FTC Informational Injury Workshop
December 12, 2017
Segment 5: Panel 4
Transcript

DOUG SMITH: So hi. Good afternoon. My name's Doug Smith, and by co-moderator on this panel is Jacqueline Connor. So welcome to the last panel of the day, Measuring Injury. As I think several times it's been alluded to on the previous panels, there are particular empirical issues or challenges that come up in trying to assess informational injury. And fortunately today, we have a wonderful group of panelists to help us understand them.

JACQUELINE CONNOR: Yes. As Doug said, we have five wonderful panelists with us this afternoon. So I'm going to give you a brief introduction to each of them. First we have Garrett Glasgow, who's a Senior Consultant at NERA Economic Consulting where he specializes in market competition, consumer protection, intellectual property, and environmental cases. He has presented and published on several topics related to damages stemming from violations of privacy or misuse of personally identifying information.

Next to him we have Ginger Jin, who is a Professor of Economics at the University of Maryland, who was on leave at the FTC in 2015 to 2017, serving as the Director of the FTC's Bureau of Economics from January 2016 to July 2017. Most of her research has focused on information asymmetry among economic agents and ways to overcome the information problem.

Next we have Lynn Langton. She is the Chief of the Victimizations Statistics Unit at the Bureau of Justice Statistics at the Department of Justice. She is responsible for overseeing the collection and dissemination of data from the National Crime Victimization Survey, a large-scale survey of US residents that serves as one of the two measures of crime in this country.

Catherine Tucker is the Sloan Distinguished Professor of Management and Professor of Marketing at MIT Sloan School of Management. Her research interests lie how technology allows firms to use digital data to improve their operations in marketing and in the challenges this poses for regulations designed to promote innovation.

And last but not least, we have Josephine Wolff, who is an assistant professor in the Public Policy and Computing Security Departments at the Rochester Institute of Technology. Her research interests include cybersecurity law and policy, the economics of information security, security metrics, incident reporting models, and insurance and liability protection for computer security incidents.

So first we'd like to have each panelist talk a little bit about the work that they've done relating to measuring informational injury. So I'm going to read a brief question, and then if each of you could go down the line and kind of give a quick response to that, that would be great. So starting with Garrett and I guess working down, can you please give us a brief description of the research or work that you've done related to trying to measure injury in the context of privacy and data security issues?

GARRETT GLASGOW: Yes. Thank you for having me here. In terms of privacy and misuse of personal data, I'm working in two areas. One which I call the easy area is misuse of information by companies.

So where a customer and a company might have some kind of business relationship, certain levels of privacy or data protection are promised, and then the company, for whatever reason, doesn't live up to that, doesn't live up to the promises. I've been working on survey methods to try to value how much of the price a consumer pays for a product bakes in this data protection or privacy. That's one area where I feel like we've made some good progress at NERA.

The other I call my blue sky research, and that's trying to determine whether there's an intrinsic value of privacy. So not the theft of data or the misuse of data. But do consumers, do individuals actually obtain some kind of utility or value from just knowing that their data's being kept private and it's not being pried into by unauthorized people?

This is an idea that comes from environmental economics, which is another area where I work where we talk about the intrinsic value of the environment. It's not something that has value because it's the hub of a lot of economic activity, but it has what we call passive use value where just we gain some kind of happiness or utility from just knowing that there's a pristine environment that hasn't been damaged.

And if it is damaged. Or somebody's owed some kind of compensation. So I've been doing research in that area as well to try to determine if privacy is similar to an ecosystem that might be damaged by bad acts or by carelessness.

JACQUELINE CONNOR: Thank you. Ginger?

GINGER JIN: Thank you. Thank you so much for having me. I have to confess that I have done little research in this area before coming to FTC in 2016. However, while I was at FTC, I dealt with a number of cases in privacy and data security, which prompted me to think a lot about information injury.

One crucial question is whether we should measure consumer harm in ex-ante or ex-post perspective. I know the second panel has already covered on this, but I still want to put in my two cents. The ex-post perspective is quite intuitive. Somebody misused consumer's information that results in, say, an identity theft or a fraudulent transaction. We can measure that by the amount of time, money, and effort that consumer's lost because of this.

However, the ex-post perspective, I would argue, is narrow-minded, because a lot of harm may not happen yet, but there's a risk there. Or the harm has happened, but we can now link to a particular firm who have done that data practice. So in these situations, we observe zero traceable harm, but that does not mean there's no consumer harm.

I would also argue that, in fact, emphasis on ex-post harm, what ends up encouraging overuse or misuse of data, because firms who engage in those misuse of data, they don't need to account to

the negative [INAUDIBLE] they are imposing on the consumers. So in my view, that's inadequate.

In contrast, the ex-ante perspective would emphasize on the increase of risk to harm consumers, even if that risk has not being realized or could not be traced back to a particular firm. So I heard a lot of panelists talking about Equifax. Myself is a victim of Equifax. I worry about my record on the black market.

I end up paying something to try to freeze my account. I even sort of welcomed the inconvenience. I have to lock and unlock my account in order to get some credit in between. And to me, it spends my time and effort and money, and that is harm to me, even though we haven't seen the ultimate harm coming back in my way. And it's probably very hard to trace that back.

And similarly, if I know my favorite retailer has some bad data practice, even if that company has not had data breach, I will be worried. I would want to probably do something to reduce that risk. In that sense, the ex-ante perspective would say there's some harm there.

And to follow on Garrett's suit, we can make an environmental analogy here. So for example, a firm has polluted my neighborhood and exposed me to a higher risk of cancer. I would say the company should be responsible for that action, even though I have not developed my cancer in my body yet, because it increased the risk I'm exposed to. So in that sense, I believe the ex-ante perspective is more appropriate.

And then a point I want to make is that the biggest difficulty in measuring harm is not measuring harm itself. It's measuring harm attributable to a specific firm. We can measure the extent of identity theft or other things, but exactly how to tie that back to a particular firm, I think that's the most difficult question.

JACQUELINE CONNOR: Thank you, Ginger. Lynn?

LYNN LANGTON: Thank you for having me here today. As Jacqueline mentioned, I oversee the National Crime Victimization Survey, which is one of two measures of crime in the US. It dates back to the 1970s when it was developed to be a complement to police statistics because of the recognition that a large portion of crime goes unreported to police.

And so if we just focus on those crimes that are reported to law enforcement, we're missing a big piece of the picture. And this is particularly true when you're talking about sensitive crimes and when you're talking about more of what we would call sort of white collar crimes or emerging crimes, though I don't think emerging is really the correct term anymore. So crimes like identity theft and fraud that are highly unreported to police.

