can add here, which is interesting, which is the value of patents. That value is estimated using effectively abnormal returns upon announcement of a patent that tells us what the monetary value of that patent is. And that is a metric that we imported from some other researchers.

So, I will wrap up there, and quickly summarize since I have a couple of extra minutes here. Making noncompetes more easily enforceable leads to a large and persistent drop in patenting and we would like to think that this represents a true loss in innovation. It's not just a reallocation or just getting rid of some useless or strategic patents. Enforceable noncompetes do appear to increase firm investment in intangibles, but it's these externalities, these spillovers across firms, this water cooler/Silicon Valley story that is changing the results in terms of the number of patents due to labor market dynamism. Taking a

step back and connecting this to the literature more broadly, there are findings out there that suggest that there is declining labor market fluidity over the past several years. There's also results out there that show that there is declining inventor productivity. Our paper shows that there is a linkage between those two things. That's not necessarily to say that noncompetes are the only or the major driver of that linkage, but it does say that there is a link between declining labor market fluidity and declining inventor productivity. So, with that. Thank you very much.
[Applause]

[Viola Chen] And next we will have Liyan Shi to discuss.

[Liyan Shi] It's a real pleasure to discuss this paper. I think that there's a really important takeaway from the paper. The punchline is that noncompete hurts innovation, and there is a lot of anecdotal evidence that we have had before the paper. Some small-scale evidence supporting these results. I think this paper provides a large-scale comprehensive evidence and really compelling evidence that shows that noncompete indeed does hurt innovation. Not only do we see a decline in quantity, but also quality, regardless of how we measure it. Either citation-weighted, text-based novelty or how investors invest the value from stock market reactions. This does reflect a true decline in innovation, ruling out any potential reallocation across the states.

So, when we look at the firm level, at which type of firms have been hurt by the enforcement of noncompetes? It is intuitive that noncompetes create additional barriers to entry, that we see startups decline. Also less patents being created by startup firms, but it also seems to some extent that the incumbents have been hurt as well. We see the non-startup patents declining also. Some would say that the evidence is not very significant. But when we look at publicly traded firms, and here we can think of these firms as being large incumbents, and we would think they might gain a little bit by having extra protections on their investment. But it seems that their innovation also declines by all measures, despite putting in more investment in dollar terms. So, it seems that everybody is losing on this and there is no identifiable winners that we can see. That nobody is gaining from noncompete enforcement.
In my discussion, I will piece together the evidence and offer my interpretation of these facts and see whether we can more concretely tease out the channels of how noncompete can hurt innovation. I'm going to focus on three aspects. Patented versus non-patented investments and the efficiency for the production of innovation and incentives. I think these channels were mentioned in Mike's presentation, but I am going to offer my own interpretation on how we can potentially give extra support for each individual channel.

Perhaps the most striking contrast of these facts is that we see incumbent firms investing a lot more, and that seems to address the holdup problem. This is consistent with previous findings in the literature, including my own work and others. Intangibles really is a very broad concept when we think about it. It lumps together many non-physical investments. We can break it down according to R&D expenses and SG&A expenses. So it is really curious that we wanted to know what the breakdown of the two types of expenditures. I would think that R&D expenses would be more directly linked to patenting outcomes. Let me make a bold assumption. Let's say that noncompete is not affecting the returns and efficiency of R&D.

If we put in more dollars, it pumps out a certain amount of patents. If that is the case, we must see a decline in R&D, which causes a decline in inventions. If that is the case, perhaps there is a compositional change that we are seeing more investments in other types intangibles. So, it more broadly relates to some of the issues that I'm curious about when it comes to noncompete versus patents. It could be that noncompetes are protecting non-patentable knowledge and investments. It is possible that -I think Michael also mentioned this issue- it could potentially be a substitution effect. We could alternatively think about a complimentary effect. I think that by breaking down these two types of expenditures, maybe we can get a better sense of the input and output on innovation.

Let me come to my second point in terms of the production of innovation. For now, let's drop out the assumption that I made earlier. It is a very reasonable hypothesis that noncompete prevents movement of workers and inventors and reduces spillover across firms. Therefore, the return to R&D at a society level goes down. How can we more concretely test the potential channels? Mike offered some of the suggestive evidence. Coming back to the citation as a measure of patent quality, that is one common measure in the literature. But we also often use citation as a measure of spillover and knowledge diffusion. In viewing the measurements, I think there is some extra digging that we can do to look at changing citations because if there is some non-compete that precludes spillover, we would also see a decline of citation to past patents. So, in terms of the measurement we are underestimating the effect, but also perhaps, given that we have these other quality measures, maybe citation could serve as a good measure for the spillover itself. In addition to citation, Mike offered evidence on the mobility decline of workers in general and you do have the movement of the inventors listed on these patents. So, looking at their movement across firms to see whether there is any decline from firm to firm. It would give us more direct evidence on this potential channel.

So finally, let me come to my third point in thinking about the incentives for firms to invest in innovation and particularly how noncompete can effect rents that firms get from the investments. Now I am going to put on my macroeconomist hat and think about a simple thought experiment. I think for this exercise we will think along the lines of a workhorse endogenous innovation and growth model. We can think about three types of innovations. I am going to focus only on innovation firms doing small incremental innovation on their existing products and climbing up the quality ladder versus another type of innovation. Creating new products or new varieties will create a destruction that innovates over other products and could potentially generate some business deals. The first type of innovation is only carried out by incumbent firms. The other two types? Entrants can contribute to these inventions as well. From the spillover channel, the last two types of innovation could potentially be more significantly affected because we are building on top of what others have done. But for the first one, the effect is not there, and now thinking about rent. From the incumbent's perspective that are carrying out small increment innovation, noncompete really gives the extra protection that they are probably going to enjoy. Getting the rent for longer and keeping inventors for longer. I would think that for existing firms that are carrying out incremental innovation building on their existing product, they are going to be the types of firms that can benefit from easier enforcement of noncompetes.

Going back to the empirical evidence; breaking down by types of firms and by types of patents, we can find some category of firms that are potentially getting some benefits out of it. Although, perhaps in aggregate, we are still losing innovation. This would also give us additional supporting evidence in teasing out different channels. Finally, there are reasons -macroeconomic evidence- that perhaps a lot of the innovations that firms are doing are small incremental innovation, building on their existing product. It would be good to find a small subset of firms that may be -and this is perhaps consistent with the empirical evidence that we see a larger decline in the big breakthrough patents, that they decline a lot, but the incremental ones decline a bit less.
I will leave it here. Thank you.
[Applause]

[Viola Chen] We have a few minutes for audience questions.

[Michael Lipsitz] I will take this time to thank Liyan. It gives me that new project excitement about a very old project. I am excited to take some of that back and get working on it.

[Q&A]
[Q: Heidi Williams] I am Heidi Williams. The language you used around reallocation is something I think you might want to expand on. I think your argument was that it is not just reallocation because we can look across these classes with the exposure measure, and I think that evidence is interesting. But you were kind of saying it, almost as if, if it were reallocation then we wouldn't care about it. But if you take it seriously, where your innovation is done really affects your productivity as an inventor and also the spillovers. And so, I think, just to second what Liyan was saying, looking inventor mobility directly and also thinking about matching with geographic areas of specialization based on the work, it looks like it could be important to look at more directly.

[A: Michael Lipsitz] That's a fantastic idea and it's something we have poked around with a little bit, at least on a very broad scale, comparing the magnitudes at a nationwide level with the primary estimates against the estimates that we get with the exposure design and thinking about the differences being due to spillovers that potentially you are identifying, but maybe doing it at a micro level and identifying where those clusters are and how that plays a role. That can be really important. So thank you.

[Q:] Boston University. I have a boring methodological question. What happens to the estimates when you don't do the stacked design, but run a standard diff-in-diff?

[A: Michael Lipsitz] I'll give you two answers. In terms of the numerical estimates, they get larger in magnitude. I don't know the exact amount. I'll leave it at that. They do get bigger. We think it's conservative to do the stacked approach. think we are picking up some of the staggered design issues when we use a standard two-way fixed effects model.

[Q: Marc Rysman] Mark Rysman, Boston University. I had two questions. You didn't talk about selection effects of workers into these jobs at all. And I wonder if innovative people want the ability to move around and start their own firm, there could be no 'water cooler' effect at all. Just worker selection could lead them into the same results that you have and I don't know how to distinguish between them in your data, but I wonder if that's something you've thought about. The second question is kind of related. It seems like the focus was on the legal ability to enforce noncompetes, but you didn't talk much about pickupable noncompetes and I have a sense it varies across industries and is not totally random which industries use them and which don't and how they might respond to these rule changes, so I'd love to hear more about that.

[A: Michael Lipsitz] Great questions, both of them. In terms of selection of inventors into jobs, I think that is a place that we could definitely expand. We can track inventors over time, so we can think about the productivity of an inventor who was at a certain job before a law change happened and see how it changes for those types of people. Contrasted with individuals who started their job after the law change or started patenting after the law change. Maybe that sort of dimension could help us answer that first question a little bit about selection into jobs. That's a really interesting thought. In terms of the second question, a little more frustrating is the data, access but we don't know which individuals have noncompetes, so in terms of understanding the exact dynamic there… It is difficult. It is a valid concern. Things that we can do are to break down the estimates by industry to try to get at that, but of course, there are other dimensions along which that will vary, not just the industry. We're not able to do too much with that unfortunately, but here's hoping for more and better data in the future.

[ ] Okay last question.

[Q] It's like another view of your results can be that the same amount of innovative activity is going on, except that nondisclosures and patents are substitutes for one another. Have you thought about doing it the other way around, to see whether there are movements in patent enforceability that then change the incidence of NDAs?

[A: Michael Lipsitz] That's a really interesting question. There are definitely data limitations that we'll think about. I mentioned that there is some noncompete uptake data that is just coming out now as part of some the big public surveys. Maybe in the future, this is something we can look at to find legal changes on the other side and see how it's impacting things. The one other frustrating factor is that trade secrets, by their nature, secret. So it's difficult to observe, and trying to figure out if there is a difference in the extent to which they use trade secrets. If anyone has any ideas, I would be happy to chat. But it seems more difficult than these very public patents which we can identify easily. But yeah, thanks for the suggestion, I appreciate it. Thank you everyone.
[Applause]
[End of Q&A]

[Eddie] Great, thanks, and now we have Kathryn Spier from Harvard Law School, talking to us about holding platforms liable.

[Kathryn Spier] Thank you very much for the invitation. This is joint work with Xinyu Hua. This is also part of an ongoing research project that I'm engaged with. So, at this stage your comments and feedback are particularly valuable.

Although we get tremendous value from online platforms, we also face a variety of risks, including privacy breaches, scams, purchasing dangerous products. -I'm still on the first page.- Platforms have been largely successful in fighting off these types of lawsuits. Section 230 protects platforms from harmful digital content that is created by others. Online marketplaces have successfully argued that they are not traditional sellers and so they should not be subject to the same liabilities that a traditional seller would face. To clarify, if you purchased a dangerous product from Target and it was defective and it caused a fire -imagine your house burned down and you suffered injuries- you would be able to bring a lawsuit not just against the manufacturer of the product but you could sue right down the vertical chain, you could sue the distributor and the manufacturer. Strict liability would apply in this situation. What's puzzling, and perhaps ironic, is that online platforms are not subject to the same liability that their brick-and-mortar counterparts are subject to. I think it's particularly ironic given that bad actors online are often judgement proof and it is difficult to identify or bring to justice parties who are causing online harm. Moreover, online platforms, arguably, have a lot of capability to detect, and nip in the bud, potential harms.

Just to give you some examples of platform related harms -these should be near and dear to the folks here at the FTC- we have Cambridge Analytica who paid $5 billion in penalties in 2019 and just last December, they paid $725 million to settle a class action claim. The potential harms that are caused, could include scammers getting phone numbers or personal information, identity theft. We also suffer from the inconvenience of having to change our passwords, of disruptions, and so on and so forth. Another example is Epic Games. Epic Games has its own free blockbuster game Fortnite, but they also carry many games from other independent vendors as well. They get revenues both from game sales as well as from the in-game purchases, they share revenues with vendors. Also, last December there was a settlement for $520 million and half of it was penalty and half of it was designated as refunds to the consumers. What are the harms here? The potential harms to users are privacy settings that are putting children at risk and dark patterns that trick kids and even adults into making unwanted purchases. A final example here is a Hoverboard. I have personally never been on one of these hoverboards but they look like a lot of fun. Apparently, they can also cause harm. There have been many fires as a consequence of these hoverboards and many lawsuits that have been brought as a consequence.

Amazon sells a lot of these hoverboards and for the most part, Amazon has been successful fighting off lawsuits for defective products, but there were a couple of hoverboard cases where the court did find Amazon responsible. This is actually interesting, because in a couple of cases Amazon was dealing with third party vendors in Asia and some of these vendors were not even using the Amazon fulfillment centers. The hoverboards were being shipped directly from China to the homeowner here in the United States. I have a quote here from one of the judges. He was a particularly economically oriented judge who said, 'Amazon is well situated to take cost effective measures to minimize the social costs of accidents.'

What does our paper do? Our paper asks a couple of questions. First. Should platforms be held liable when participants are harmed, and secondly, how much liability is appropriate? Let me give you an overview. We are going to dig deeper into this. We are going to be considering a two-sided platform, where users get utility from being on the platform and on the other side there are firms which will get interaction benefits from interacting with users. Some of the firms will be bad actors and may cause harm to the users. If the firms themselves have deep pockets and can pay in full for the harms that they cause, then it isn't necessary to hold the platforms liable. If firms have deep pockets and can be brought to justice, then we should just impose liability on the firms themselves. If you can force them to pay in full for the harm they cause others, then the firm will abstain from dangerous behaviors. They'll take precautions to avoid harm, etc. Problems arise when the firms are judgment proof. Either they don't have deep enough pockets to cover the harm that they are causing or they are outside of our jurisdiction and we cannot force them to do so. Then, holding platforms liable has some significant benefits. Our paper focuses on two things platforms can do. First, platforms will be given incentive to detect and remove bad actors from the platform. Secondly, platforms may have an incentive to use the price mechanism. They may be able to raise the interaction price in order to encourage the exit, or deterrence, of the bad firms. The relevant factors are going to include things like consumer information, market structure, the type of platform, etcetera. That's just a basic overview.

Here is an outline of my talk. I've already covered the introduction and the next thing I will do is a very brief literature review. Next, I will show you a simple numerical example to illustrate the main insights from the paper. Then I will generalize my numerical example to show you some of the nuances that come out of the model. Then finally, I will focus on three very relevant extensions. My baseline model has homogeneous consumers, all of whom participate on the platform. I will look at heterogeneous users who make participation decisions. We can think about that margin. I will also think about retail platforms where consumers must consent to the interaction and there is pricing going on between the firms and the users. Then I will think very briefly about platform competition.