So the NCVS is a household-based survey. It's conducted for BJS by the US Census Bureau. We go to an incredibly large sample of households and interview all persons age 12 and older within those households about their experiences with victimizations. So our sample size is-- right now, we go to over 200,000 people every year, and we get response rates in the high 70s. So if you do survey research at all, you know that that's pretty remarkable at this point in time still.

And we ask questions about their experiences with victimization, and also the nature of that victimization and the victim response to that victimization. So as Ginger was just speaking about, this is the ex-post perspective very much.

The National Crime Victimization Survey focuses specifically on violent and property crime, but we do conduct a number of supplements to the survey. And the one that I think is most relevant for the conversation today is a supplement that we've been conducting for a number of years on identity theft.

Of course, identity theft is just one type of informational injury. Again, ex-post. Very specific harms associated perhaps with a data breach, and often just because of misuse of personal information through other means.

But I think there are some important things that we can glean from the identity theft supplement that are relevant to the broader conversation here today. So the supplement collects information about the attributes of victims and victimizations and the response to victimization.

But it also asks questions about the harms experienced by victims who suffered from identity theft. And on top of that, it asks victims and non-victims about behaviors that they engage to prevent any type of misuse of their identifying information, and also whether they have experienced any kind of data breach in the past 12 months.

So when we look at that information, we can see that among individuals who have experienced a data breach, the risk of identity theft or the prevalence of identity theft is double that of those that have not experienced a breach. Not that surprising, but putting numbers behind what we know to be true, or assume to be true.

And then among identity theft victims, about 20% of the victims in our survey say they also experienced a data breach during that reference period. So there's certainly a correlation there. Not surprising.

I think the other point that's important to take away, though, is that even among identity theft victims, that experience no financial losses-- so they have no financial losses whatsoever-- there are still harms. So we asked them questions about not just their losses and the time spent, but also about how distressing they thought that the incident was.

And even among those identity theft victims that didn't experience financial loss, about 30% still found the incident to be moderately to severely distressing. So that suggests that in and of itself, the experience of having your information misused, having your information out there has some harm to victims. And so I think that's relevant for our conversation today.

CATHERINE TUCKER: All right. So I'm Catherine Tucker, and I'm an empirical economist, which means that I'm going to behave in a predictably economist way and say two things. The first thing is a lot of my research has been focused on how hard it is to actually measure this. And in particular my research, which is focused on the use of algorithms and AI, suggests this is

just going to get harder and harder as the boundaries of an individual's data gets fuzzier and fuzzier.

The second thing which I'm going to say, which is predictably economist, is that my research has really focused a lot on documenting the trade-offs which come from trying to protect consumers and their privacy. And let me tell you-- give you sort of three typical unintended consequences that I found in that research.

So the first unintended consequence I want to highlight is often when we have privacy or data breach regulations, we're very focused on the idea of big companies and the idea that they need to be provoked to do something better. Well, my research suggests it's actually smaller companies, startups, which tend to be most adversely affected in terms of the costs of this kind of regulation.

The next piece of research for unintended consequences I want to highlight is some research I did into data breach notification laws, and in particular what happened when they gave exceptions for encryption. Now in this research what we found is that actually, there were more data breaches as a result of these exemptions. And why is that? Well, they just focus on encryption, and it meant you had a whole lot more doctors losing laptops from their car as a result of people focusing on just one dimension of data security.

Now the last thing I just want to highlight as an unintended consequence is some research we've recently done about patient data in hospitals, which highlights that actually, hospitals are using data security and privacy regulations to actually stop patients themselves getting access to their medical data records they need if they try and transfer hospitals. And some people in the room, I see you nodding, right? You've experienced this of trying to get your own medical data out of the system.

JOSEPHINE WOLFF: Thank you. So I work on a couple areas related to thinking about the injuries and the types of harm that come out of data breaches in cybersecurity incidents more generally. And the first time I sort of encountered this was in a project looking at the aftermath of security incidents and trying to trace, who ends up suing whom and who ends up paying whom?

And a lot of this is at the corporate level. A lot of this is the credit card companies and the payment network suing breached retailers and demanding certain amounts to settle the fraud costs. And then there are also a number of the consumer class action suits that you heard talked about a little bit at the panel before lunch. And for all of the reasons that they brought up on that panel, those are often very hard to see sort of go through or even heard by courts.

And I got interested in this question of sort of how different types of policy agencies, different parts of the government, and different parts of industry think about and understand these types of harms or injuries. And so I did another project looking across the different ways that people had tried to measure them, and bucketing a couple different kinds of data.

There is a lot of survey data in this space. A lot of sending out reports to different companies and saying, how much did you lose due to breaches in the last year? And putting that together as sort of self-reported data to think about how much we think these breaches cost to the companies, to their customers sometimes.

One of the big sources that some of you may have seen is the Ponemon study that comes out every year. Again, focused on self-reported numbers and trying to stack those up against some of the other data sources that we sometimes use to think about quantifying these types of harm.

One of those was insurance claims. NetDiligence has a data set of companies that have filed cyber insurance claims for harms due to computer-related losses, and looking at sort of how the numbers change if you look at what they file for insurance versus what they report in a survey about these losses. And then finally, looking at a lot of the analysis that's been done around stock prices and market value of breached companies, and trying to understand sort of how that tells a slightly different story if you use that as a proxy for thinking about the costs or the consequences of breaches.

And then to come back to sort of thinking about the consumer injury. And not just the losses to a company, but also to individuals. I spend a lot of time looking at the firms that are offering cyber insurance policies now, and how they price those, and how they think about the kinds of harm or losses that they will or won't cover.

And increasingly, we're seeing a lot of them get into the still relatively small market of personal cyber insurance, selling policies to you individually or to your family to cover certain kinds of losses. And trying to understand how it is that we understand the ways people are purchasing those types of policies, what kinds of things we're covering with them as a way of thinking about the types of losses that people are concerned about.

And also coming back to sort of some of the skepticism you heard on the previous panel from Omri about, what do people actually do? Not, what do they say they care about, but what are they actually willing to pay for? What can we see when we look at this emerging personal cyber insurance market about the types of injury customers seem most concerned about and most willing to actually pay for protection?

DOUG SMITH: All right. Thank you, guys. So now we'll sort of jump into the discussion. We're going to start off, I think, by talking about-- kind of trying to measure preferences, right? Trying to understand what consumers value.

And so I think a good place to start that would be sort of how to measure just what consumers say about what they value. And Garrett, you talked about doing surveys on this, so maybe you could tell us a little bit about how that works.

GARRETT GLASGOW: Sure. So my work in measuring privacy, measuring the value of data has primarily been survey-based. Before coming to NERA about three and a half years ago, I was a professor in the Political Science Department at UC Santa Barbara where I did lots of

survey research. So maybe it was natural that I fell into a survey-based approach to looking at privacy and the value of data.