Literature review. This is a paper that fits into two literatures, both the law and economics, and the platform economics literature. Within the law and economics literature, liability is a mechanism for correcting negative externalities. By holding bad actors, injurers, responsible for the harms they caused, it gives injurers an incentive to take precautions to avoid harm and it also gives them an incentive to scale back on the level of activity and abstain from engaging in dangerous activities to begin with. Judgment proofness is an idea that is very important. So, if the bad actors don't have deep pockets and cannot pay for the harms they are causing ex post, then it may make sense to extend liability to parties who can help to serve as gate papers. I have contributed to the literature by thinking about gun manufacturers and their role potentially in raising the price of guns, holding gun manufacturers liable in order to keep guns off the street. Others have looked at lenders, at banks that are lending money to dangerous operations. Or within managed care. You may want to extend liability to managed care organizations who can monitor the activities of doctors and others within their networks. The paper also fits into the multi sided platform literature. The most relevant thing for us right now, is that, in a multi sided platform, you may have cross subsidization. Indeed, it may be optimal for one side of the platform to pay a negative price or if that's not possible to have a 0 price. That's going to be a feature of the model that I am going to develop. One side of the platform will pay nothing. The platform will monetize through the other side. The literature on platform liability is in its infancy. There is very little written about it. A couple of policy papers have been written in the EU context. There are just a small number of recent working papers thinking about platform liability. So, this is a very exciting and policy-relevant area to be working on.

The next thing I will do is show the numerical example, just to illustrate the main ideas. It is highly simplified. We have a two-sided platform. On one side we have the users. They are going to join the platform for free and they are getting utility. They are getting these benefits from perhaps interacting with each other, as well as benefits that are coming from the platform itself. On the other side of the market, we have firms. We can think about these as being app developers or advertisers or others. These are the ones that are paying a price to the platform. This is how the platform monetizes the strategy. What we see here, is that firms on the other side of the market are getting an interaction benefit of $30, so 30 is the interaction benefit when they are interacting with each of the users. There are two kinds of firms. There are safe firms and there are bad actor firms. We imagine that some of these firms are potentially going to cause harm to the users. This is private information for those firms. In the picture, you see that the harmful firms are going to cause a loss of $40 to the users. Notice that, these harmful firms are socially inefficient. We would like them to be deterred from joining the platform to begin with. Because 40 is bigger than 30. In this very simple example, as in my baseline model, I'm thinking about a monopoly platform. User consent will not be required for the interaction. So, if the users join the platform, then they are going to be sitting ducks. Once they join, they are going to have to interact with these firms. The users are getting enough utility that is worth their while to join the platform.

Why do we need liability at all? What happens if there's no liability? This is a very static example. If there is no liability and the users are sitting ducks, then the platform is going to charge a price of close to $30. Because that's the willingness to pay of the two types of firms. The bad actors will join along with the good actors. And this is a bad outcome. The best solution would be to hold the firms themselves liable. If you can force the firms to pay for the damages that they are causing to the users then this will solve the problem. Not even thinking about the price that will be charged, think about the bad actors. They're going to get a gross benefit of 30, but they'll have to pay damages of 40. So, they will be deterred. They will choose not to join the platform at all. They will be deterred. That's great, but the problem is that it may not be possible to hold the firms 100% liable. The firms may be judgment proof. They might not have the money or they might be outside of the jurisdiction. If firms are judgement proof, then platform liability makes sense.

Let's talk about platform liability. Right now, I've cut off damages for the firms. We're just holding the platforms liable. We are holding them fully liable. You see it with the green line, they pay 40 whenever there is a damage caused. So, what is going to happen in this situation? What is the price that the platform will charge? The platform is in a tricky situation because they don't want those harmful firms on the platform, but they cannot screen them out on the basis of price alone. The platform is going to end up charging a price around 30, and they'll get both the harmful types and the safe firms joining the platform. What the platform can do, is engage in screening activities. They can audit. They can spend money to try and get rid of those harmful participants. So, that is what they would / will do. They will spend money to try to identify the harmful firms and kick them off. That's socially efficient. We want the platform to do that. Platforms do have the ability to detect bad behaviors. This is what the platforms themselves say. I know that Meta spent about $5 billion in 2022 and they employ tens of thousands of people to develop technology, to use machine learning, etc. Amazon also spent almost a million dollars in 2021, hiring many people to try to block suspicious listings, dangerous products, counterfeiting, etcetera. So, they do have the ability. Putting legal liability on them is a way to make them pay even more attention and to bring their incentives into better alignment with what is socially efficient. Ideally what we want to do is give these platforms the right incentives.

I just talked about what happens when the firms are totally judgment proof. I want to show you something else, which is a bit more subtle. Let's imagine that the firms are not totally judgment proof. What you can do, is force them to pay 25% of the damages. Now we have firms that are moderately judgement proof. They can pay $10 of the harm that they cause, not the full 30. What's going to happen in this case? This looks more promising than the previous examples. It's more promising because the harmful firms have a net willingness to pay per interaction of 20. The gross benefit is 30, the damages are 10, so the net benefit is 20. They have a lower willingness to pay than the safe firms. One thing that the platform could do in order to deter the harmful firms is to charge a high price for interaction. All the platform has to do is charge a price of $30 and the problem is solved. If the platform charges 30, then the harmful firms are going to be deterred. They won't join voluntarily. But here's the problem. It may not be in the platform's interest to charge a price of 30. If there are many sufficiently harmful firms out there -imagine that it's 50/50, 50% of them are safe and 50% are harmful- then the platform would rather accommodate the harmful firms, charge a low price of 20 and get everybody to join the platform. There's a problem when the firms are moderately judgement proof. Adding in residual liability for the platforms, forcing the platform to pay the residual damages -that is the extra little green line here- causes the platform to have to pay the 30. This is going to create the right incentives. Now the platform is really suffering a loss when a harmful firm joins the platform. They're losing money of the harmful firms. The platform then has a financial incentive to raise the price to 30 and the harmful firms are deterred. That is the basic idea of the basic model in the paper. It is a simple framework. The next thing I am going to do is show you a little bit of the more general model and some of the extensions.

In the baseline model in the paper, we are also imagining that users are bystanders. They are going to join the platform. They are homogeneous. There's a fraction of harmful firms, some are harmful and some are safe. We're allowing the harmful firms to have higher interaction benefit than the safe firms. In a retail context, that makes a lot of sense. If we are thinking about retailers and manufacturers, manufacturers that have unsafe products may be cutting corners on their costs. They have lower marginal cost of production, and hence, may have a higher willingness to pay. So, the alpha, that's the interaction benefit that is higher for the harmful firms. Harmful firms also cause harm with a greater probability, theta H is greater than theta L. We are assuming that the harmful firms are socially inefficient. That the net benefit associated with the harmful interaction is negative. Therefore, we would like to prevent them from joining the platform. The platform can prevent harm in two ways. They may be able to do it by raising the interaction price to deter harmful firms. Or they can tap into a detection mechanism. Auditing. They can take effort E where the cost of effort follows all the usual properties. The basic timing of this game -it's a very static game. First the platform chooses the price, then the firms learn their types, and the firms or users decide whether to join the platform, the platform then chooses its effort level. That's after the parties have joined. The platform chooses its effort and then blocks the detected harmful types. Then the firm user interactions take place and lawsuits occur if users were harmed. The liability rule will allocate liability between the firms and the platform. WS is the fraction of harm being paid by the firm and WP is what the platform is paying. It is a Stritch liability rule that we are looking at.

The basic insights of the paper can be seen by looking at a few equations. I only have two slides here on this general model. The first idea I want you to see here, is that either the harmful firm or the safe firm might have a lower willingness to pay to be on the platform. A type I firm is willing to join the platform when their interaction benefit minus the firm's expected damages, the theta iWs, is greater than the interaction price they have to pay. What this is showing us, is that if firms cannot pay anything and are totally judgment proof, as in the WS is equal to 0, then the harmful firms have a higher willingness to pay, because alpha H is bigger than alpha L. In that case, it's the safe firms that are marginal. They are the ones that are on the cusp and it will be difficult for the platform to exclude firms. The platform will have to engage in auditing in that case. On the other hand, if we have a full firm liability, the harmful firms are going to be completely deterred. They will not want to join the platform, even if the price was equal to zero for joining. In general, either type can be marginal and this will hinge on the degree of firm liability, whether the firms are themselves judgment proof.

I want to show you two cases. In this first case, where the firms are quite judgment proof, so WS is small. In that case, it's the safe firms that are marginal. The platform in this case will set an interaction price that is tailored to the marginal firm, the safe firm. Then the harmful firms are capturing information rents. We are going to compare the social and the private optimum. We have a social planner in our model. The social planner is concerned about everybody. They are concerned about aggregate welfare, the welfare of the platform, the welfare of the firms, the welfare of the users.

If the platform is in control and the social planner could choose how much effort to put in to remove harmful firms, the social planner would choose its effort to equate the marginal social benefit. The marginal social benefit here, that first expression on the right side of the equation, are the damages that arise from a harmful firm being on the platform. Alpha H minus theta Hd. They are going to equate the marginal social benefit of kicking off a harmful firm against the cost of engaging in the auditing. That's the C prime. That's what the social planner would do. What about the platform? The platforms' incentives are not necessarily aligned with the social planners'. The platform is interested in profits.

In the paper, you can see how we drive this, but you can show a real divergent between the private incentives and what the social planner would do. There's partial alignment on the right side for the platform. What we see is the S prime, the marginal benefit from a social perspective. What the platform is not taking adequate account of, are the benefits that will accrue to users who can avoid being harmed. The D-W on the right-hand side are the uncompensated losses for users. When the platform audits a little bit more and kicks off a firm, they save users from having to suffer uncompensated harms. The platform is also not taking into account the rents that are accruing to the harmful firms that are remaining on the platform. We see that auditing is conferring both a positive externality on the user bystanders, but also a negative externality on the harmful firms. What is the upshot of this? One thing you can see just from that equation is that full liability on the platform is not going to be optimal. We don't want to make the platform fully liable for a residual harm. If W is equal to D, then that first expression on the right-hand side would be equal to 0. The platform would engage in excessive auditing. They're going to be overzealous in their efforts. You don't want this excessive overzealous auditing. What you want to do is to scale back and have less than full liability on the platform. Then you can fine tune it to balance out the negative externality and the positive externality. This was the case where the firms were very judgment proof.

On the next slide, we have the other case. This is the case where the firms are moderately judgment proof. They can pay for some, but not all of the damages that they are causing. In this case it's the harmful firms that are marginal. The harmful firms are lower on the demand curve than the safe firms. The safe firms are potentially getting rents. There is promise here, because if the platform were to charge a price that is tailored to the safe users, then we can deter the harmful types from joining to begin with. The problem, is that the platform may not want to raise the price to deter the harmful firms. What the platform is going to do, is to think about; what are our benefits from accommodating the harmful firms? What kind of margins are we going to get from the harmful guys. Then they're going to weigh against the information rents that are being captured by the safe firms. It's a standard type of trade off. In this world, holding a platform fully liable for the residual damages is going to align the incentives and get the platform to deter the harmful firms.

We do several extensions in the paper. One of the extensions is user participation. In what I showed you before, the users were essentially sitting ducks. They all chose to join the platform because they were getting so many benefits for being on it. We have a detailed extension where users are heterogeneous and make a decision whether to participate or not. Platform liability has an additional benefit in this world, if users believe that a platform is being held liable and they believe that the platform will take more effort to make the platform safer, then users are more willing to join. Users are going to rationally anticipate that the platform is safer and are also going to be getting a rebate in the future from liability. As a consequence, they are going to be increasing their participation. What we show is that the platform liability is weakly higher than in the baseline model. Weakly higher because the platform is not considering some of the benefits of increased participation including the rents that are accruing to some of the firms. Relevant for the FTC -I have about 59 seconds to ago- is our retail platforms, so we have an extension where consumers must consent and there is transfer pricing between the firms and the users. They are being matched together. Here it's more involved, because now interactions are going to require that the users participate in these individual interactions. We showed that if the harmful firms are marginal, then platform liability is unnecessary. If the harmful firms are marginal, the platform has an incentive to raise the price, which is observable to consumers as well, who are hyper rational in our model. That's going to lead to improvements in profit for the platform. On the other hand, if it's the safe firms that are marginal, then we do need to have platform liability.

I'm just going to get right to my concluding remarks. Should platforms be held liable? Maybe yes. We are making the best law and economics case for platform liability. We have ignored some other things,

including litigation costs. If there are significant litigation costs, then our results may be reversed. Depending upon how many inefficiencies come from that. We have many detailed extensions within the paper as well that look at that issue too. I'm going to end. Thank you.
[Applause]

[Eddie] Great. Now we have Marc Rysman from Boston University to discuss.

[Marc Rysman] Thanks so much. It's really wonderful to be here and to be part of the conference, and be asked to discuss such a thought provoking and really groundbreaking paper. So I don't see, do I do? Great. So there's my name.
Yeah, so public policy establishes who's in trouble when bad things happen. Kathy gave some examples of how internet platforms are not liable, but other platforms are really liable. Banks have 'know your customer' rules. Credit card platforms can often be held liable for transmitting or facilitating bad stuff happening. These rules of which platforms are liable and which ones are not, are really strikingly different across different categories industries. As Kathy pointed out, a lot of bad stuff can come across platforms. Misinformation, counterfeit products. The point of this paper is; can we use the rules about damages to know who's owed what and when, in a way to optimize outcomes in some way?

I'm going to rehash the very basics of the model. Kathy did a great job. I'm just going to set up some basics of the model to make some points. Then I have one slide of comments to think about. As we saw, the platform connects buyers and sellers. There are two types, the high and low type. The high types are going to do more damage to the consumers than the low types. The consumers don't do anything, they just sit there and take it. The consumers are getting this value V from the platform, and then they get randomly matched to either a bad or a good seller. Then they face some damages they and the damages they face are a function of this theta, whether it's a high type and a low type. The sellers are getting some benefit from interacting with the consumer. You can think of them as advertisers in this case. They can be on Facebook or something. That's the sense where consumers aren't really making a choice whether or not to interact with them. The advertisers are getting some benefit alpha and they have to pay a price platform. What you can see here, is the consumer suffers the damages and the other people really don't care about them. Because they don't face any harm from these damages. As long as the consumers are signing up. They are going to face the damages, but it doesn't really affect the seller payoffs or the platform payoffs whether they face these damages or not.

We will make some assumptions. The high types cause a negative payoff. The consumers are losing out. That's a bad thing for the world and for the consumers when they face these damages, but the consumers are still happy enough that they're going to sign up for the platform and absorb the occasional or even possibly frequent bad damages. The big result here, in Kathy's paper, is that the platform does nothing to prevent the bad cells. They just don't care. They're benefiting from the sellers paying them. The baseline of the model is platforms do nothing to prevent bad sellers. I think, in practice, without damages the platforms would have some incentives. Obviously, this is a distraction and I'm going to focus on the main points in her paper. But just as you think about what's going on in this paper, where the baseline is that the platforms are just letting it happen.

Now we are going to allow this extra tool. The government can come in and set damages. They can set this payment that goes from the sellers or from the platform to the consumers. Now I'm just adding this stuff in pink here. The sellers has to pay damages, the platform has to pay damages and the consumer gets those damages. That's going to change the payoffs in this model. That's the government's problem, to set these damages in a way that gets the optimal outcome and then somebody has to respect these ability constraints. Perhaps the sellers can't be sifting, maybe we kind of WS, stuck at 0. Because we can't go find these people, or maybe there's so many of them and they are little. Then the usual enforcement mechanisms don't work. Consumers will never bring lawsuits against little sellers. That's going to give us a reason to assign some level of damage responsibility to the platform.

Then what happens if the high types are less profitable for the platform than the low types? If it was just WS, the platform wouldn't care. They'd just let things go. But now, when there's this WP, these damage that the platform has to pay, now the platform's like, 'wait a minute. I don't want to have to pay damages. This seems bad. I got to get rid of these bad guys.' One way, is to set price so that they don't want to join the platform anymore. But if the level of damages and benefits is out there, where there isn't a way to deter the bad guys without also deterring the good with price differences. Then you get this alternative choice, which is to engage in some screening. This optimal level of screening is a big part of the paper. Screening can get rid of half of the bad guys and pay a cost of C half. And this cost goes up very quickly. You never want to screen out all the bad guys, but that's kind of where the paper is at. It starts deriving first order conditions and we get this result -I'm not going to go too much deeper with the math- but the platform can engage in too little or too much screening. It depends on the size of the benefits and the size of the damages and how the government has set these damages.

It's very elegant, this part of the paper. I'll just say, if I had written this paper, it would be this big mess of W, lambda, and other stuff. But she somehow organizes this it in the slides, so that you can see the economic effects of how the platforms are different from the social planners. I thought that part of the paper was really nice. A bit of economics.

Let me take my last three minutes to make some comments. One of the assumptions in the paper is that when you screen out the H types or when the H types don't show up, the consumer loses out and they don't get any of their interactions. You might think that once those H types are gone, they can go better find their low types. They can go find the good sellers. Now the benefits to screening are much higher and the benefits of getting rid of the bad guys are much higher. That might change some of the outcomes of the paper in an interesting way. Much of the paper is written as if there is full coverage, so all the consumers want to buy. They want to be on the platform. That's where you see those results. If there's no damages at all and there's no incentive to screen. Exactly how those things change when that's relaxed, is not part of the paper, but it could be. I guess one set of results that I kept waiting to see, is this spends type of result where the incentive is to engage in screening. The social planner cares about the average effect and the platform cares about the marginal effect, and that's kind of waved away. I think there was an assumption of homogeneity. At least there could be a comment about that point.

Then the last one. I think the biggest point is that, a lot of times the damages -the paper is using this concept of damages- and it makes me think about the consumer suing either the platform or the seller in court, and being awarded damages in some cases. But I think a lot of times the harm from this bad behavior that Kathy is talking about is not really to the consumer directly. So, for example, Cambridge Analytica, let's say they're stealing data. The consumers are harmed by that, but there's a lot of cases of a spread of misinformation and we want to punish the platform for allowing misinformation to spread. I'm not sure that's exactly the harm, in the sense of a consumer suffering harm. In many of the cases that Kathy mentioned in her examples, there wasn't always damage that she was describing but fines. Those were more for breaking the law. And those may or may not be driven by consumer damages and the way that we normally talk about it.

So just blowing up, this last point about what if sellers cause damages, but not to consumers. Political information, counterfeit products -which is this great extension that she touched on for a minute where the seller is actually selling something to the consumer and they're buying something. In that case, the consumer may want to have the option to buy counterfeit products. There may even be a real benefit there. The person that's being harmed is the brand holder, the brand owner/IP holder is the one that's being harmed. They are a separate player and they are the ones that want to sue for damages and the sale price that's going between the consumer and the seller in that case may not be a good measure. There may be no damages there at all from their perspective.

The I baseline model where there's no interaction, actually doesn't change very much if we think of that damage not being to consumers. Here the damages are to the buyers, but if that damage was kind of more general to society and it wasn't really sustained by buyers, I don't think things changed that much. I think the model could accommodate that. You could write the paper to say that it's more general than just the consumer suffering damages. When you get into the case where there is selling between the buyer and the seller and there's a price that's being set- there my intuition ended with how far I could get without trying to solve a new model. So, I don't really know if it's going to extend so naturally to that case.
But I'm out of time. So, I will stop here.

I will just conclude here. I just want to say, really clever and thoughtful paper on a really important topic. There's a lot of really cool extensions. She sort of touched on a few of them at the end that really contribute to the robustness of what's going on here and really suggests that how we set damages are a really important element of policy that we have for incentivizing platforms to do the right thing. Thank you.
[Applause]

[Eddie] Great. We have time for a couple of questions.

[Q&A]
[Q: Ben Castner] Hey, Ben Castner, FTC. This is a really great paper. I'm having trouble choosing from among the many questions that I have written down here. The thing that struck me, is that there is this part of the platform governance literature that goes over the benefits -that market platforms can have- from having low quality sellers in terms of additional value extraction through selling prominence or

softening competition between sellers. I am wondering if that is going to have an effect on your result that when the harmful sellers are marginal you don't really need the platform liability?

[A: Kathryn Spier] That's a great question. In our model, our users, our consumers, all have homogeneous preferences. In the literature that you are referring to, consumers presumably, there's opportunities for vertical differentiation and maybe some price discrimination. At that point, I would also have to think about how that dimension of differentiation is then correlated with the harms that are being caused by the products, I mean, there are some low-quality products that are low quality in a good way. You know? You get what you pay for. It's not that they're dangerous, per se. The question would be; are those low-quality products correlated with causing harm to people? I would have to think much more carefully about that. Thank you very much for the question.

[Q: Flora Nedra] Flora Nedra. I very much enjoyed this. Particularly the result that you can also get too much screening because it reminded me, in a way, of the walled garden that Apple had, and that sort of alleges to exclude some of the sellers that would be valuable for some of these consumers. I wonder if you can comment or conjecture a little bit about what would happen with platform competitions since you only have a monopoly platform. Would those same insights also extend to the competition case?

[A: Kathryn Spier] That's great. I'm afraid that I ran out of time at the end and I didn't get to my platform competition part. We have a very limited treatment of platform competition. We look at a hoteling model where we have two platforms, one on each end and they are competing for users. Having some platform competition can really help, in so far as the platforms want to make platforms look more appealing for users. So, they are going to try to create an illusion of higher safety and if they can, in a visible way, screen out harmful firms, then they have an incentive to do that. Now, of course, if they are very differentiated, there are going to be limits to what they can do. This is very underdeveloped in our paper, we have just a toy model, and I see this as a real growth opportunity. A new research opportunity to think and engage more fully and comprehensively with competitive issues. What we have currently is a consumer protection kind of paper and not a competition policy paper. Thank you for the question.

[Q: Michael Schwartz] Michael Schwartz from Dice. Here. Over here. First of all, thanks for this very interesting talk. I was wondering if maybe you had thought about endogenous benefits because, of course, firms can adjust their pricing and whether then some of those cases are more likely to arise than others.
[A: Kathryn Spier] So endogenous benefits to who?

[Q: Michael Schwartz] To consumers.

[A: Kathryn Spier] So we've done some modeling of network effects. Same side network effects. Thinking about, on the user side, how having more users creates more value. And it surely added some complexities. In our playing around with it, we weren't getting interesting interactions with the liability rule. But we didn't push that far. I would welcome further suggestions from you or others about which types of benefits we might want to build into the analysis.

[Q: Michael Schwartz] Thank you.
[End Q&A]

[Eddie] Great. Thanks everyone. We are going to break now until 2:50. So, thanks so much.

[Tom Koch] Welcome back. I'm Tom Koch. I'm a staff economist here at the Federal Trade Commission and I have the honor and opportunity to introduce the participants in the session organized by Heidi Williams. Today we will be starting with her keynote address. Doctor Heidi Williams is a professor of economics at Dartmouth college and the director of science policy at the Institute for Progress, a nonpartisan think tank focused on innovation policy.

[Heidi Williams] Someone mentioned to me on the break that they had noticed that I was not an IO economist which is true. But don't worry because I'm not talking about IO today.
The field of economics provides a language and toolkit for tackling fundamental questions about how to best design public policies to achieve outcomes like economic growth. When we think about research, we often think about purely curiosity driven research, and what I want to talk about today is, that curiosity driven research can often miss opportunities to inform key policy debates where rigorous nonpartisan objective economic analysis can have the most impact. Some people argue that those kinds of research are very distinct and that curiosity driven research can be done in universities, which are driven by their own sense of incentives around institutions like tenure and pure recognition of people's work. And that more applied researchers working in non-economic institutions can focus on these shorter-term problems. In my view that characterization is a mistake, in the sense that it misses broad swaths of where important impactful research can happen. The view that I hold is more closely associated with a political scientist named Donald Stokes, who put forward a view that he thought we needed basic research in what he called Pasteur's quadrant. So, this book that he has from 1997 basically categorized research in two different dimensions. First was relevance to immediate application, and second was relevance to generalized knowledge. So, in is very simple two by two framework, he talked about this upper right hand side quadrant as being Pasteur's quadrant, which was inspired by the fact that Louis Pasteur and a lot of his fundamental basic research on pasteurization was itself inspired by directly observing manufacturers that were struggling with bacterial contamination of milk and wine.

So, at a broad level. What is use-inspired economic policy research look like? In my view, it stems from understanding the bottlenecks and challenges that are facing the relevant government agencies. So, it's a funny thing to talk about here because obviously at places like the FTC, I would say that translation happens almost immediately. What I mean by that, is that the problems themselves almost immediately present themselves in language that we as economists speak and understand. I would also characterize a lot of macroeconomic policy analysis the same way. Almost by construction there's not any translation that's needed. There are some cases, like healthcare, that I would put in between. I think a lot of healthcare policy research requires lawyers to translate things into language that we as economists

understand, but I think there tends to be a pretty rich ecosystem for that translation to happen. It means that a lot of economists are working very directly on these policy relevant questions.
So, what I'm talking about today is the area of innovation policy, which is the area that I work in and the problem that I perceive, is that, a lot of the most important policy bottlenecks present themselves in ways that we as economists often don't realize, that are amenable to our toolkit our and our ways of thinking. That means that we can often miss opportunities for where research can have incredibly high impact.
So, what do I mean when I say innovation policy? Institutions like the National Institutes of Health and the National Science Foundation fund basic research. Places like Patent Office and the US Food and Drug Administration have regulations that shape which scientific discoveries get translated into real world impacts. Immigration policy dictates how many STEM advanced degree holders like PHDs can either stay in or work in the United States. Each of those institutions is itself incredibly complicated to understand, and so if you take just the National Institutes of Health alone, the NIH consists of 27 different centers and institutes, each of which has its own set of rules and regulations for how it funds scientific research and basically you would want to study each of those in isolation, if you are going to be optimizing how they give out research funding.

So, policymakers and politicians have long debated how to best structure those kinds of innovation supporting institutions, asking how we as a society can encourage scientific progress for the social good. Just to give one example, in the mid twentieth century, Vannevar Bush famously clashed with Senator Harley Kilgore over where federal research investments should be focused. So, should the government focus, as Bush argued, on funding basic scientific research, with the idea that private firms could translate those basic advances into real world outcomes? Or should the government instead focus its public dollars, as Kilgore argued, on applied research that was aiming to make direct progress on society's problems?

So, I think of our role as economic researchers, as laying the groundwork for those kind of Bush/Kilgore debates to be debated based on facts and evidence rather than just on people's ideologies or political views. But we as researchers don't generally get to influence whether or when those debates take place. It's funny to pick up these very old terms like, Stokes and Pasteur's quadrant and all these things. But just use one more, people sometimes use the phrase the Overton window to define, at a given point in time, the range of ideas that the public and politicians on their behalf are willing to consider and accept.
To give one example for innovation policy. In my view, the COVID pandemic created an Overton window to raise discussions of whether the way that NIH and NSF fund scientific research could be improved. Because I think the urgency of needing to get grant funding out the door in the middle of a crisis really laid bare some structural challenges that are inherent in the way that we use scientific research funding on a day-to-day basis, not during a crisis.

So that one was obviously obvious for everyone. However, I think a lot of shifts in the Overton window are not that obvious and so, the example that I'm going to focus on today is for high skilled immigration. So, a casual consumer, which I would classify myself as until I started thinking about this problem. When I think immigration is this very politically intractable issue, where Democrats and Republicans have just been at loggerheads for decades, but in fact motivated by concerns around US competitiveness and China and national security, there's actually emerged in recent years, a very strong bipartisan consensus on one key pillar of immigration reform, which is attracting and securing highly skilled foreign born STEM degree holders to the United States. That shift started a few years ago and moved high skilled immigration reform into the Overton window. Unfortunately, as best I can tell very few economists were there to notice that shift. Even more unfortunately, we as economists doing largely curiosity driven research basically failed to lay the groundwork for the data and research that would have made it easier to have informed policy debates now that we are at this moment of opportunity in the Overton window.

I realize this is an unconventional structure for a talk. I am going to talk about research, but mostly I'm going to tell you a story about basically trying to motivate where I think inspired research from economists can be quite helpful. Sometimes I kind of talk about policy bottlenecks as if they are very mysterious elusive things, and I think very much in the spirit of Steve's remarks, that's kind of the opposite of how I view it. I feel like if you go out in the world, and you kind of ask non economists, who work adjacent to you, about what's in the way of the thing that they think should be happening, I think they will often give us answers that we might not have thought of on our own.

To me, many of the most important questions in this Pasteur's quadrant space are in the heads of staff at government agencies. By staff I mean non-economist staff. I think encouraging them to articulate and publicly disclose areas of empirical uncertainty and the decisions that they've based on their job can be a really important source of questions for us as economists to think about as research questions. That is one reason that I asked and applied to be editor of the Journal of Economic Perspectives because I thought that running a solicit journal would give me an opportunity to ask people to share things that I thought we as a profession could learn from.

In that spirit, I reached out to Jeff Clayton last year, who is the director of research at the Congressional Budget Office, to ask if Jeff and the CBO staff would be willing to write a paper for the JEP. They came back with a proposal that I loved, which was essentially saying, 'here are some examples of ways that we use economic research and here's why the stuff that economists are doing is not as useful as it could be and where more and different directions of economic research could be useful to the CBO.' They also disclosed the substance of that on their blog last July, so even though the paper is not published I'm not talking about anything that's not publicly disclosed. If you read through those blog posts, you will see that the work of CBO is really broad. They're asking about how best to model the impacts of federal permanent reform underneath. They're interested in how to think about estimating borrowing responses to changes in student loan repayment terms. There are some more entertaining ones including a puzzle about why inflation for shipbuilding has been so much higher, consistently, than regular rates of inflation. Because it turns out the CBO is tasked each year to study the costs of shipbuilding for the Navy.
The question that most caught my eye was that CBO expressed interest in understanding the relationship between immigration and productivity. Unpacking why I thought it was interesting that they asked that question, and why it's important that they ask that question, requires diving into some details that I hope you will be patient for. But I think the details of this are incredibly consequential.

I think it's fair to summarize that the state of economics literature is not just intuition, but data suggests that the US economy would grow more if we admitted more high skilled immigrants to the economy. There are a lot of papers that you could point to on that. However, in essence, large numbers of skilled immigrants are eager to move to the US. When they move here their research looks highly productive. They generate valuable innovations at disproportionately high rates relative to natives and they start new firms, that lead them to act more as job creators than job takers for US natives.