And there are two different approaches that I've looked at. There's what's known as conjoint analysis, and then there's contingent valuation. To different survey-type approaches. And why don't I just briefly talk about each then, the strengths and weaknesses of each of these approaches?

Conjoint analysis is well known in consumer research. It's accepted in court as a reliable method for uncovering truth in a lot of contexts. And what conjoint analysis involves is we ask our survey responders to make choices, as if they were in the market making choices from among some set of hypothetical products. And these products are going to have different features or attributes. And so we can see, if they make enough choices, which attributes they seem to value and which ones are unimportant.

I'll give you an example from my own research. We did research on streaming video services. And we present people-- we say, pretend you're in the market to update your streaming video service. Here are some services that might be available to you.

They might have different streaming speed. Some might have high definition available. Some might have more television shows. Of course, different monthly prices, and so on. And we just present these different services. We'd say, well, pretend these are the three that are available in your area or that you could possibly purchase. Which one of these services would you purchase? And they can make those choices.

Now when we want to bring this to privacy, one way we can do this is just regard privacy as just another feature of a product. Obviously, this is only going to work if we're looking at a situation where it's a consumer engaging in a market transaction with some company that's made some kind of promise about privacy or how they're going to treat your data. So we can say-- what we did with this paper that I'm talking about is part of that streaming video service was.

And then there's different possible privacy policies that these services offer. They might say, we never share your data. Others might say, well, we'll collect your viewing habits. Not your personal information, but we'll collect your viewing habits and package those up, and use those to help content providers decide what kinds of shows to make. And then a third option is, maybe we collect your personal information as well, and we might share that with third-party marketers, and so on.

And what we were able to do with this survey was see, what value do people place on protecting their data? Or conversely, what kind of discount do we need to offer consumers to get them to agree to share different types of data?

Strengths are this is a really straightforward, reliable way to measure consumer preferences. It's widely used. It's widely understood. I regard its main weakness is narrowness. I said, these are consumers engaged in a market transaction. What do we do in all the privacy cases where there isn't a market transaction like a data breach?

I mean, I suppose we could design some kind of conjoint analysis. We'll say, now pretend there's a certain percent chance that this company's going have a data breach in the next year. But I don't think most consumers think about their purchases that way. It's a strange hypothetical to pose to people, and I think you'd get really strange results with a survey like that. Maybe it's possible to do, but I think that's one of the main weaknesses.

But at least in areas where we can apply this, this conjoint analysis lets us measure the value that individuals are placing on their privacy or on different data sharing policies. It gives us what we call a willingness to pay to protect their data.

And then we're talking about quantifying damages. If a company doesn't live up to that promise, we can use that willingness to pay to calculate damages, that we can use that as the base just to say, a certain percentage of this price was privacy protection. You didn't live up to privacy protection. You owe consumers some refund or some amount of damages based on your failure to live up to your promise.

And that's-- I mentioned earlier, there's easy cases and hard cases. I think those are kind of the easy cases. I think we've made some headway in terms of measuring the value of information, the value of privacy there.

Hard cases are all the other ones, things like data breaches, and so on. And one possible approach to that is what's known as contingent valuation. And a really rough description of that would be asking people how much compensation they would need if some bad event happened.

This is very common in the environmental setting, in areas where there isn't a market transaction. There's There isn't a market for the environment. But there can still be damage done, and there might still be compensation owed.

The classic example of contingent valuation came from the Exxon Valdez spill. So the Exxon Valdez was a big tanker that spilled a whole bunch of oil in Prince William Sound up in Alaska. It's a place most people have never seen. There's not really much economic activity going on there except for the tankers going through. But people still felt aggrieved. There was still damage. How do we measure what kinds of compensation Exxon might owe to the American people?

And one way that this could be done would be contingent valuation. And they went out and conducted surveys. They said, well, suppose we had to increase taxes next year to create a system. Maybe a tugboat system that would make sure these kinds of spills didn't happen again. How much would you pay to prevent this from happening? And people would give us answers, and you could tabulate that up and say, this is the value of that environment to these people.

You could easily see taking this over to a data breach environment or a privacy environment where we'd say, a data breach has happened. How much would you pay to do business with a company that has some lockdown system where they've reduced the chance of a data breach to almost nothing? We could see designing a story that would be something along these lines.

The strengths here are that unlike some other surveys, this is something we can at least apply to something that's not a traditional market transaction. A data breach is not a traditional market transaction.

The weaknesses are that these are kind of strange hypotheticals that we're presenting to people. And people, especially with issues that are a bit emotional, things like environmental damage or privacy data breaches, people often give what are known as socially desirable or protest answers. How much would you pay to clean up Prince William Sound? Well, if people say, I'd pay $50,000 to clean up Prince William Sound, well, their income was $45,000, they're just trying to signal to us that they're angry about what's happened.

And I would imagine with doing surveys on data breaches, the same kind of anger and the same kind of protest answers where people want to signal they're angry at companies for allowing a data breach to happen or for leaking their data. And it might not be very useful in terms of measuring damages. But that is one possible survey approach to dealing with data breaches. I'll talk more about that later on. But that's an overview of two major survey approaches.

DOUG SMITH: Great. And I want to open this question up to all the panelists. But first, Catherine Tucker, you've done some work in this area. So I was wondering if you could comment on what you've done and found.

CATHERINE TUCKER: Oh, yes. So just as I previewed, it's going to be about the difficulties in this. And the work I did was basically investigating something we call the privacy paradox, which is that if you ask people in a survey how much they care about privacy, everyone cares about it. But then we see all this actual behavior which is inconsistent with that stated preference.

Now what we did in this study is it took place among the MIT undergraduate population. And we asked them to take part in a study, and as a reward, they got $100 in bitcoin. And this was in 2014. So those of you who can do the math, this is a lot of money right now. We made a good call there.

So they were doing this study. And one of the questions we asked them for was some data on some really quite personal and sensitive information, which was data about their friends' contact details.

Now in the ordinary setting, when we asked this question, a surprising number of our MIT undergraduates really decided they were not going to answer this question, but instead they were going to give us fake information. And we knew it was fake because of their use of swear words and the use of swear words employed to describe what they thought about us for actually asking this question as part of the email. So it was definitely deliberate.

Now that was the first setting. And then for half of them, we changed the context somewhat in that we gave them an offer of cheese pizza. And if you wondering how much cheese pizza, sort of a slice of cheese pizza.

And what we found was that when there was no cheese pizza involved, people tended to behave in a way which was consistent with what they said in a survey about their privacy preferences. But the moment we gave them the offer of cheese pizza, even those people who said they really cared deeply about privacy, started giving away this quite sensitive data.