More directly relevant to CBO, is not those economic impacts, but rather the federal budgetary impacts. In terms of fiscal impacts, a recent National Academy's report concluded that, on average, high skilled immigrants and their descendants contribute hundreds of thousands of dollars more in tax revenues than they receive in benefits. Yet, if you go talk to hill staffers that work on immigration legislation, they describe it as an open secret that when they give something to CBO to be scored for admitting more STEM PHDs to the US, for example on additional green cards, what CBO reports back to them is that that's a huge cost to the federal budget, rather than a benefit. When I first heard that, I was quite confused because I really like CBO and I think they do great work. I was thinking, why would that happen? Why would the CBO and the staff of the joint committee of taxation be reporting back something back that is so obviously at odds with all of the research that we have produced? And that's all explained. CBO and JCT didn't make a mistake, they were following the rules that are specified by Congress and the rules that we have down basically direct CBO to do something that doesn't really make sense in this case.

If you're not that familiar with budget scoring, which at least I wasn't started until I started thinking about this, when Congress wants to know what the federal budgetary impacts of a proposal are, they ask for a score of the proposal from CBO and the joint committee on taxation. CBO provides the spending estimates. JCT provides the revenue estimates. Conventional scores, the default scores, are including two things. First is, mechanical effects and the second is behavioral responses. I want to emphasize that because I think there's a lot of misunderstanding about what and is not captured in conventional scores. For example, conventional scores for flu vaccine subsidies, would include if you hold people's vaccination behavior fixed. It's more expensive if you're giving federal subsidies, but they do also incorporate these behavioral responses that if you subsidize flu vaccines, more people are going to get vaccinated. So, both of those are included in conventional scores.

What conventional scores don't take into account, is how a policy could affect macroeconomic variables, like employment, growth and productivity in the US economy. At a very fundamental level, what that means, is that conventional scores take GDP as given, even if the policy that we are asking CBO to score, is by construction intended to boost productivity and growth. Those kinds of effects are included in so-called dynamic scores, but not in conventional scores, and that distinction matters because conventional and dynamic scores can be quite different.

Most of you, those of you that have heard of this, I'm sure are familiar that this mostly comes up from conservative lawmakers when they want to talk about tax cuts. Conservative lawmakers argue that tax cuts could increase economic activity and spur economic growth. Effects that wouldn't be included in conventional scores. But if you did a dynamic score, in their view, which hasn't come out, they can make tax cuts less costly for the government or even save the government money. Public debate, when it has happened, over dynamic scoring, largely focused on that tax example, and dynamic scoring became branded as this very conservative partisan idea. The problem with that is that dynamic scoring shouldn't be partisan.

In many cases conventional and dynamic scores are going to be similar, but in cases where they differ, good dynamic scorers are going to provide more accurate evidence to Congress and to the public about what the federal budgetary implications are, of a given proposal. It's that true dynamic scoring is more time consuming and more complicated, rather than just holding the economy effects. Yes, merging on a macro model takes more time, and because of that reason, it's not the right tool for all occasions. We should not just dynamically score everything.

My co-author and collaborator on this work, former CBO director Doug Elmendorf, has argued that dynamic scoring should be applied to major legislation. Which is essentially saying we should apply it when we think the macro effects are big. But a complementary approach that we have now argued for, is that we should try to anticipate context in which we expect conventional and dynamic stores to differ. So even if the macro effects are small, if they are large relative to the conventional scores, that means that the dynamic scoring could flip the sign of the conventional scores. That's essentially the case that we are in with high skilled immigration.

More generally, policies that want to spur research and innovation in productivity growth in the economy are often going to look like cost to the federal budget on conventional scores but are going to look productivity improving if you do a dynamic model. So, when we apply conventional scoring to growth policies, we are essentially choosing to tell lawmakers and the public inaccurate estimates of the federal budgetary impacts. Dynamic scoring matters for high skilled immigration in particular because the current guidance that Congress gives to CBO and the GCT on how to score immigration legislation holds population fixed at this baseline projection.

What do immigration policy changes do? They admit more people to the economy than were here otherwise, but if you're holding population fixed, you're not changing the population. The way that flows through in practice, is that CBO does tabulate that those people are eligible for federal benefits. For example, they might take up subsidies under the Affordable Care Act. I think that's the largest component of what those benefit eligibilities are, but Congress doesn't ask CBO to tabulate the wage tax revenues that we would earn from these relatively high earning STEM PHDs coming to the US. That's essentially where the slippage comes in, where we are tabulating the cost, but we are not tabulating the tax revenue. That's where we are getting this negative score on these STEM green card PHD proposals. That's just the wage tax revenue component. There are tons of studies saying these people would come here and really improve productivity growth and the economy. That's even before accounting for what you would get from a more fully dynamic model.

Just to illustrate the magnitudes, I want to give two examples. It's not exactly the same proposal scored two different ways, but it's quite close. The Skills Visa Act in 2013 and something called section 8303 of the America Competes Act. They both essentially proposed uncapped green cards for STEM advanced degree holders. They were scored using different methods. The section 8303 proposal was scored conventionally, so the default way, as increasing the deficit by about $3.1 billion over 10 years. Whereas the Skills Visa Act was scored as what's called partially dynamically, which just means that you are adjusting for the wage tax revenues that people bring in, and that was scored as deficit reducing by $118 billion over 10 years.

To be clear, that's a sign flip and the magnitude of the gap is around $121 billion over the 10-year budget window. You could say, 'well, that's not exactly the same proposal,' and I'm sympathetic to wanting a more precise comparison. So, Doug Elmendorf and I basically went to the Penn Wharton budget model team and said, 'can we contrast the same proposal getting scored in different methods? Just so anyone who wants more precise answers about how much the choice of scoring matters in this context can have a really crystal-clear comparison of those two.' Our goal in publishing network is just going to be to spur a better-informed debate for anybody who cares about this, on how much the guidance from Congress matters over the scoring methods.

That was a lot of budget details, which I realized may have been a little bit in the weeds. At a high level, I would call this issue of the choice of budget scoring methods for high skilled immigration, a policy bottleneck. What I mean by that, is that Congress is setting rules that are unintentionally having this consequence, but it's kind of getting risen in a case where we as economists weren't even paying attention enough to recognize that this was something that was happening in proposed legislation.
I think that matters directly for me as an economist because I really care about CBO reporting correct estimates to congress, and to the public. Because I think we should be debating the merits of policy proposals based on accurate information. If Congress doesn't want to provide green cards to STEM PHDs because they don't want to, that's fine, but if they're not providing them, because we are telling them the wrong answers about the federal budgetary impacts, that's terrible. And I care about that. But I think it's also important to say that CBO scores matter indirectly. In a much larger sense than that. Because Congress is complicated and budget processes are complicated, many types of requests and proposals considered by Congress are subject to these very specific budgetary and procedural rules. Changes in the estimated budgetary impacts of a given proposal can also affect what types of requests and proposals are made of Congress in the future.

Just to give a concrete example. In the past few years, there's been a provision in the National Defense Authorization Act, which is this 'must pass bill' for defense funding, where each of the individual agencies under the department of defense flexibility to request STEM advanced degree holder green cards for science and engineering workforce needs in the defense sector. That provision has been removed from consideration because of complication stemming from the negative score that came in from CBO. The reason for that is that defense provisions in the NDAA have to be deficit neutral. So, you get this big negative score coming in and that fell out of the policy discussion. This is something that I would call a policy bottleneck, and I think it's a policy bottleneck that matters. The data and empirical work needed to inform Congress on the implications of different scoring methods are things that we, as economists, need to provide. Whether we are internal to CBO or external to CBO. The problem is, if we as economists aren't even aware this is a problem that needs our toolkit, we are not even paying attention to this. Just to loop back to some language that I used earlier, I first learned of this issue earlier from talking to someone who said, 'oh yeah, everyone on the Hill knows that,' and I was like, 'how could none of the economists that work on immigration have never even heard of this.' But it turns out they were completely right. You go talk to anybody that's worked on immigration legislation on the hill, and they're like, 'oh, yeah, that weird thing that we didn't understand.' Of course, hill staffers are for the most part not economists. They lack the time and patience to run down what the underlying issue is. I will say that it's a testament to CBO's transparency, that as an economist who literally knew nothing about budget scoring, I was able to comb through all of their past reports. They are so transparent that you can trace out the exact history. How and when has conventional and partially dynamic and dynamic scoring been considered for immigration legislation overtime and basically being able to piece out that comparison that I mentioned on how much it looks like this matters. That research formed the basis for me asking Doug Elmendorf to do this collaboration. And I'm hoping it's going to provide the basis for more informed collaboration among the relevant Congressional committees about what scoring methods they want to ask the CBO to apply to immigration legislation.

So, imagine now that we're in a world where some key politicians on the relevant Congressional committees decide that they would like to do more dynamic scoring of immigration. They would need to decide, as Doug says they teach at the Kennedy school, that dynamic scoring would be technically correct, administratively feasible and politically sustainable. So even in that situation, we as economists, have not provided the data and research that CBO would need to actually have on hand to do the dynamic scoring models. Let me take a moment to take explain why that's true. The question that CBO disclosed is that they wanted to know about the relationship between immigration and productivity. They specifically said they wanted to look at that separately by skill groups. Because I think there's a lot of evidence in the literature that the productivity effects of immigration would depend on different skill group levels. That's a hard question. It's not that nobody's ever worked on it, but it's a classically tricky question to answer.

Let me just unpack why. You can go back to at least to the Solow 1957 paper. Economists were really interested in understanding the drivers of productivity growth in the economy. There were a lot of really nice microeconomic studies that have documented evidence that really suggests that the kinds of mechanisms that would encourage productivity growth are very much present for high skilled immigrants. One recent example was a paper by Bernstein, Diamond and others. There are several papers that you could point to that are providing these microeconomic mechanisms. However, quantifying and measuring the contribution of immigrants to US productivity growth on the whole, is hard. Let me just take as a point of contrast. If we think about how do R&D tax credits at the state level impact in venture mobility. You can look at it using a standard design of state law changes. But for productivity, we think that if high skilled immigrants move to one area of the US, they're going to generate spillovers across technological areas and across geographic areas that make it much harder to quantify the full contribution to productivity growth in the economy. Given the challenges that we know are inherent in measuring that kind of paper trail of spillovers it's perhaps not surprising that there's extremely few research papers on this topic. When I thought about CBO's question, it struck me as a perfect example of work in this Pasteur's quadrant space. So, absolutely understanding the contribution of immigrants to productivity growth in the economy is something that I think is an A+ research question. If some of the students came to me, and I was their dissertation advisor I'd be like, 'yes, that's a great question. I think that would publish well in a good journal, it's something academics would really appreciate.' But as I read CBO's paper, I realized that to CBO this is a question of really direct practical import. So, they get a request from someone, what would happen if we uncapped green cards for STEM PHDs, and these are people that are going to move to the US. They are going to patent, start firms, and increase the growth rate of total factor productivity in the economy. That's going to have federal budgetary impacts that they need to quantify and measure very directly. In CBO's disclosure of interest in this question they focused attention on a paper by Giovanni Perry, which is a state your panel version of thinking about TFP. I want to be clear, it's a nice paper, but for the reasons that I talked about, I think that conceptual approach is misguided if what the objective interest is total US factor productivity growth. Because it's just not going to be captured on a state your panel. Much more conceptually appropriate, in my view, would be a more macro approach. Maria Prato's recent paper is a really nice example. But mostly when I started thinking about this, I went to everybody I know who works on immigration and innovation. And I think it's fair to say that they mostly had feelings similar to mine, which is that it's quite embarrassing, for me, that I had never thought of that question. I had never taught that question to my students. I had never encouraged people to work on that, even though it's something that absolutely is very fundamental to where I feel like curiosity driven research should have guided me.

The combination in parallel of reading the CBO disclosed request for research and working with Doug Elmendorf for this budget scoring of high skilled immigration, you immediately run into the same problem, which is that for either of those, you need some data. That data really doesn't exist right now. What that means is that, when policymakers want to debate questions about STEM green cards, we really just don't have any basic facts for them to use as the basis for how many people would be affected.

Just to give some examples. Of new STEM PHDs earned in the US, what share leave the US versus initially work on temporary visas versus immediately transition to permanent residency status. Of those who leave the US initially, what share ever come back? Of those people who initially work on temporary visas, what share stay and eventually transition to have legal permanent residency status? And how long does that take? Of course, that data would be useful for a lot of people, other than just CBO. To give one example that I've been thinking more about recently. The national institutes of health set their statutes for eligibility for postdocs in the 1970s when foreign nationals were a tiny share of the students getting trained in the US, but the statutes haven't been modified. So, a lot of NIH funded postdocs are not eligible because eligibility is restricted to US citizens. Those kinds of data could also inform the costs and benefits of those eligibility restrictions for places like NIH. More broadly for researchers this data could be used to develop a much better understanding of the fundamental role that immigrants play in patenting, entrepreneurship, immigration, and growth in the economy.

The good news is that, in recent years, both Census and Treasury have made a lot is of progress in starting to compile data sets that measure a lot of aspects of that. The census is, business dynamics statistic of innovative firms project is one example. Those do lay the groundwork for starting to tabulate those kinds of statistics, but the key thing that they are missing, is that they have never had data on temporary visas. They are just fundamentally limited in their ability to speak to that kind of question. Which students in the US are on F1 visas? Which researchers in the US are on J1 visas? Which STEM PHDs are employed on H1 visas? Those kinds of data reside at departments like the department of homeland security, department of labor and department of state. Each in their own proprietary data systems, but in principle, you can imagine linking those to the census treasury data to have an integrated way to measure and estimate some of these questions.

If you ask most economists about the feasibility of doing those kinds of cross agency data linkages, they would look at you like 'I could spend some better five years of my life doing something else.' Those cross-agency data sharing MOUs are especially hard when they are proposed by outside researchers rather than by someone internal to the agency for a project that is explicitly of interest to one or both of the agencies. Luckily for me, a mutual friend introduced me to Amy Nice, who is an immigration lawyer, who spent her 35-year career essentially working in a bunch of those agencies and other key institutions in the space. Including the office of general counsel at the US Department of homeland security and most recently, leading work on science and technology workforce issues at the White House office of science and technology policy. What I realized, shortly after meeting Amy, is that relative to a lot of people, even though she's not a researcher -she reads our research, she cares about our papers, she knows her favorite economist by name, by having read their papers- but she recognizes research opportunities where other people basically just see that this area is kind of hopeless. When people ask questions to which there aren't any good answers, we can either choose to discard those as unanswerable. Or we can think it's in society's long run interest if we kind of make an investment at this point to try to create better data or create research so that the next time someone asks this question, we are going to have better answers.

When CBO publicly disclosed this interest in the immigration productivity question, Amy and I started working with Nathan Goldschlag, who is a principal economist at census on exploring whether those data linkages can be done in principle, and started conversations with the relevant perspective federal agencies about whether they can be pursued in practice. If that linkage is successful, then Amy, Nathan and I are going to collaborate on a disclosure, essentially describing first how to use the data correctly, because as it turns out, this temporary visa data is complicated. Thankfully Amy is a lawyer. So we will get the details on that more correct than if Nathan and I were doing that on our own, also disclosing some initial descriptive tables that I hope would be useful to CBO and others that are just data hungry for facts to inform current decisions that need to be made on the ground.

Nathan, Bill Curran and I are also going to use that data to reapply some of the quasi-experimental data from Bill's past work on H1B expansions that he had with Bill Lincoln. In the census you can combine that with standard productivity metrics that people construct all the time so that we can basically try to take care of the focus on the 10 year budget window and loop back to try to answer the question that CBO disclosed, that motivated this whole line of thinking in the 1st place, in a macro framework that -unlike the Perry paper- would appropriately account for the relevant effects. Even more importantly, one thing that's really rewarding for me about working with census on this, is that those links census reference records can go to the FSRDC. There's a really nice, recent paper from a student of mine and his PHD student which suggests that access to new data in the FSRDCs can really meaningfully shift the direction of scientific research in different topics going forward.

I'm describing that work with the census data as research work, which of course it is. And we are trying to publish those papers in standard peer reviewed journals, but I do want to be clear that my motivation in doing all of this work is to provide better data and better numbers to CBO to improve the accuracy of their scoring methods for immigration proposals. As I mentioned, Doug and I are collaborating with the Penn Wharton budget model team and essentially doing what's a public use version of what congress asks CBO to do, which is to score the types of proposals that I see coming up in congress.

The question that we're starting with is, if we uncapped green cards for STEM advanced degree holders, what would the implications be for the federal budget? I just never thought about budgets scoring before we started this. It turns out no one gives you a list of 'here's the channels for federal budgetary impacts', you just have to reflect, and it turns out me reflecting -not an immigration lawyer and I don't work on immigration- that was a very challenging problem. In our case, we're modeling 2 broad channels for federal budgetary impacts. The first is those people that will be in the US who wouldn't be otherwise. So that includes new people that come to the US who either wouldn't have come or would have been delayed in arriving. Second, a much larger group is that that there's would people have stayed in the US on temporary visas who we will now stay with permanent resident status, which changes their eligibility for public benefits. Which you also need to account for in terms of the federal government impacts. That's the high level. Here's two broad channels. Each of those has 15 or 16 cases, it's like every other day we think of a new case of some spousal eligibility pathway that we haven't really remembered. The economics of those cases is quite interesting. It raises a lot of elasticities that I hadn't really thought about. Just to give an example, is there a discouragement effect that people don't apply for a green card because the wait time is so long in India? Could you estimate that using cross country variation of the wait times and think about actually quantifying the discouragement effect in a way that we could take into account in the modeling. Just to give another example, you need to account for the fact that there's huge backlogs of individuals with approved green card petitions, especially in India and to a lesser extent China that are waiting to enter the US as a one-time stock adjustment. I think often times when people start thinking about these STEM green card proposals they just think, 'I just need to tabulate new degree holders by year as a flow.' No, there's a huge stock of people that you need to account for, in addition to these discouragement effects that I mentioned.

So, I'm part of a team doing that modeling work, independent of Penn Wharton. We are going to provide Penn Wharton with our best estimates of these policy counterfactuals in a firm that will also make available to CBO in case our work can be useful to them. The team that we pulled together for that modeling work is essentially unconstrained, in the sense that, you can think of who's your wish list of immigration lawyers and immigration economists and I'm willing to ask anyone to collaborate with me.

You can get all of the subject matter expertise that you can have on that topic. In addition to being essentially unconstrained on subject matter expertise, unlike CBO, we also have the luxury of time. We are not trying to do a budget score in the middle of the debt ceiling negotiation. We can take as long as we need to try to do a good job. And yet, even with an unconstrained team and an unconstrained amount of time, I have found this to be incredibly intellectually challenging and just really rewarding, but a very interesting hard problem to work on.

I came into this work with a really deep respect for CBO and the work that they do, but the process of actually trying to generate a score ourselves is really redoubled my appreciation for how hard their job is. I've always found CBO to be very thoughtful and humble with how they discuss what they often describe as their misses. If you want a recent example, Chapin White, the CBO director of health analysis, gave a Congressional testimony to the Senate budget committee a few weeks ago. At the end, Chapin shared what to me, came across as a very genuine, humble, and thoughtful discussion of how CBO thinks very deeply about their misses and what to learn from them. Which I thought was just great and lead to him having a really thoughtful exchange with senator Tim Kaine on that topic. CBO often gets criticized for what they get wrong. CBO is going to get things wrong in the same way that I and our team are going to get things wrong with the immigration budget scoring case study. It's hard to predict the future, even if, in our case, time and people are essentially unconstrained in the sense that I'm shameless in being willing to ask literally any outside expert that I can arm twist into collaborating with us.

When those misses are big, they matter, because it means the policy debates in Congress take place based on inaccurate information. As an institution, I think CBO and its staff do an exceptional job of learning from those inevitable misses. But what I think is on us as researchers, is that we can contribute to CBO having less misses by developing better data and better empirical studies that can inform them on how to be able to do more accurate budget scoring models.

Of course, not all academic research should be driven by what policymakers need to know. I think there's an incredibly important role for basic creativity driven research. My concern is that the structure of our profession often shields us from discovering some of these questions were economics research that's use inspired can be quite important. In this case, I would conjecture, that this entire Overton window for of feasibility for high skilled immigration reform could easily have been entirely missed by the economics profession. Partly, there's a very specific structural reason for that, which is that the key reforms in this space are being pushed forward from that national security community, as science and engineering workforce needs, rather than as immigration reform. Which is something that economists tend to be a little bit more attentive to. Unfortunately, outside of interactions in very specific places, such as the economics of national security seminar that Marty Feldstein used to teach at Harvard -which I sat in for five years of 'Tuesday Valentine's Day dinners with Marty Feldstein.' Or Stanford's Hoover Institution where they do have some interaction between the economists and national security institution. In general, we as economists don't talk very much with people who work on national security. It's just not an area of focus for us. Under the assumption that that's true, that highlights the value of economists not just working in traditional roles in DC, where we as economists get together and spend time together. But rather, we as economists also need to be more proactive about trying to listen to and learn from non-economists, who are engaged in the practical day to day work of legislative and executive reform efforts. Even if they are happening at institutions like the National Security Council, which are generally not perceived as being in our wheelhouse.

You could, of course, think this is just immigration, but when I worked through CBO's draft for the Journal of Economic Perspectives a few weeks ago, the first question they posed from their energy and environmental team actually struck me as being an even more extreme version of economist and economic research almost totally missing the boat on the Overton window for a different example, which is permitting reform. Thankfully Doug Elmendorf humored me asking him to work with him on a similar project and we're teaming up with Zach from Yale Law, who I think is one of the most thoughtful researchers on the topic.

Just to close, economics by construction obviously have academic freedom to choose to do whatever they would like in terms of their work. Speaking for myself, I'm much more motivated to pursue work on these kinds of use inspired research questions for which I know there's an audience, be it CBO or NIH or others, that would see my findings as relevant to their work. I would conjecture that creating more of a home for that kind of work in academia could also attract a wider set of students to want to train and do research in the field of economics, and to apply their economics training to a variety of career paths that are high social impact, such as government service careers.
Thank you.

[]We will now go on break until 3:50, thank you all.

[] (No sound) A&M University and the title is, Selling Subscriptions.

[Ben Klopack] OK, so thanks so much for including our paper in this terrific program. This is joint work with Liran Einav and Neil Mahoney at Stanford.

This paper is about subscription products. For the purposes of this paper is what we're thinking of is, services where consumers sign up initially, they get a flow of goods or services, and they sort of pay on a regular interval and that kind of subscription service continues until the consumer actively decides to cancel. So the key element for us is the autorenewal piece. To fix ideas, here is a list of 21 of what this industry report considers the top subscription categories in the United States. What you can see here is that there is a set of products which perhaps are legacy products that were always in a subscription, things like newspapers or perhaps mobile phone service. There is also this other set of products, which might have traditionally been sold in a pay as you go model. So, I used to go to the pharmacy and buy razor blades. Now I can instead get them delivered to my door with dollar shave club. There were similar examples in the clothing category or in pet food, and a variety of other things in which we traditionally bought these products in a pay you go to model, but now maybe we buy them as part of a subscription. So, there has been an expansion of this model into other categories.

By some measures, the subscription economy is growing rapidly. Between 2012 and 2020, the sales of subscription product grew roughly 3 times as fast, or a little more than three times as fast, as overall retail sales and this trajectory has continued as we extend this line path to 2020. Lots of growth in this model. This paper is really about exploring one possible rationale for the growth of this subscription model. To be clear, we think that there are multiple possible explanations for this growth. One thing could just be that, as the set of things that people buy has shifted towards digital products, these perhaps lend themselves naturally to a subscription model. Think about a streaming service, like Netflix, where the marginal cost of delivery of additional quantity is low. This might just naturally be well suited to subscription model.
We also think there's a role for an increased demand for convenience. If I subscribe to something on a recurring basis, I save this transaction cost every month if I'm going to buy the same product anyway. This auto renewal is potentially convenient. Indeed, you ask consumers why they sign up for subscription products, many of them say that it feels nice to see receive them something every month without physically doing the entering of the payment information online.

But today's talk is going to be about a different, less benign potential explanation that we think could be possibly amplifying this trend. Which is, that if consumers are inertial or inattentive, and they sign up for a subscription and then don't remember to cancel even after their evaluation for that service has dropped or if there are potential hassle costs associated with canceling a subscription. This could potentially increase revenue for firms relative to selling that same product in a pay as you go model. Indeed, if you ask consumers in surveys to reveal how much they spend on subscription products, their initial estimate is wrong by a factor of 3. They say that they spend $80.00 a month and then when they go through their credit card bill, they find that they actually spend $250 a month. This is usually the part the in talk where everybody starts looking at their phone and trying to cancel a subscription that they've forgotten about for the last year. That's really the economic phenomenon that this paper is about.

So, what do we do? We explore the effect of this inattention on firm revenues for subscription companies. We are going to focus the margin of subscription cancellations. To do this, we are going to use a large transaction data set from a major credit card network. Our empirical analysis is going to exploit a key feature of this data, which is that we can observe when a consumers credit card is replaced, because it expires or it is lost or stolen. When a consumer's credit card is replaced, typically, -let's say Netflix- requires the consumer to go in and update their payment information in order to continue the subscription. We are going to think of this as a shock to consumer attention. In most months consumers subscription is auto renewed in a passive way. But in the month that the card is replaced, this triggers an active decision where the consumer decides 'do I actually want to continue this subscription.' That's going to be the key variation of our empirical exercise. We are going to show you that there is a sharp and discontinuous drop in consumer retention rates, right around this month that a consumers credit card is replaced. We're going to take this data pattern and embed it in a fairly stylized model of subscriber behavior and use that to think about what would happen if consumers were fully attentive. We are going to show you that this has a large impact on firm revenues, but there's huge heterogeneity across different services. Then we are going to use the model to think about the impact of possible regulatory remedies that we think are in the spirit of some things which have been proposed by the FTC.

First, I will talk about the data. The data comes from a major credit card network. Our sample goes back to the middle of 2017, but this indicator for whether we can see a card that is replaced as part of the same account is only going to be available from the middle of 2018, so we are really going to look at 2018 to 2021. You should think of every row in this data as being akin to what appears on your monthly credit card statements, so we are going to basically see which store you went to and some information about the store, the store category, and also the transaction date, and we are going to see the dollar amount. But we are not going to see exactly what you bought. This means if you made a purchase on Amazon, we're not going to be able to tell whether you're paying for Amazon Prime or whether you were making another purchase on Amazon. Then, crucially, for identification, if there are multiple credit cards with different numbers that belong to the same account, we are going to be able to link those and we are going to use that as this key variation in the empirical exercise.

We are going to create a set of subscription services that we analyze. As a starting point we're going to use the table that I showed you on the first slide. It has these 21 categories of subscription services. And we are going to take the services on that slide, as well as any kind of additional ones that we were able to generate from industry reports, that had high market shares. That's going to generate a candidate list of 69 possible subscription services. Then we are going to exclude a bunch of them for various reasons. There are 12 that we're not able to find in the credit card data. There are 31 that are fairly small, so they have fewer than 500,000 monthly transactions on average. There are 4 services that are in categories that are typically associated with long term contracts where consumers are contractually obligated to continue those subscriptions, things like cell phone plans or ISPs. There are 6 services associated with both subscription and non-subscription transactions. This is Amazon prime example. If I bill my Amazon Prime subscription or just make a purchase, we're not going to be able to directly distinguish those transactions so we will drop services like that have a that. There are 2 services that have a short average subscription link length, less than six months, 2 services with recent launches and 2 services with non-monthly billing. After dropping those that will lead us the total sample of 10 relatively large subscription services that range across a broad array of different product categories. They include both digital and non-digital goods. We have newspapers, security firms, entertainment, and streaming services, etc. Unfortunately, as part of the NDA, I'm not allowed to reveal what services we have, so I'm going to refer to them as letters. So, you can spend the rest of the time guessing which service is which.

With those ten services, we are going to construct these monthly survival rates of subscribers. Here is what we do. For every service, we will construct a set of cohorts of subscribers and every subscriber is going to have their credit card replaced at some point during the sample. We're going to follow every subscriber in the cohort for 25 months from the date of their initial subscription and every cohort will be defined by two things. One is the month at which they first subscribe to this particular service and the second is, the length between the initial subscription date and when their card was replaced. We're going to focus on cards that were replaced exactly 6, 12 or 18 months after the date of the initial subscription. This is not for any particular reason except it looks nice on the graph. We are going to require accounts to be active in each of the 25 months after the date of initial subscription. This means they had at least one transaction. So, we will drop accounts that become inactive. There's some data cleaning that we do, if a card misses one month of a subscription, we will fill in single month hole, we're going to assume that they continue to stay subscribed in that month hole. But if they miss two or more months, we will consider them unsubscribed and drop them or think of them as lapsing from this cohort of subscribers. Then we see a large drop in retention rates right after the first month, so there are many cards that signed up for just one month. So, we will condition on cards that stay for the first two months and consider this first month as a trial period. Then we're going to aggregate across cohorts for each of these 10 services that we have, which will yield these monthly survival rates for every service and every cohort.

Let me show you what those look like. This is service 'A' and these are cards that are replaced exactly 12 months after the date of initial subscription. Each of these lines is a cohort and you can see that in the months before the card is replaced, subscribers are dropping off at a relatively smooth rate, but then right in the month in which the card is replaced there's a discrete jump. For this particular subscription service, about 25 percentage point drop in this kind of survival rate, around this month 12. This is going to be the main data moment that's going to inform our analysis. We aggregate this across cohorts for service A and this is what the aggregated survival rate looks like. You might notice that this discrete drop happens over two months and the reason for that is related to the timing of how we think about card replacement. Imagine your Netflix subscription is initially billed on the 6th of every month and your card is replaced on the 12^{th}. In month 12, the billing is going to be successful, but in month 13, Netflix is going to try to run your card and it's going to be unsuccessful and that will look like a lapse. If those dates are reversed, as in, your card is replaced on the 6^{th}, but your Netflix is billed on the 12^{th}, then that will happen in month 12. Here's the same graph, layering on the set of cards that are replaced at exactly 6 or exactly 18 months. You can see there's this discrete drop around the time of card replacement and a remarkably stable slope in the non-replacement months. One thing you might be worried about -or at least that we were worried about- is that when a card is replaced, whether it's lost, stolen or expired, what consumers are doing is moving their transactions on to some other card. The credit card people call this being 'top of wallet', meaning I have some card in my wallet which I always reach for to do transactions and if I am missing that card, I might switch to another card. One thing we can look at, is to look at transaction activity around this replacement date. Here the lines are all stacked around the date of replacement, which happens at month zero. And you can see that the transaction activity drops in the month of replacements and also in that next month because of this messiness in how we observe the timing of things, but it recovers almost to the original level. There's a permanent two or three percent change in the number of transactions per month and amount of spending per month, but we don't see big effects in card activity and part of this is because we are conditioning on cards that remain active during the whole period. But we don't see these big shifts.