Now in terms of interpreting it, it's really quite hard, right? One thing you could do is sort of take an economist response, which is that stated preferences, it's hard to use them to measure anything. We should use revealed preferences.

Another interpretation of the study is, oh my gosh. If MIT undergraduates behave like this, people who should really understand technology, maybe we need to really protect people to actually get them to behave in the line of their stated preferences. So no conclusions. Just some difficulties in terms of using survey responses.

DOUG SMITH: Thank you, Catherine. Ginger, do you have any thoughts about--

GINGER JIN: Yeah, I would just want to follow up with Catherine's commons on the privacy paradox, and how actions seem to differ a lot from stated preferences. I think the biggest question is to distinguish different explanations behind this, because consumers don't know the risk, so they're willing to give it away, or because they know the risk, but they somehow believe the benefits dominate the risk.

Or they see the extra risk of giving away this to one more person, given that their data probably already been breached multiple times is so small that they feel helpless, and so therefore, they give away. I think those explanations would sort of imply very different policy actions. I would be really interested to see future research be able to distinguish those explanations.

DOUG SMITH: Thanks, Ginger Lynn?

LYNN LANGTON: So as I'm sitting here thinking about measuring actions, I'm realizing that there's more that we could potentially be doing with our survey. So we're not getting at consumer preferences in any way, but we are asking a series of questions about behaviors that they engage to-- again, this is in the context of identity theft. But behaviors they engage to prevent the misuse of their personal information. And I think, again, this speaks to what they're actually doing.

So you know, these are pretty basic things, but I think the proportion of people that do these things regularly is still pretty surprising. So if you ask them if they check their bank and credit statements, OK, 76% say they do that, which makes me wonder about the other quarter of people, I have to say. Who doesn't ever check their bank statement? But that's a different subject.

[LAUGH]

But then when you move into things like, do people regularly change passwords on financial accounts, about 30% say they do. Checked their credit reports. That's about 40%. Purchased identity theft insurance or a credit monitoring service. Again, maybe this gets more into that willingness to actually pay to spend money on something that might prevent this from

happening. That's only 5% of the residents that are participating in the survey that say they actively do this.

So now I'm thinking, what could we ask? Could we ask some questions related to data breaches specifically that would get at the behaviors that they engage in? So even just, how often do you consider whether a company has had any sort of breach that you can find information about before you provide your information to them? Just to kind of gauge whether or not there are actions being taken to actively avoid potential situations.

And again, that companies can change over time, and so those are challenging things to measure. But I think there are still some actions that we could think about concretely getting at that would address this issue a little bit more.

DOUG SMITH: Thanks, Lynn. And I think with talking about that, you sort of broadened the conversation to looking directly at people's actions, or at least their reported actions. So Josephine, would you like to comment on either sort of stated preferences or revealed preferences, or both?

JOSEPHINE WOLFF: Sure. Well, I think that a lot of what we have when we look at the data that people have analyzed and collected around the injuries and the costs associated with data breaches is very much stated preferences. Most of what you see, especially say in industry reports, is going to be kind of self-reporting, this is how much this breach cost, or this is how I was affected, these were the injuries. And so I think that we start out from a position of basing a lot of our estimates, basing a lot of our understanding of these types of costs on those stated preferences just for practical reasons.

That that was sort of-- and still is, in a lot of ways, the data that's easiest to collect in large volume. And it's definitely true, as several of the other panelists have brought up. There's a huge discrepancy when you start trying to compare that to what people are actually doing. And it's very hard, as Ginger says, to get at why exactly that is.

And I think one of the things that's complicated when I look at legal cases around these class action suits in the aftermath of data breaches is that getting involved in one of those class action suits is not always just a sort of straightforward question, do I care about my data? Do I value my privacy to some amount of time?

And the strongest example of this, I think, is probably the Ashley Madison class action lawsuit where the judge actually decided if you want to go forward as a member of this class action lawsuit, you're going to have to use your real name, right? So it was this sort of complicated decision for everybody whose data had been stored by Ashley Madison and breached of, OK.

On the one hand, I value that data and the loss of privacy that came with everybody in the whole world potentially knowing I was a member of this website for adulterous affairs. But on the other hand, if I now move forward with trying to pursue that as an injury in court, I'm going to lose this privacy in a slightly different way.

And that kind of decision, I think, is not actually as straightforward as saying, oh, well, I guess nobody cares about the injury that was done to them in this data breach because they're not willing to pursue this lawsuit. But more about trying to understand, what are the actual decisions that people are making when it comes down to acting on what they think are their preferences, on what they think is what they value?

DOUG SMITH: So in some sense, understanding sort of the costs that they're trading off, the benefits of privacy is sort of important to figure out how to measure it?

JOSEPHINE WOLFF: Absolutely. And I think gets at some of that discrepancy we see between sort of how much I think I value my privacy and the things I'm actually maybe willing to do in practice to protect it or pursue my losses.

DOUG SMITH: OK, great. Thanks, Josephine. Yeah. So do people have any further thoughts about sort of measuring consumer preferences through their actions? Garrett?

GARRETT GLASGOW: Yeah, actually. I think we've heard several different stories now of a mismatch between consumer stated preferences and consumer actions. And this is a regular feature of privacy research. We see this a lot, that people seem to take actions that contradict what they've stated matters to them.

And I think that it's often used to dismiss the concerns of consumers in what I think is an unfair way. It's a complicated issue, whether you-- people might be willing to take some actions in some contexts, but not in others.

I'll give you one example from my research, the conjoint analysis I mentioned on streaming video services. About 10% of our survey respondents said, I would never choose a streaming video services that collects any information about me, whether that be which movies I watch or personal information, or anything like that.

Well, one of our screening questions to get into our survey was, do you currently have a streaming video service? People had to say yes to get into the survey. So 10% of people already had a streaming video service that was collecting information about them. I don't know how many of you saw the Netflix tweet from yesterday, the day before, where Netflix sent out a tweet saying, for the 57 of you that have watched the Christmas Prince 18 times in a row, are you guys OK? Something like that.

[LAUGH]

So they clearly are collecting information on who's watching what. And they even sent out a tweet kind of teasing people about movie choices. So what's happening with these 10% of people in the survey that said, I would never choose a streaming video service that shares my information? Yet they already are doing business with one.

One possibility is just lack of information. They weren't actually aware that these streaming video services are collecting this information. And that's one possibility for this mismatch

between action and preference. Although if somebody's really a privacy hawk and says, I would never share information, they probably know who's sharing what.