I showed you these graphs for service A and these are other 9 services, B through J. You can see there's a huge heterogeneity in these graphs which reflects, presumably underlines things about the service. Two graphs to maybe call out are our merchant I, which has a relatively flat slope in this decay rate of the consumers who stay subscribed and has this huge drop right at the time of expiration. In the model we are going to think of this drop as being evidence that consumers for service I are extremely inattentive. Because they stay subscribed at high rates, but in the month they are replaced, lots of them drop off. We are going to interpret this as evidence that there's lots of subscribers of merchant I that didn't want to be subscribed, but they either forgot or merchant I did things to make it difficult for these consumers to cancel. Contrast that with something like merchant G, where there's almost no drop at the time of replacement. We're going to think of -in the model- we'll interpret this as consumers who are relatively attentive. In fact, there are relatively few consumers who have lost track of the subscription to merchant G Then, of course, lots of cases in between. This heterogeneity in the counterfactuals will translate into very different effects of this inattention on firm revenues. Some firms will benefit hugely from inattention and other firms are going to benefit relatively little because they have pretty attentive subscribers to begin with.

Then we'll take these retention rates and fit a pretty stylized model that actually fits the data quite well. We're going to think about a model of subscriber behavior. Every subscription service charges some monthly price, p, and a subscriber gets some flow utility from being subscribed, which follows a Markov process. That utility is going to be denoted as Uit. In order to have signed up in the 1st place, a cohort of subscribers must have had a positive evaluation in that 1st month that exceeded the monthly price. So, in that first month, a cohort of new subscribers is going to have a distribution of initial utilities, which is strictly greater than p, which we will call Ui0. We're going to assume that conditional on subscribing, consumers are myopic. This will rule out behavior like I'm attentive this month, I know I'll be inattentive in future months, so I canceled this month that I'm attentive. We won't be focused on this sign-up piece, so we won't think about whether consumers don't sign up because they know they're inattentive. Actually, Avner and co-authors have a nice paper that shows that that margin on sign up is also important. But that's not something we'll study here. In a given period, subscribers will be probabilistically attentive, so if they are inattentive they will automatically renew, because the subscription service will passively bill them. If they are attentive, they will cancel if their monthly evaluation has fallen below the price. Every month, consumers are attentive with a rate lambda. That will be a key parameter when we think about counterfactual. When the card is replaced, we'll assume that consumers are attentive with probability one. In the paper we do some robustness exercises to think about what happens when consumers are in fact, not fully attentive at the month of card replacement and we showed the robustness of our revenue numbers to those assumptions.

We're going to parameterize the model as the flow utility that I get in each month, net of the price, as an AR1 process with a parameter row and a normal innovation shock. The initial distribution evaluations in the first period of the utility net of the price, we'll assume it's distributed exponentially. So, it is strictly positive. Then we'll estimate this model service by service. And we will recover three parameters for every service. Lambda, how attentive my consumers are, what share of my consumers are attentive in each month. Ada, which is what is the distribution of my consumers initial utilities. Then Rho, which is how quickly are my consumers initial utilities decaying.

Here are our estimates. And we estimate Rho to be fairly close to one for most services. Lambda, which is an important parameter for us, has huge heterogeneity across services, which you might expect, given the heterogeneity in the survival graphs that I showed. The service with the least attentive consumers is service I for which only 4% of subscribers are attentive in each month. And for service G, the graph that I showed you, which had almost no drop at the month of replacement. Roughly half of service G's subscribers are attentive in a given month. Wide variation in our estimates of Ada, which describe how much do consumers like the service in the first month. That together with Rho is going to generate how quickly our consumers are dropping off in non-replacement months. Here's the fit of the model relative to the data for service A. When we fit this with a method of simulated moments, we'll match the retention rates in the months except for the replacement month and the following month. There is nothing in the model to inform the fact that we don't have this timing exactly right, so we won't match on month say 12 and 13. The model in general fits the data quite well, given how stylized it is. Here's the model fit for the other graphs, calling out again, service I and G.

Next, we will take the model and go to thinking about counterfactuals. There will be two main sets of counterfactuals, the 1st is really more of a quantification exercise to think about what is the impact of inattention for revenues and it's something we're going to do service by service. What we do is we simulate a set of consumers and then simulate their subscription activity for 120 months after the date of initial subscription. For the simulation exercise we're going to assume that cards are never replaced, so if I'm inattentive, I will be probabilistically inattentive in every month and I'm never going to get this attention shock as in the empirical exercise. For each of these services, the first column, the share unaffected column is the share of subscribers who are never marginal, because they always have positive evaluations. Remember, we have this Rho parameter which is close to 1 and a positive initial evaluation. So, if I get this big initial shock, it is quite likely that I'm just going to stay positive forever. And I'm going to stay subscribed throughout this whole 10-year period. This fits what we see in the data, which is that many services have very high subscription rates, even after our 25 months of sample.

For the remaining consumers, we'll calculate the average number of months that they stay subscribed under two different scenarios, one of which is the scenario in which consumers are probabilistically attentive at rate Lambda and the other is if they were fully attentive in every month. When consumers are fully attentive, they will cancel as soon as their evaluation drops below the price. When they're inattentive, they're only going to be able to cancel if their evaluation becomes negative, they're not going to be able cancel until they have a month in which they are attentive. By construction, there's a higher number of months that I stay subscribed and if I am inattentive then if I am attentive. Again, there's large heterogeneity across services here. Sticking with merchant G and I, G has only about only a 10% or 20% drop off in average months subscribed if consumers are fully attentive and for merchant I, there's an enormous effect in how long subscribers are sticking around if they were fully attentive. Then we will combine the share unaffected column and the average month subscribed, which is conditional on me being marginal, and compute what we're going to call the revenue ratio. This is the ratio of my revenue when consumers are inattentive to my revenue if consumers are fully attentive. Huge variation across merchants. Merchant G is getting only a 14% revenue bump from inattention and merchant I is getting more than 100% bump from inattention and merchant B is getting an even larger bump from inattention. On average, among these ten services, the average service is nearly doubling their revenue, getting a + 90% effect on their revenue because of this inattention. We discount our future revenues. This makes relatively little difference for the counterfactuals if we discount or change the discount rate.

Then we think about a policy wood forest an active choice every X number of months. We think of this as being in the spirit of recent FTC policy guidance on negative options. You can think of this policy as becoming an opt out of your subscription activity after a certain amount of time. This is not free, it's associated with lessening of this convenience, which is presumably one of the reasons that we see people go to these subscriptions in the first place. What you can see is that if I make people opt in every single month, I'm at the fully attentive counterfactual. And if I never make people opt in, I'm at the inattentive counterfactual where I'm probabilistically attentive with rate lambda. On the left panel is the revenue ratio, the ratio if I am inattentive but reminded after every X number of months relative to the fully attentive version. It's hard to see perhaps from this graph, but if you wanted to reduce the revenue ratio for the average merchant by half, you would have to remind consumers every six months. Force them to opt in every six months. You can get most of the benefits of this policy from a six- or three-month duration. Huge heterogeneity across services, because of the variability in how much this inattention impacts firm profits. We don't do anything here with consumer utility and I think that's a promising avenue for future work, but we think of this revenue impact as in some ways being closely related to consumer revenue. Presumably these consumers who are staying subscribed when they're attentive, they're being heard in some way by this inattention.

This inattention could also presumably vary a lot across consumers. We try a couple of heterogeneity exercises and one thing which we found with what I think are somewhat striking results is to break up consumers into those who took a cash advance on their credit card versus those who didn't. Taking a cash advance on your credit card is a relatively expensive way to borrow money. You pay an initial fee to pay to pull the cash out and you begin to accrue interest immediately, as opposed to your credit card statement where you have a grace period of a month. You usually pay interest at a higher rate than your typical credit card APR. When we do this and estimate the model separately for cards that took cash advances versus those that didn't, we find that consumers who didn't take a cash advance tend to be more attentive than those who took cash advances. We think of this as a proxy for some type of financial sophistication, although it could also be related to credit constraints and other things. We think of this as suggestive evidence that there is scope for possible distributional effects of these types of inattention policies. This would also inform the sort of policy discussion in the FTC and other places about things like simplifying cancellation procedures or perhaps kind of reminding people about their subscriptions, making these auto opt in and opt out pieces more clear. If we also compute the revenue ratio, we find that the revenue ratio is much higher for consumers who took the cash advance, so that's what the plot on the right is showing. Each dot is a service and the 45-degree line corresponds with no difference between cash advance and non-cash advance on the right the panel. The dots aligned above the 45-degree line are services where the cash advance consumers are generating higher revenue from inattention.
With that, I will wrap up. We show evidence from 10 subscription services that consumer inattention appears to be an important thing affecting the subscription cancellation margin. But the impact of inattention varies a lot across services. You can imagine without me telling you who the merchants are, that if I sign up, for example, a meal delivery service and forget, the groceries are piling up on my doorstep versus if I sign up for a streaming service that I never use. So, we think there's a lot of natural variation in what these services are that's generating this heterogeneity. The implication for that that is the revenue impacts are also highly heterogeneous across services. We also show evidence that some type of intermediate policy like reminding consumers that they are subscribed perhaps opting them out if they're inactive could have some usefulness, but we note that they should be traded off against the

convenience benefits of subscriptions, but we think there's potentially scope for some intermediate policy. I'll end there.
[Applause]

[] Speaking of the discussant, you are fortunate to have Avner Strulov-Shlain from the University of Chicago of the school of business to discuss.

[Avner Strulov-Shlain] Let's see if I can get there. So many appendix lines. Look at those figures.
This is not going to be a review of your discussion, I'm a big fan of the paper. I like it so much that I saw Liran present it, I saw Neil presenting it, now I saw in Ben, so I'll collect my prize money later.
I'm taking the liberty of musing here and looking at the big picture of the paper.

In one sentence, this is a a very clever, clean design. The data is great, look at those figures and see the results. You don't need any fancy metrics. Yet the simple model fits the data and you can learn a lot from it. One thing I've heard in previous presentations is presenters being asked what explains the heterogeneity between the services. The paper doesn't go into those differences and I think rightly so, they only had 10 services with very different clients. We're not going to learn much from it, so I think that's just -there's a negative reviewer request, do not do something.

The big take away is that many, in some services most existing subscribers are inert, which means that the same person who has been a subscriber for months will not resubscribe even if it were for free if their subscription is cancelled. It's not exactly what happens here, because it is costly to resubscribe and you need to reenter your credit card details, but it is very close. What it means is that much, and in some cases most, of sellers' revenue, as they show us, is from consumers who actually don't want to be subscribers and that's not a good thing. It's not exactly a welfare statement because you know, if I'm not using Netflix at all, then I'm just losing -I don't know how much Netflix cost- but probably $15 a month or something, but if I'm watching only one show and that is worth $12 a month, and I'm still forgetting to cancel, and the welfare loss is likely smaller. It implies some loss. What you do about it, depends on why consumers do not cancel and the approach that the paper takes, which I think is a very plausible one, is inattention. You can't explain almost all the results also with switching costs and those are the main two reasons for why we observe this kind of initial behavior.

Switching cost story is that it's just going to be constantly to take an action, I need to find the cancel page, I need to click a button and maybe I need to call customer service, maybe I need to send a fax or a homing pigeon or something. That's going to be a lot of trouble. And that's why people don't cancel and kind of generating those patterns, you can also think of lambda as, in a stochastic cost model, that will generate this gradual decline and then the non-renewal after a while. So, really teasing apart that against inattention or inaction where there's just this- So switching costs are coming from this cost benefit analysis that is it worth it for me to devote the effort to cancel or is it good enough or not too bad that I'm willing to stay subscribed versus inattention, which is more I think of it as more inherent that I just forgot, or I never looked at the price of Netflix even though I've been subscribed for a few years and it was 8 and now it's 15, but I had no idea. What happens in practice is that most papers will assume one channel or another, because they can just speak to which one it is. I want to say that there are a few papers Brat-Goldberg et al, looking at Medicare part D, drug prices and found strong evidence for inattention to switching costs. Miller et al and the asterisk is Paul Simon saying ‘you can call me Al.’ We also find support for inaction over switching costs, but it's not the cleanest. We all know what a big hassle it is to cancel services. It's kind of tricky knowing which is which, and policy can address both. What the impact will be in general equilibrium will be very different. What the impact will be on welfare will be different. But we can think of both. An assessment can be made, but it's tricky and requires some effort into separating those different channels.

One thing I wanted to touch on, is the importance of the market structure and the competition and how inertia inducing different services are. To the extent that consumers are sophisticated, that they know that they’re going to be locked in on Amazon. But they have other options, so they can be on eBay without a subscription double -I don't know if that's true. That will lead this paper to give us an upper bound on the revenue effects, because we're losing all those consumers that will not sign up for Amazon in the first place. It will also underestimate the benefits to consumers from safeguards, because by eliminating or reducing the level of inertia, whether through switching costs or increasing attention, more consumers will take up the service because it's less risky. My risk of getting stuck and not being able to cancel when I want to is going to be lower. That means that theoretically, it might even flip the revenue effect, which should be good news for a regulator. Because if we think consumers are being heard, then businesses will not oppose it as much, because it's not as bad for their business and that should be easier to do as a policy.

Now I want to push track on myself a bit. Well, if that's true, why do we see so much exploitation? Why don't businesses do it on their own? Here I answer myself. Just an aside, as my view on firms’ decision making, I want a stamp of question mark max pie if I had reviewed papers in prints, that's what I would put on it. I don't think companies, necessarily, always do the best things. Now, our paper will find that there's revenue equivalents between exploiting this inertia and not, in some particular way. I think a lot of it is just coming from the firm's initial behavior of the industry knowing that this is how you run a subscription service. You get people through the door, and then you just keep them in and you increase the prices gradually and worsen the terms of service, even though it might not be the best thing in the long run. In our paper, what we see is the effects on revenue are very different in the short term and long run. In the short run, there's locking in, or exploiting, those initial consumers is very lucrative. Because they do stay for those few extra months, as Ben just showed. In the long run, it might backfire because of this extensive margin of consumers who know they will be locked in and won't join in the first place.
You need to be a pretty savvy firm to realize those two margins. You need to actually think about it and measure it. And usually, you don't do that kind of measurement of asking what happens if I change my business model from one edge to another. I do want to know that we do see newer actors move in the direction of more transparency, easier to cancel. So maybe they do see some value there that the legacy actors do not see.

Just in the last minute or so. What are -I can allow myself to be less tied to actual research right now- so what can be done? What the paper does is assess the effects of forced attention. What will happen if we

make people decide every month, 3, 6, or never? We can think of the mandatory reminders. Some companies do this voluntarily and that's nice. We can think about auto cancellation or reminders due to inaction. Those target the inattention channels. If it is cost that leads to the locking initial behavior, then we can think about the FTCs click to cancel policies. Some of those were already implemented in the EU and in Britain. Other third-party services do that as well. I just want to suggest another idea, which is to think of subscriptions as a business model. I don't think it's a regulatory relevant concept perhaps. But think of running the subscriptions as a soft optic. So, I pay every month, but if I don't use the service this month, I get reimbursed all of it or part of it or something. And then I get the convenience of the subscription with lower risk of overpaying. Maybe everyone is happy. With this optimistic message, I will conclude here. Thank you.
[Applause]

[Eddie] Okay, great. Do you want to take a couple questions?