Another possibility is this is that signaling of importance that I mentioned earlier, where you give a protest answer. Maybe those 10% of people, even though we tried to be careful, they figured out, this isn't a survey about streaming video services. This is a survey about privacy. So I'm going to send a signal that I value privacy by saying, I never share anything, and that'll let those researchers know that this is really important to me. That's another possibility for the mismatch. It's not that that consumers are being inconsistent. It's just we failed as researchers to try to hide what we were trying to get at and honestly measure that.

And then the third possibility is the context. In some situations, you're willing to share. In some situations, you're not. And consumers would like to pick and choose. Or in some cases, you might perceive a threat. In other cases, you might not.

And with another little story, this is from about three years ago. An artist baked some really fancy cookies that had the Facebook logo and the Google logo and so on, and went out into New York City and gave them away to people.

Well, didn't give them away. She didn't sell them, but she said, you have to give me personal information if you want this cookie. And about half of the people were willing to let their photograph be taken in exchange for a cookie. Half of people were willing to give what they said were the last four digits of their social security number. A lot of these people were willing to let her write down information off their driver's license in order to get a cookie. And one third of people were willing to give up their fingerprints for a cookie.

So you say, well, how do we reconcile that with people that say they care about privacy? Well, maybe there's a context issue here that an art student probably isn't going to take those fingerprints and then go rob a bank using them, or something like that. There's probably a very low risk there. But if a data hacker gets the same kind of information and it winds up on the black market, that's a very different risk.

So I don't have answers to this necessarily, but I would say there's often a mismatch between preferences and actions when it comes to privacy. It doesn't necessarily mean that consumers don't know what they're doing. It's a complicated issue.

DOUG SMITH: Thanks. Ginger, do you have any further thoughts?

GINGER JIN: Yeah. I just want to emphasize that as economists, I often believe actions are louder than words. But in this particular area, I really believe the survey approach. And the revealed preference approach should be really complements to each other.

We've actually already seen some interactions here while Catherine's talking about sort of MIT students actually in the lab while Lynn was trying to say about how you can design a better survey question to get at that sort of behavior in a national representative sample. And I guess the feedback could go the other way.

That's if we know, for example, consumer lack of information or feeling helpless is one of the primary reasons for them to behave a certain way, maybe Catherine can design a very clever lab experiment or field experiment to really tease out people's action. I really see the positive complementarity between the two approaches.

DOUG SMITH: So we're kind of go along on this topic. But Lynn, do you have any sort of last thoughts on this?

LYNN LANGTON: I'm sorry. What was--

DOUG SMITH: Oh, do you have any last thoughts on this before we move on to outcomes?

CATHERINE TUCKER: Yeah, I would just like to also add, to make it worse, that actions don't always translate to actions in the way we would like. If you sort of move along in the bitcoin study I was talking about, we then tested really how much laziness affects people's privacy preferences in that we asked them what wallet they wanted to put their bitcoin in.

And we gave them a variety of wallets, some which protected their privacy, something which completely didn't. And the biggest predictor of what wallet they chose was not their privacy preferences. It was instead how far up on the page that wallet was.

Then we then said, OK. Well, maybe it's just like a lack of information. So half of the students saw lots of information about the wallets. Basically, everything you could ever want to know.

And there was some good news in that that reduced, shall we say, half of the laziness. But I guess the bad news is there was still half of the laziness there, even for people expressing a lot of privacy preferences. So in other ways, you can actually change people's actions a lot simply by the context in which you get them to make the choices, making even this action data [INAUDIBLE], Ginger, I'm afraid.

[LAUGHTER]

DOUG SMITH: Just any last thoughts on--

JOSEPHINE WOLFF: Well, I mean, I think what Ginger says about needing both the survey and the revealed preference data is really important, because I think there are certain types of injuries. And you heard a lot about that this morning, all of the different ways that people can be affected by data breaches. I think there are some of those injuries that we really have no other way to get at, especially when we're talking after a breach has occurred when we're not necessarily going to say that all of the different types of emotional or psychological harm you may have suffered are things you can obviously mitigate, right?

It's one thing to say, well, people didn't go and change their passwords, or people didn't go and file a class action lawsuit. But if all of your emails have been released in the Sony breach, then it's not obvious what the action you would take would be to demonstrate that that really affected you, that that was a real kind of injury.

So I certainly think it's true that the self-reported data has a role to play in this. I think where it becomes tricky is when you're looking at really kind of crisp quantification for things like legal remedies or policy interventions, whether there's any way to turn that into something that can be calculated precisely enough for those types of remedies.

DOUG SMITH: Thanks, Josephine. OK. So we're going to have to leave the topic of preferences and move on to an alternative, which is to just straight out measure outcomes and try to understand things that way. So Lynn, I'd like to hear more about how your study allows us to get at that question.

LYNN LANGTON: So again, we're talking in the context of identity theft here. So beyond just any sort of data breach or privacy-related issue. But we ask a whole series of questions related to both tangible and intangible harms. Again, ex-post harms associated with the misuse of personal information.

So of course, we ask about financial losses. That's an obvious one that you have to include. We also ask about the amount of time that individual had to spend clearing up the issues related to the victimization. And then we ask a whole series of other questions, trying to get more at some of these intangibles.

So I mentioned already that we ask about distress. We also ask about whether the incident resulted in any problems with family, friends, work, school. And then we ask questions about whether they experienced any credit problems related to the incident, whether they experienced any legal problems related to the incident, whether they had to deal with debt collectors. So these are some of those more intangibles.

And then I think to get more at those intangibles too and really the impact of an incident on an individual-- and again, this goes broader than identity theft. You know, the other thing that we can do is then sort of cross these different types of outcomes and different types of harm.

So when we look at, for example, how much time an individual has to spend dealing with an issue and then we look at the level of distress, I mean, there's a positive linear relationship there that's pretty strong. So when an individual has to spend six months or more dealing with this misuse of their information, a large portion of them say they were severely distressed. Whereas if they're able to resolve it in a day or so, you know, a smaller portion say that it was severely distressing.

So again, it's indirectly translatable to data breaches, but I think sort of the same ideas. And you can use survey research to sort of tap into these ex-post responses and harms that victims experience as a result of these incidents.

DOUG SMITH: Thanks, Lynn. Catherine, do you have any thoughts about this? About sort of measuring outcomes, what data we might look for, or how we might take things that are sort of less tangible and convert them into something we can sort of quantify?

CATHERINE TUCKER: Well, so my thought actually was that really when it comes to measuring injury, I think we lack a good model of the supply side of people who want to do injury. And what do I mean by that?

Well, I've seen some really intriguing research at the Workshop on the Economics of Information Security, which is all about your potential for actually being injured as a result of a bad actor on the dark web. And you should go and look at this research, because it's fascinating in that these great computer scientists go out there and they actually try and work out, if I'm trying to do identity theft, how do I do it? And so it's all very interesting.

But what they found out, which really surprised me, is it's just not a very scalable profession. And as a result, basically they take so long to do their identity theft that it tends to be 100 people each time who are injured, and it doesn't vary as much with the numbers of records that are left or lost as you might think.

And so I think, if I [INAUDIBLE], because Lynn's collecting such wonderful data, I think if we were to really try and have a measure of identity theft or the risk of it, which is sort of the first point of measuring the injury, I would actually like to sort of highlight this research, which talks about the supply side and some of the surprising insights, that it's not very scalable, which may explain why sometimes the incidence of identity theft from a data breach is less than we might think.

DOUG SMITH: Thanks, Catherine. Josephine. Any thoughts on these measurement questions?

JOSEPHINE WOLFF: So one of the things I think is really interesting about outcomes-- and I come up across this a lot when I'm both talking to people who sell insurance and people who buy insurance in this space-- is that outcomes are much easier to tie to an individual person than to an individual incident. And so if we're talking about, what were the consequences of this particular breach of Equifax or whoever else, that's a very, very hard question to answer, because your information has been stolen so many times. Earlier today, they referenced-- they think most of the Equifax information had probably been available on the black market even before that breach.

So it's hard to say, this incident of identity theft or this particular even sort of emotional or psychological toll associated with that was the fault of this particular company and this particular breach. And it's much easier, though it's not easy by any stretch, to say, this is the outcome associated with this individual person. This is the amount of financial fraud. This is the amount of time. All of those things are easier to measure in terms of people.

And the reason that's important, I think, is because all of our thoughts and all of our sort of ways of understanding how policy can come in and try and correct these externalities have to do with, what are the levers, or what are the pressures we apply to the companies? And say, your breach caused this amount of damage.

Which is actually a much harder question to answer than, what were the outcomes for this individual person as a result of the sum of all of the incidents that they've been involved with?

And I think that's a big part of the reason that we're starting to see a lot of these solutions-- or at least the insurance market certainly thinks of them as solutions-- coming to center more on the individual than on the companies.

But I think one of the things that is perhaps problematic about that is even though that aligns with the ways that we can perhaps calculate or collect data more accurately, it means that all of the entities that have the most power to decide how this data is being protected are not necessarily the ones who have the real incentives to be executing that in the same way.

DOUG SMITH: Thanks, Josephine. Garrett, do you have any thoughts on this?

GARRETT GLASGOW: Sure. I'll just briefly-- I keep using the words easy cases and hard cases if we're talking about outcomes. I think in some cases, I wouldn't say easy, but it's easier to measure harm.

I've talked about market transactions with companies that don't live up to a promise. So say, for example, you do business with a company that has a certain privacy policy. They say, we will not share your data with marketers, but in fact, they do. And that's the extent of it.

Measuring the outcome there and measuring the harm that came from that outcome is pretty feasible. I think that's something that we can do. Sort of a benefit of the bargain-type argument. I paid you a certain amount. Some of that amount was for you to protect my data. You didn't live up to your part of the agreement. It's the same kind of techniques you would use for a defective product, say, in a class action case. My washing machine is leaking. One of the deals is when you buy a washing machine it doesn't leak, so you owe me some kind of compensation for that. One of the deals when I bought your streaming video service was you wouldn't share my information with marketers. You did, so you owe me some compensation for that.

Where it gets a lot harder is when we start to introduce a lot more uncertainty, which is, you didn't sell my information to a marketer. You lost it to hackers, and it's now on the black market somewhere. The possible harm that might happen, it could be well beyond the value of the product. I'm paying $10 a month for a streaming video service. If a hacker steals my information and drains my bank account, that's probably going to be a lot more harm than $10 a month.

Now I will say, some researchers have tried to measure this sort of harm by looking just at the value of the information, which is, how much can you sell somebody's personal information for on the black market? What does it cost to buy a social security number? There's actually prices out there. You can buy someone's information.

I don't think that's a measure of harm. I think how much a criminal would pay for a record to open up a fake bank account or take out a payday loan and then run away, that amount is much less than the harm caused to the consumer than the criminal is paying for. Just like if somebody breaks into your car, they break into your car and they steal your sunglasses, well, you have that loss, plus you've got to fix the window and your car's in the shop for a while, and so on. There's a lot of additional harm that's not measured by just looking at the price that that information can be sold at.

So that, I think, is the hard case, once we introduce this uncertainty. And we can measure things like the time spent changing records and changing passwords, and so on. But I guess I'll leave it there. There's easy cases and hard cases. And unfortunately, most of them are hard cases.

DOUG SMITH: Thanks. So Ginger, I'd like to get your thoughts on this. But also maybe if you can transition us into our next question, which is basically this question of causality. When information has been stolen, how can we measure the risk? And how can we kind of maybe sort of back out causality? What events may cause that? Or just get any kind of sort of way to clarify sort of the harm that happens in those kinds of cases.

GINGER JIN: Yeah. I started the panel by emphasizing the difference between an ex-ante perspective and ex-post perspective. In fact, estimation on an ex-post harm, I would say, is extremely important for us to understand the ex-ante expectation of harm, because we want to know, what was the financial loss of identity theft?

We also want to know, how many percent of people actually suffered that loss? How many did not suffer that loss? We want to know the distribution of those losses across the population. That would give us sort of a big picture distribution so that we can form an ex-ante expectation on this, OK?

But I would say that ex-post harm estimation is not enough. We also want to sort of quantify the risk. I talk about the [INAUDIBLE] panel to also talk about [INAUDIBLE] of risk that a firm's data practice would increase, right? If there is an increase of risk, we want to measure that. We also want to measure how many people got exposed to that risk.

So I just want to point you to two sort of research directions I see as quite promising in doing that. We know it's very hard to tie back to a particular firm's data breach or data practice.

One promising direction I have seen is a study by a group of researchers from Google, from Berkeley, and from the International Computer Science Institute. They actually started-- they actually studied the dark web, OK? They're passively monitoring the dark web for a year, and they were able to identify 1.9 billion usernames and passwords as stolen from the previous breaches. They're also able to, if you look at their paper, they're also able to give a list of the top 20 leakages, which identify the companies that got the data breach, and therefore got the records on the black market.

So this suggests there is a way to sort of link a data breach to what kind of records are on the market. Of course, we need other links from that record on the market to some real fraudulent transactions or identity theft or other sort of tangible or intangible outcomes that eventually happen to consumers. But I would say this is sort of a good one step forward to connecting the dots in the dark web.

Another study that was fascinating was about the blockchain. I think this might be a technology solution in the trackability of data flows. If blockchain can be used as technology to track bitcoins--

[LAUGH]

--ownership, maybe that technology could be used here to track sort of how the data changed from one hand to the other. I'm not a computer scientist. I don't know exactly how to do that. But I have seen other people, both computer scientists and economists sort of try to work in this area. To me, that seems a pretty promising direction to really do some research on.