[Q&A]
[Q: Matt Leisten] Matt Leisten, FTC. I've seen this paper twice and I think that this is an epsilon neighborhood of a perfect paper. In the counterfactuals, when you force attention, for instance, as I understand, prices don't change. But you could probably assume that maybe these things are 0 marginal cost. And you could think of the supply side response when it's easier to cancel. I also thought that, well maybe not as policy relevant, you could also compare the counterfactual to the old model of you buy office or whatever and then you charge a price based on the perpetuity, net present value of this flow of utilities that you get every month and that would be a useful benchmark as well.

[A: Ben Klopack] Thanks, we thought a little about the pricing, but I think there's more to do potentially in a counterfactual. And also, going to Avner's great discussion by the way, thinking about the competitive effects would be useful as well, so thanks.

[Q: Jonathan Hawkins] Hi. Jonathan Hawkins, CFPB. In other markets like mortgage refinancing, we think about how people who are inattentive maybe subsidizing people who are attentive. And so, if you force attention, the inattentive people are better off and the attentive people are worse off. And there might be distributional implications. Have you thought about that at all in this model?

[A: Ben Klopack] We are getting at that in a very reduced form sort of way with this breaking consumers into groups based on the cash advance stuff, but there could be a way to formalize it a bit more, so thanks.

[Q: Mike LeGower] Hey. Mike LeGower, from FTC. Can you observe in the data, individuals that have multiple subscriptions going on and can you use that as either a way to get at heterogeneity inattention within consumers. But also, I was thinking about it as a way to think about welfare effects in a ballpark sense of- do we know how many people have multiple subscriptions and how costly it would be for them to pay attention to all of the subscriptions that they have?

[A: Ben Klopack] We haven't thought about the welfare piece like that, but that's an interesting suggestion. But we have looked a little bit to see if consumers are switching between subscriptions at the time of their attention. We don't see a lot on that margin, but we also have been limiting to these particular timing assumptions, so there could be more to do if we relax some of those. Like only looking at cards that are replaced exactly 12 months after a subscription and stuff like that, so there may be way to expand the sample to look more at that.

[Q: Marc Rysman] Hey. Marc Rysman, Boston University. This is my first time seeing the paper. So, I feel really out of it, relative to these other people. I have a data question. Can you tell if people switch credit cards or switch to their debit card, if they switch to subscription over?

[A: Ben Klopack] Unfortunately, no. Not right now. In past iterations of the data, they had some ability to do that and we hope to be able to do that again in the future. The credit card firm sometimes buys data from a credit bureau that allows the linking of multiple cards to the same household, but in the current access to the data, we are not able to do that.

[Q: Sarah Webb] Hi, Sarah Webb, from the University of Maryland. Again, excellent paper. You spoke to it a little bit in the… What do you call it when they talk after? But is there some way that you could make an index of how difficult it is to cancel these subscriptions? Recently I tried to cancel a subscription with CVS and I had to call. 'Dear consumer, are you sure you want to lose the great value?' And it was like a 15-minute process. You might be able to decompose in some way what's the *[inaudible - microphone crackling*].

[A: Ben Klopack] We did some version of that. We had RAs sign up and canceled all of these subscriptions. In the cross section, to Avner's point as well, it's difficult to get much because there's so much heterogeneity, so we didn't see a clear correlation with our lambda parameter or something like that, but I think there are potentially other things we can do. I think there's been recent policy changes, for example, requiring firms to make cancellation more simple, which is sort of some high profile offenders have changed their cancellation procedure. If you look on Twitter for New York Times cancellation, there's lots of stories on- I think there's potentially more to do there, so thanks for the suggestion.
[Applause]
[End of Q&A]

[>>] I need to organize papers up here. Now Parker Rogers from Indiana university and NBC are discussing regulating the innovators approval costs and innovation in medical technologies.

[Parker Rogers] I think you might have the clicker. Is this right here? Thanks so much to the organizers for having me. A pleasure to be here and take part in this excellent conference. New technologies have the potential to dramatically improve our lives, but they can also cause harm. One way to mitigate harm is by regulating technologies through mandatory premarket testing. That's an approach taken by the US Food and drug administration, FDA. For more groundbreaking and high risk technologies, this is particularly compelling. This is shown in a recent paper. For establishing low risk technologies, the benefit of this regulation is less clear.

Critics of regulations say it shows innovation and competition through the expensive approval process while the proponents of regulations say that it is necessary for improving product safety. And it may actually lose innovation and market competition because it allows new homes at credible signal of product quality without having long standing reputation. The FDA is a focal point in this debate and we've seen this play out recently in the COVID-19 pandemic with regards to the safety first access and affordability of new COVID-19 technologies, but this debate is much broader than COVID the FDA regulates about $2.8 trillion worth of goods each year and in this project, I'll focus on the FDA's regulation of medical devices, for examples of these are things like x-ray machines, COVID-19 tests come I'll get into more examples as well as we move through this talk.

The question in this project is harvest fba medical device regulation effects innovation, market structure and product quality of these more well-established lower risk medical device types. A broader question. I think I'm speaking to is how does strict product regulation compared to lenient regulation. And then how does the regulation environment compare to the no regulation environment? No I'll answer these questions is by examining a set of depopulation for regulatory regimes that switch these extrinsity of regulation from class 3, a strict regulatory environment, what you can think of clinical trials, to class 2, which is a moderate regulatory environment.

We'll call that switch from three to two down regulation and we'll consider events that move medical device types of class 221, which class one is an environment where there's no approval process required and I will call that event type from 2 to 1 deregulation. (inaudible) Various degrees of technological sophistication come across these different products, but also how the different margins of regulation effects regulation and these different outcomes I'm looking at. My analysis of these events leads to the following core results double down first, related to innovation, I find that the number of patents filed in effective device types and the number of submitted devices to the FDA increases by about 200% after these regulatory shifts.

The value of this increased innovation is quite large when you extrapolate markets. It can be as large as $32 billion or 20% of the market. Our father's outsized increases alcohol and inexperienced firms. And these are the ones that face complexity and financing frictions with the approval process, I also find that the innovativeness of these technologies increases after you decrease the amount of regulation that these firms face. You see the innovativeness increases by 200% and that's as rated by patent citations and also patent valuations. A reason why this might be happening is because I see that the largest increase in innovation is coming from these small firms, which tend to be the most disruptive firms. On the marginal market structure, I find a tenfold increase in firm entry after down regulation and corresponding to this increase in firm entry for both down regulation and deregulation, I actually find that after deregulation, they dropped the price and the price is paid for medical procedures that use medical devices of about 25%. I don't see any change in the prices paying for medical procedures that use down

regulated technologies and I'll talk about why that might be later. These are all reasonable attitudes when you look at the literature that ties firm entry to prices, and then lastly, product safety --
I find that after you deregulate a product, you actually see that inventors are emphasizing product safety and empathic text much more readily, so an increase of about 100% on that margin. We see a decrease in the number of severe adverse events about 20% rapidly regulation as well unless even at the margin the of more dangerous device types before deregulation. The mechanism. I'll argue about why this is occurring is the fact that firms are being exposed to legal liability risk after deregulation. To provide some background on these policy changes that I'm analyzing in this project, the FDA regulates products and medical devices in three different classes and class three of those devices deemed high risk -- to take 54 months to approve the restrict process, like a clinical trial, it costs $75 million on average to navigate this process. When you receive FDA approval, you are not subject to any product liability related to product design defects comment because you've received the federal approval through double dashes because the -- of federal redemption and for class 2 devices, those are deemed as moderate risk, and they take ten months to clear, $24 million to navigate that clearance process. In the era I'm setting in this process, they are protected from legal liability risk after clearance and lastly, you have class one devices deemed as low risk, there's no approval or clearance required.

They just need to register the products with the FDA and pay a small fee. His device manufacturers are exposed to all legal liability risk, because they don't get federal approval. One thing I want to mention is the timing of the events can be considered as good as random and what I mean by that, for example, a example good of this is in 1994, the FDA decided publicly announce a down regulation of soft contact lenses and this took them ten years to finally internally bring this about. It takes quite a bit of fun to do these down regulation, deregulation events and for these costs two to one events, these deregulation events, this is basically based on a crude measure of product safety and there is some discretion involved, but I'm unable to observe that measure so I'll talk about that as well.

Another piece of background here is about the legal liability context for medical devices. Judge sweet of the US District Court said like an economist would, it must be recognized that state tort actions remain a powerful incentive for improving product safety. It's regal liability and miscontext / context is quite important how about litigation percent of annual medical device manufacturer revenues and has been shown to fill innovation. Examples of meaningful legal liability lawsuits that have occurred in this space are the $1 billion payout in 2014 for this Stryker hip implant and the $3.2 billion payout for the Dow Corning breast implants in 1998, and so interestingly I'm going to leverage some changes in the legal liability environment as well in this project to show you that legal liability is the things driving the improvements in product safety.

Before 1996, virtually every court held that cross three and cross two approval clearances protected from injury towards, so as you can see on the ground on the left, I'm showing the number of medical device cases in the light blue line, which is a large number of these cases. And arising over time, and in the dark green line, I'm sure the number of those cases that were preempted by federal preemption because of their FDA approved. After 1996, this preemptive protection of class 2 devices disappeared and you can see that the number of -- to share of those device cases that were created drops, while in orange I'm showing the number of pharmaceutical cases printed and those continue to rise over time, just to show you a plausible control group those were not affected by the Supreme Court case in 96 that altered the legal environment.

What I want you to take away from the slide is first come up legal liability matters and 2nd, the environment changes in a way that allows you to rule out exactly if liability is the reason why we are seeing (*inaudible*). The last bit of background that I'll provide you is that I think there's mutual innovation happening even in these well-established product categories. This is an instance of meaningful incremental innovation after down regulation. This is a US package soft contact lens, and the invention says it is a method of making a coded contact lens with desirable physiological performance. Miss catheter was violated upon the release at about $1 billion and it was important because it this contact warns us to retain shape while withholding moisture, which greatly enhance user comfort and which led to one of the top two brands on the market today, which is the Acuvue. What I want to say is gross innovation isn't necessarily radical but it's important because it bridges the gap between these groundbreaking technologies and consumer adoption, which is kind of what I'm focusing on today. I did it. I look at in this project, I could compile a rich set of data to comprehensively analyze the different margins that are important in the space.

First, I have FEMA institute of data combo execution data on average events related to product safety, device approval submissions to the FDA, which kind of gives me the innovation activity and also the firm's activity as far as which types of firms are entering and how long they've been navigating the FDA's regulatory processes, I have US Patent Office patent database, which gives me filed patents in medical device categories that are affected, I have also patent citations gives me me a patent or individual quality measure it allows me to be able to change an innovation direction and then also allows me to measure activity to new entrants in a particular space etcetera.

I had the market scan database come up. So, this is a large health claims database that covers about two decades. And this gives me the price of procedures across time for a large number of employer-sponsored health insurance groups. And I have CRSP Compustat data, which allows me to measure firm size and heterogeneity and I have the Kogan at all estimates. We talked about earlier today, which measure the patent market valuations, basically it attributes the value of a patent to the change in the patent assigning stock market price when the patent is announced. The methodology.

I use here is a stack difference in differences design and event studies. And it's also been mentioned earlier, and this confronts the staggered adoption bias that has risen recent econometrics literature. This compares 3 to 2 untreated groups to avoid inappropriate comparisons between impairing these different types of groups. I have 4 control groups I'm analyzing, (inaudible) also an intuitively similar device types, which treat similar diseases and indications, I have later treated device types, which are those that are double dash by the FDA beyond my sample window, and lastly, I have a match control group, which is basically the control group that is match based on baseline averages of the outcomes.
Importantly, I don't find any of the big differences in the outcomes estimates in that I have across all these control groups, and then I also see that there are no diverging pre-existing trends in the outcomes of

interest across -- all my outcomes. And then for both types of regulatory shifts that I'm analyzing. Want me to dive into the results today. I'll give you results of innovation, market structure and product safety. On the innovation front, how do these like practicing race and device submission rates? And then how do firm traits shape these innovation effects? Done regulation, so moving from high to moderate regulation, that boosts innovation and also deregulation seems to do so as well, but to a lesser extent.

On the bottom side. I'm showing an event study we're on the x axis, you have the years since the down regulation event occurred and on the Y axis, I have the change in the outcome listed at the top. On the left, you've seen the effect of down regulation on patenting within these affected device types. At the base, we have 8 patents per year in these affected device types and using an increase of almost 15 patents per year, which is a large increase in high tech space that I'm analyzing. On the right side. You're seeing the effect of deregulation comments across 21 on the number of presenting rates.
At base rate of 19 is much lower of an increase, and if not significant at the 5% level. I interpret this thing that deregulation is exposing firms to legal liability after that deep regulation events, to the effects on innovation might be less strong than when down regulation occurs. I'll say the innovation, I see that innovation is increasing most for the small and experienced firms as I mentioned earlier, and when I look at FTA device submissions I see a story that is consistent with these patenting outcomes as well. The last thing I'll say is that Catholic quality also increases come up.

So, when we look at citations on these patches and patent market evaluations, you are seeing increases as well after these events. The second set of results are related to market structure -- how do these events affect firm entry and prices. On the left. I'm showing the effects of down regulation on new entry. This time. As you can see off a base rate of about 3.8 firms, every year you're seeing an increase of about 6 to 8 new firms each year in the long run, this gradual increase. I think is natural in the space and some time to actually bring these products to market and to develop them.

On the right-hand side, you are seeing an increase in new entrants after these deregulation events. Your feet is smaller. But it is significant. And what I can ultimately hear is that almost all the increase in patents after deregulation seems to be coming from these new firms, which I think is interesting whereas there's not a lot of action going on with these existing firms. I do find similar estimates using FDA administrative data on new firm submitting devices to the FDA. What happens to prices? Here I'm looking at the price of medical procedures that effective device types. You should see, for example, a frontal fusion surgery that uses a spinal implant and I'm looking at, for example, the price of this spinal fusion surgery, so on the left-hand side.

This is the effect of down regulation on the log average price of those procedures. I don't find any discernible change the log average price for these down regulated technologies come up, up, which I think is interesting. Or is it the right-hand side, the effective deregulation seems to be negative and a significant, over the static estimate, and you can see it kind of reducing over time as more firms enter the space. And one of the reasons why I think you don't see the change on the left where you do on the right is because non regulation is affecting more differentiated products that are high tech where's the deregulation effects more commoditized products or generic products, which it may be the case that physicians are more price sensitive to these more high tech products, these tools that they are used to whereas these lower tech products that they are less price sensitive to, it also could be the case that some of these products that are down regulated form a smaller share of the procedure costs, then these technologies that are deregulated, so technologies for you can think of something like a salmonella test, it was a salmonella test itself would be a large one of that procedure.

Lastly, is going to increase innovation simultaneously, so that might have an effect on price. The last set of results I provided today are those related to product safety. So we find large innovation in market structure effects, but the natural next question is how does it affect product safety since that is the goal of the regulation in the first place? I'm showing here that downregulation may decrease product safety. But then deregulation seems to improve it, so on the left-hand side, you're seeing the effect of down regulation on the number of serious events, so what serious events are, things like life threatening events, hospitalizations and mortality.