DOUG SMITH: Thanks, Ginger. Lynn.

LYNN LANGTON: So both Ginger and Josephine made the point that causal ordering is really difficult to establish when you're talking about individual incidents, and I would certainly echo that. And it's something that we've wrestled with quite a bit with the supplement. Are there ways that we can more directly try to tie experiences of data breaches to individual incidents of identity theft?

And the reality is, I mean, when you're collecting survey data, the data are only as good as the responses you get. And we know, because we already asked respondents if they know how their information was obtained, that the majority of victims don't have any idea. So we have about 30% of our victims that say they have some idea, even if they're not sure, about how their information was obtained.

So using a survey like the NCVS to try to get at this causal relationship between data breaches and identity theft is really not the best vehicle, unfortunately. I think you have to look for sort of these other technological, computer science-based solutions.

I mean, among those that do say they know how their information was obtained, about 20% of our victims say that it was obtained through personnel or other files, personal information being obtained through a company that had their information. But again, that's such a small percentage of the 30% that knew how their information was taken that we can't really use that to draw any conclusions. The causal ordering issue is a big problem for that.

DOUG SMITH: Thanks, Lynn. Catherine.

CATHERINE TUCKER: Oh, no. I think, you know, Ginger and I completely actually agree in that what she was saying is that to measure a causal link, we have to have a better understanding of what actually causes injury to happen. And I think if you look at too much the research we're trying to come up with an average effect, rather than understanding what leads to those few incidents which are really bad that happen. So I think we're in agreement. I don't need to take any time.

[LAUGHTER]

DOUG SMITH: Great. Josephine.

JOSEPHINE WOLFF: I would just say, I think the flip side of that, which is also an important area for more research both from academia and from government and industry is, what prevents harm from happening? Which I think is also a pretty underdeveloped area.

To take two examples that you've probably encountered, one are the chips in your credit cards, right? We had a big liability shift in this country a little more than two years ago about the kinds of credit cards you use and implanting those microchips in there to help prevent future data breaches from being able to steal those numbers as easily. Another one you've probably encountered at some point is multifactor authentication, the idea that for some of your accounts now, you log in not just with a password, but with another code or with a tap on your smartphone or something.

And I think that trying to understand what the impact of those types of defenses actually is in practice, right? Does it actually cut down the harm? Does it cut down the number of data breaches? Does it cut down what the injuries look like when data is stolen?

I think we have a long ways to go still in trying to understand which of these different mechanisms actually work for preventing the types of harm we care about. And especially if we're moving away from being able to hold companies liable directly for things like identity fraud, because it can be hard to trace back to individual breaches, I think being able to have very clear expectations about what the ex-ante protections should be based on empirical evidence is going to be really important in that space.

DOUG SMITH: Thanks, Josephine. Garrett.

GARRETT GLASGOW: Sure. Just briefly, I think if we are going to understand the-- if we understand causality and understand the increasing risk from any particular data breach, we need to understand the baseline risk, and I don't think we have a good handle on that right now. The Identity Theft Resource Center has reported over 8,000 data breaches this year alone. More than 20 a day. Some are large, some are small. But there's this constant background of leaks of information out there. And perhaps becoming impossible at this point to know if your information shows up in some identity theft case where exactly it came from.

To apply this to my-- I had this analogy of privacy as Prince William Sound and the Exxon Valdez spills oil there. Maybe imagine Exxon Valdez runs aground, spills oil. The cleanup crews show up, and it turns out 20 other tankers ran aground there overnight and were towed away by the owners, but left all the oil behind. So how much damage does the Exxon Valdez do if it spilled oil where 20 other tankers have already dumped a whole bunch of oil?

Maybe that's a really negative viewpoint. But you know, we're talking about the intrinsic value of privacy. Maybe that value's been severely eroded by this sort of permeating background radiation of continual data leaks.

And that's one thing I think we need to get a handle on is, what's out there? Maybe in the aggregate we can see identity theft, increasing, but how much of this is due to breaches six

months ago, a year ago? We just don't know. So we don't have a good handle on baseline risk, and that makes it very hard to establish causality or increases in risk.

JACQUELINE CONNOR: This has all been incredibly interesting, and we have a few audience questions. The first one actually, I think, could implicate both consumer preferences and revealed preferences, or stated versus revealed potentially. So the question is, are you considering studying the increase in purchases of ID theft products or cost for freezes, lifting freezes, delays to obtain credit while waiting for a freeze to be lifted? And I don't know if anyone in particular would like to take a first stab at that.

DOUG SMITH: I think the question could just sort of be rephrased as, would that be a way to get some kind of measure of how much customers value or how much they're harmed from ID theft?

JOSEPHINE WOLFF: I think that's possible. I think one of the things that's complicated about purchases of those products is that most people who have them have them through a breached company. So the people who are usually the big purchasers for those are the entities that get breached and then purchase a contract for hundreds of thousands or millions of their customers. And it's then usually offered free of charge for at least a period of time to those individuals.

So there's some interesting work to do looking at how many people actually take the company up on that offer and how that's changed over time. But I'm not sure it's exactly an economic decision on the part of the individual consumer.

JACQUELINE CONNOR: And Josephine, I wonder if your answer changes when you're talking about personal cybersecurity insurance. I know you discussed it quickly at the beginning, but I know I didn't really know about it before I spoke to you for the first time, so maybe you can kind of explain it, and whether or not a measurement of that would really change your answer to this question.

JOSEPHINE WOLFF: I think it could. I think at this point when we talk about personal cyber insurance, we're talking about probably tens of thousands of people in this country. So it's a very, very small population of people who have actually purchased these policies. They're mostly fairly high net wealth individuals, so it's not something that low-income people are looking around to purchase, or even most middle-income people are looking at.

I do think it's a different measure, because it is something that people are paying for out of pocket. So it is them deciding, you know, this is a type of risk I'm concerned about, and I want to have some protection from all of these different ways of thinking.

But I think the other reason it's interesting is if you look at how those policies are being structured at least right now, they encompass a lot of things that I at least would not necessarily have associated as risks of sort of computers and cybersecurity. For instance, we're seeing families purchase insurance to cover home schooling costs in the event that their children are cyberbullied and don't want to continue going to school. So I think it's this very complicated set of costs that's totally dictated by what people want and are willing to pay for, as opposed to what we think the company should be responsible for protecting you against.