This effect on the left-hand side isn't statistically significant. But it's economically meaningful, where it's on the right-hand side, the effects of deregulation significant. And I think, as I said earlier, this is possibly driven by the fact that after deregulation, firms are being exposed to this legal liability risk. I also find that inventors are focusing on product safety in the patent test, so they are saying keywords that are related to product safety more frequently. When I normalize the numbers of serial diverse reports for the utilization of these devices, I do. So, by dividing by the number of medical procedures that use those devices, I find that the increase in serious adverse events after down regulation on the left hand side largely disappears, so it's relatively flat here, whereas on the right hand side, this negative effect of deregulation on serious adverse events seems to persist.

I'm only looking at a subset of my device types, would include I think 46 out of 123 for which I have pre-event claims data. Just kind of reassuring. And I also mentioned that when I look at different reporter types, whether it's a provider or a user or a manufacturer, I still see these kinds of decreases in serious adverse events after deregulation, so the takeaway it doesn't seem to be evidence that their regulation has much of an effect on product safety. But deregulation seems to improve product safety, at least in the era that I'm analyzing.

I'm going to show you two different ways of me trying to show you is likely the mechanism that we are seeing the product safety improvements. The personal show is that using variation and exposed exposure to legal liability using the idea that these large firms are not or cannot use bankruptcy as readily as small firms, so this is appealing to this judgement proof issue that talked about earlier. These small firms, these large firms have pockets. So, they can pay all the legal damage that are presented in them, but these small firms can use bankruptcy to has the right distribution risk.

So, they face lower expected damages, so on the left hand side, you can see I'm showing the change in safety effort of these smoke firms relative to large firms. Those in the first asset turf style are small and those are the third are the largest. The small aren't really significantly increasing their safety effort in the patent text but the large are. During that effort, you see the change in serious -- in the Third asset tercile (inaudible). The standard errors are large. But this seems to be suggestive that these legal liability mechanisms are driving these improvements.

The public might also that liability is likely the reason we are (inaudible) deregulation events in which the class 2 devices that were affected, removed from class 2 to one were already exposed to legal liability because the protection for class 2 devices had already been removed. Once you look at the change in serious adverse event rates or adverse events after this 2015 events, you don't see a market change at all in the series adverse events whereas before we were seeing a decrease. That seems to suggest that the legal liability is, again, the reason why we are seeing these product safety improvements.

For the 2015 events. I want to emphasize there is larger Innovation -- possibly because there's a benefit coming from this because they are already exposed to legal liability risk before deregulation. What do we do with this information moving forward? Can we generalize these effects to other device types moving forward? The way I tackle this. And this is suggestive, so there's still some discretion by the FDA beyond the score that I'm using, but the FDA uses or at least in the past for the event analyzing, the FDA uses something called a device priority model, and what this does is it assigns a score for medical device types that basically coincides with how dangerous that device type is, the larger that ADPM score is, it means the more dangerous the device type is, it has more life threatening events, deaths and hospitalizations involved.

On the left-hand side, I'm showing it's changed in the series adverse events as you move up towards higher DPM scores. (inaudible) As you can see, I suppose more marginal device strikes, we're seeing a greater decrease in serious adverse events. And if you drop the outlier, on the right hand side, you can do in right,, (*inaudible*) at baseline. 5th, exam discussion evolved in the choice of deregulating certain device types for this is the partner measure they used to evaluate whether a product is safe or not. To interpret these findings, I'll give you a back of the envelope calculation.

Are you appraising the value of additional packets and I take the value and discount it dramatically to account for -- and then the value of live saving cost savings, in panel I am showing the costs and benefits of down regulation come with the benefits of down regulation seem to be about three times as large as the cost however, I flag this as being something that might not be as generalizable given that there's some margins here where the FDA has discussion. The innovation effects. I find maybe even stronger when you're kind of moving through the unaffected devices, because those that are more risky tend to have larger market sizes, so you may see bigger increases in innovation after down regulation, though I think it's a lot less clear what you would see with these adverse event and the safety of these devices.

For panel B. Reflect A deregulation is over $50 million per year for device type. But there are virtually no cost because product safety to extrapolate these seems to improve. If you were benefits to all current class 2 devices, you could see there'd be $132 billion net benefit per year across the current class 2 devices. If you were to deregulate all of that. I could say something *(inaudible*). Proxy regulators happen printed for the national academy of medicine and there's a fluid brief papers coming out criticizing them as being ineffective, so my paper simply is suggesting that a good alternative could be relying on the legal liability environment.

This context. In this paper it highlights the importance of considering innovation and market structure when you are doing or considering public policy. This type of great literary that most aspects already. I find that medical device deregulation increases innovation and market competition. And I also find that the frictions associated with regulation may magnify the costs, particularly for small and inexperienced firms that have difficulty navigating these regulations. Our friend launch regulation and every effect and what sister suggests, given the literature on the demand side policies, the effects of demand side policies on innovation, at the supply side policies that listen to constraints on firms tend to have really large effects on innovation and market structure relative to (*inaudible*).

This paper suggests that litigation is an alternative tool in more established markets where we already know quite about these products, and these results may generalize to other regulated markets for example, plus 3 devices are regulated like brand name drugs. So, they both acquire similar clinical trials and class 2 devices are regulated like generic drugs and genetically modified foods and crops. Legal liability is -- with FDA approval in these markets and lastly, the tension between regulation, litigation is present in a lot of industries like aircraft industry, automobiles, and for over 15,000 products regulated by the consumer product safety commission.

Some of the lessons learned here are applicable especially for some of these lower risk established products where we have -- or a safe record has already been shown. Hopefully this will inform policy in some sense. Thank you for having me.
[Applause]

[>>] To discuss we have Matthew Fielder.

[Matt Fiedler] Get to the appendix slides, I guess. Well, I'm very happy to be here to discuss a really interesting paper. And I enjoyed reading this paper a lot and learned a lot about medical device regulation and its consequences in the progress process. I'm a fan of this type of paper like that's deep in the guts of the federal regulatory process. And I think it's valuable in general. But perhaps particularly so for the FDA device regulation regime which I think maybe hasn't gotten the love that the drug approval regime.
I want to start with a quick summary of the paper's results. And there are really sort of two event types we are looking at here, we are looking at these Class 3 toClass 2 events where we are moving from something where we will require clinical trials for a device on the market to regime where the device just has to show that it is similar to some existing device, suitably similar to an existing device. Let me see big increases in patent flows come alongside that we see big increases in product entry. Little changed interestingly, in products and procedures that use the devices and perhaps some evidence of more adverse events from although I would describe the results as fragile and inconsistent across different measures.

On the Class 2 to Class 1 events come as a move from a regime where the product needs to show that it is substantially similar to existing device where there's no FDA approval in advance, here. There are some increases in practical as well, although they're quantitatively smaller relative to the sort of baseline level of patenting. I was just using to follow the procedures here, and find other reductions in adverse events. I want to offer a few main comments. I want to start with a data quirk related to the adverse event data and then offer a couple thoughts on how to interpret and apply the empirical results and finally, close with some thoughts on what device regulators might take away.

First the data quirk, as a starting point I want to highlight something that initially puzzled me when reading the paper. I'm showing you the main table on the paper that reports the address event results. And one of the things you'll know is that the estimated effects of reclassification on adverse event rates is far larger than the -- for treatment units. We're going to be most striking for the cross two one to events for the -- is more than 10 times the pretreatment mean and for mortality where the estimated reduction in mortality is about 60% larger than the pretreatment mean, and since these are outcomes that are always sort of positive by definition, this implies that the control group must have been experiencing 3 large increases in adverse rates and the posts relative to the pre period, which on its face is surprising. But I suspect we're seeing here (inaudible). The issue that there are changes in what types of event reports are included in that data over time. So, ‘user facilities’, the FDA term for health providers in the context are included in the data. The data's inception of 1991, distributors and various tips and voluntary reports coming in 1993 and manufacturers coming in 1996. If we look at the accurate number of fabrics event reports, we see that it rises sharply as more reporting sources come online, especially large jump as manufacturers start reporting in the middle of 1996.

The timing of those changes is notable, because a lot of the reclassification events, particularly the Class 2 to 1 events are occurring in 1986 and subsequently in 1988, I think it's also worth noting there seems to be a rapid escalation in total adverse events late in the period comment which might reflect other changes in reporting behavior. Waste powder and data can account for treatment effects are so much larger than the pretreatment Knees come up. But the more important question is really whether they pose a risk of bias to the adverse event results. I don't have a clear view but a couple thoughts to offer.

I think one is thinking about the identifying assumption in different contexts is that the pre period difference is between treatment and control is a good proxy for the sort of counterfactual post period difference. Where there are major changes in reporting, I think there's reason to worry. But it's not clear what direction the bias goes or whether there is bias at all, but I think there's reason to worry. The other thing to think about is that the scale of the outcome variable being so low in the pre period relative to the post period may mean that there's a sort of pre period -- is not as reassuring as it might ordinarily be, Just because as real violations and it turns may get lost in the noise of a small relative to these much larger scale post period outcome variables. For these reasons, I think it would be worth thinking about how to address some other limitations of the FDA adverse events. And then data, I think one good option if it's feasible and I don't know if it is, is to try to get manufactured reports prior to 1996. As I understand, FDA while receiving reports from the manufacturers during this period, but during through a different report system.

It’s conceivable FDA has that data. And it could be obtained, so I think it's worth exploring. Claiming that it would be interesting to see results that limited two types of events that are more likely to be consistently reported over time or exploring whether it appears to be the case that different types of devices are differentially affected by these changes in the reporting regime. The next thing I want to do is unpack some of the welfare estimates that are presented in the paper. What I'm showing you here is in the figure at the top of the paper is the underlying components of the welfare estimates were just presented come with the estimates here reflect the most recent draft of the paper with the caveat that the price bars are adjusted to reflect the price estimates that Parker shared today, I think in a way that's consistent with what Parker showed on his slide.

On the consumer side, there are two effects being taken into account; the one is value to the consumers - and the other is the value of the change in prices. When the firm set the market, there's a change in firm profits, which is what's really worth driving the overall welfare here. I think it's worth at least free until they enter those changes and firm profits and what the author saying here is pretty creative, to use these proactive estimates as a way of combining -- to try to estimate the change in foreign profits might be and the way he's done that is take the market value in the patent and multiply it by 20% to as he said, sort of account for business.

The question I walk away with here is whether 20% is the right factor. I was thinking about this paper. And I tell stories where 20% is much too low, and I can tell stories where something 0 or even negative is the right factor. Another thing we're thinking about here is whether there are costs associated with developing these new patents that sort of need to be netted out of the patent values since those costs are obviously sunk at the time the patent is awarded. I think both of those are places more thought would be valuable. And then I think looking beyond firm profits, I think there are some other welfare channels that might be worth thinking about here.

One, is effects on consumers for product quality, changes that aren't captured in adverse events. I think it's possible that those were magnified the value to consumers of some of the adverse event changes and then the broader scientific value of new patents. I think there's also a question of whether some of these reclassification events could have effects of going beyond just the affected device categories, in particular moving more devices into Class 2 or Class 1 might signal to firms that are thinking about more radical innovation that they're more likely to be in a regulatory regime in the future in which they are going to be exposed to -- and that could affect the incentives for developing some of those fundamentally new products.

I think I'll skip over the last point, in the interest of time. In terms of what device regulators should take away, I think for me, the sort of clear message from this paper is that these regulatory decisions and perhaps particularly the Class3 to Class 2 decisions really have big effects on patent activity and product entry. That's an important fact and I think it's one that should affect how the FDA thinks about these decisions in the future. What's a bit less clear to me at this stage is how these events affect welfare, and what I skipped over, I do think there's some question given the amount of discretion that FDA had in deciding which devices get reclassified of how generalizable these events are to other devices in these categories. Even with those caveats, I don't want that to over shadow the fact that this is an interesting paper that adds a lot to our knowledge and an important and understudied area, so thank you.

[Parker Rogers] I want to say thank you to Matt for his feedback. You don't know that he's been giving me feedback beyond just his discussion, and it's been valuable. So, thank you so much. I just want to say one thing that we've also communicated about offline, I was able to do an analysis using -- excluding the manufacturers, because they showed there was a big increase in number of manufacturer reports. And I was focusing on users. And the other reporters and I was able to show that the effects were consistent and still a reduction in serious adverse events after deregulation. But point taken for sure about potentially expanding to pre 1996 manufacturer reports, so thanks a lot, I appreciate it.

[Q&A]
[Q: Michael] This is Michael from the FTC, I'm curious about the relationship between innovation, competition and prices, and you sort of present them as separate endpoints, but possibly especially for these 3 to 2 events, the increase in innovation and the increase in -- could push prices in opposite directions, so I'm curious if you thought of ways to potentially decompose that effect. And see if there are competing forces there?

[A: Parker Rogers] In this paper, I felt that I didn't have the bandwidth or whatever, but I couldn't really do it here, but I absolutely think it's an important thing to disentangle given that I think that the innovation margin is quite important in determining the prices, given that is potentially shifting up the demand curve for that device, so it's kind of ambiguous what would happen there if demand and supply is shifting. It's in the books to look at that sometime. But I haven't yet.

[Q: Dave Schmidt] Dave Schmidt, FPC -- my question is about do you know anything about what the FDA says when they are down regulating or deregulating, like what is their reasoning? Do they do this when they think hey, maybe if we lower the entry costs, finance has gotten to a point for this particular technology that we think more entrance will be able to come in now whereas five years ago, we didn't think the science was there?

[A: Parker Rogers] Further deregulation events come I will say these events happened that there's mass deregulation that happens once every couple of decades, I just don't have it very often. And usually, it's the results of Congress pushing the FDA to do something and then they find that -- they're kind of under resourced for some of these big, huge down regulation, deregulation events. So, they use cruder majors evaluating the safety of products and that was kind of what I was talking about with the device --, they use these models to evaluate which devices to deregulate and that's what they say in the federal register notices about why they are doing it.

It's mostly because Congress and they are using this measure to see which ones they should deregulate. For Class 3 to Class 2 events. It's much more nuanced, it's more about product safety. And it hasn't been around long enough, do we have the safety -- and if that's the case, they sometimes choose to down regulate some of those products. Cortana. You like much more indulgence in that sense, I think mostly on the average event side, less on the innovation side, although I think that's been changing recently, for deregulation. I think it's a lot more crude in that regard as far as what they are using as a justification for that deregulation.

[Q:] And I'm wondering if there are also benefits from this reducing the burden on the FDA. Things get done faster because they're not doing these things that were marginal?

[A: Parker Rogers] That's an important point. And if you look at the approval timelines across like how long it takes for the average device to be approved, you see really big spikes in the approval timelines of exactly around the times when they need to come up when they're doing these mass deregulation events, so absolutely they're having to shift the workload around meaningful ways when they do this. And I think yes, reducing the cost on the FDA for approving and clearing these devices in the 1st place probably is valuable. I will say when you move to a litigation framework, there's also the cost associated with litigation on society.

And so that's something that I haven't also considered, so I think it's a little bit more direction as far as the production of administrative burdens on the regulatory agency, but also increase in burden on the legal system, but yes it's a fair point.
[Applause]
[End of Q&A]

[>>] That concludes our presentations today. We have a reception outside and we'll see each other tomorrow morning, too.