And that when you shift the question and you say, OK, what is it that we think people are actually willing to spend money on, it turns out there are a lot of different kinds of harm or injury that people might be interested in trying to insure themselves against. And that at least some people are willing to spend money on, which I think does give you a sense of what people really care about and want to protect themselves against. But again, at this moment, we're talking about a fairly small population.

JACQUELINE CONNOR: OK, great. And we have another question. And maybe, Catherine, given that you opened up talking about data breach notification laws, you might want to take a first stab at it. But the question is, can panelists discuss the multitude of state breach reporting rules and how that complicates setting a national standard of harm or injury?

CATHERINE TUCKER: Well, I'll just start off by saying, so as part of this study we did about hospital data breaches and data breach notification rules, we spent a lot of time trying to decode the different texts of the laws, and there's an amazing amount of variation. And the other thing I would say is that in my very lay opinion, many of the laws seem rather inconsistent if you know anything about technology in there were exemptions which make no sense, and it looks like, I don't know, people were just taking a random word generator sometimes to actually try and describe what they wanted to happen.

Now maybe this is an opportunity for something better. But certainly, I had a certain amount of disquiet, having seen the lack of standardization of language in these laws. And I'm an economist, not a lawyer. So that's probably a bad thing.

JACQUELINE CONNOR: Ginger?

GINGER JIN: Can I add? I think given the discrepancy we see across states, there's definitely a value to standardize the notification law, just to make sure that firms know what to do after a data breach.

But I want to ask probably a harder question, that we got a concern about what happened after notification. If the firms sort of meet all the obligations stated in the law and have disclosed the information they know at that moment. So what? We're relying on the consumers or the media or the public somehow respond to it so that they would feel embarrassed and therefore improve their data security? I mean, to me, that sounds like pretty wishful thinking.

As we know, if the harms we have in mind cannot be traced back to the individual firms, and they've already done their duty in notification, it's almost like, OK. I have done what I can, right? The rest is up to you. It's up to the consumers and vigilance of sort of having their own preventive measures or other things. I think that question has got to be coupled about the data notification law itself.

JACQUELINE CONNOR: Does anyone else have any thoughts on that question?

GARRETT GLASGOW: Just on reporting requirements in general, say, you know, of course we want-- if companies lose control of our data, of course, consumers need to know about that. But I think I have nothing in particular to say about any particular state's requirement.

But one thing that we have looked at at NERA is the correlation between media coverage and both class action lawsuits and enforcement actions. If your case goes viral, you can expect lots of class action lawsuits, and you can expect a lot of attention from regulators.

So I think one thing we might want to think about is, are we creating any perverse incentives with the way that these disclosure laws are written? Do companies want to try to game them, to try to do things to minimize media coverage or meet the letter of the law without providing any information that can be seized upon by the media?

It's something that [INAUDIBLE]. I'm not sure that we've created a perverse incentive. Maybe it's just one of those things that goes viral sometimes, or the biggest cases tend to pick up the most coverage. But that is one thing that I know a lot of people in the industry are aware of is the data breach is bad enough. But if the media gets wind of it and then the class action lawsuits and the enforcement rules, that's a whole bunch worse. So that's one to think about in terms of what kinds of reporting we require and how we require that reporting.

JACQUELINE CONNOR: OK. Thank you. Well, we have a few minutes left. So I think as a final topic, we'd love to hear what you all think should be the focus or the focuses of research on informational injury going forward. And I know you've sort of answered this in your answers already. But maybe specifically, what should researchers and the government work on to improve our ability to measure and assess informational injury? And maybe this time, we'll start at the opposite end of the table. So Josephine.

[LAUGH]

I'm trying I'm trying to think of a new answer.

[LAUGH]

Well, one thing that we haven't talked a lot about that I think is important for an agency like the FTC is thinking about sort of which injuries people have some protection from, right? And so back when data breaches were mostly about payment card information, we had a lot of policy protections in place to say, if somebody's using your credit card fraudulently, you're not going to be liable for at least most of those charges. Usually any of those charges because your bank and your payment card network are going to want to keep your business. They're going to want to cover that for you.

And now as we've seen, say, a strong shift towards ransomware and other types of cyber crime that are hitting individuals more directly, that they have less insulation from the direct economic costs of, I think there's an advantage there in that sometimes it's easier to calculate those costs. Because if you're talking about something like ransomware, you can put a price tag on it more directly.

But there's another question that comes up about, how much do you sort of weight the fact that individuals are going to be bearing those costs directly? And does that change the kind of injury or the kind of remedy that you want to have in place, depending on why that ransomware infected my computer in the first place?

And was it my fault, or should my ISP have had some inkling of that? And how the question of sort of distributing the costs in that way is going to change the incentives of the large, centralized, powerful actors who might have the most potential to implement widespread security controls or defenses.

JACQUELINE CONNOR: Thank you. Catherine?

CATHERINE TUCKER: Well, so if I go to say what you've got to research, which is just like a wonderful thing to be asked, I think I would flip the focus of research in that there's a very natural tendency to try and think, well, how can we measure most accurately an average effect?

I might encourage you instead to focus on the question of, how can we identify the occasions or incidents or people where there is no injury, where there is no effect? Because I worry that we may be going towards a role which is very natural to try and understand, how do we measure an average, from average baselines, and so on? But I think this is such an intriguing field, because so oftentimes we don't see injury. And trying to understand and pinpoint those occasions seems to me incredibly both exciting and valuable.

LYNN LANGTON: An interesting perspective. So I think there have been a lot of issues raised today about how we can measure different aspects of data security and the harms associated. And in some ways, the NCVS, the National Crime Victimization Survey is limited in that it can't address these full range of topics that we would want to measure, the harms that we would want to measure.

But I think as a government researcher, we want to try to measure as best we can we can measure. And the thing that we can do is look at level and change over time in terms of identity theft, which is obviously a direct harm. And then we also have another supplement that I'll just make the pitch for that we just implemented-- it's in the field right now, looking at financial fraud, which is another potential outcome of a data breach.

And so just to be able to track over time whether we're seeing any changes in the prevalence of these particular types of outcomes that we know are correlated, though we can't look at the direct causation and the direct causality to data breaches, but we know that there's still a correlation there. And to see if the risk of experiencing some of these outcomes and the nature of some of these outcomes is changing over time.

And so that's really what our focus is right now. We'll be putting out data from our 2016 Identity Theft Supplement in January. And then also in January, our 2018 Supplement goes in the field. So again, we're trying to measure this consistently over time so that we can track those trends.

GINGER JIN: Well, one thing we haven't touched much in this panel, but has been touched in the previous panel is some similarity between the problem of privacy data security and the problem we have seen before in, say, food safety, drug safety, product liability, and [INAUDIBLE] laws. And so I would like to see probably more interdisciplinary research to sort of summarize the lessons we have learned from those areas and to see to what extent we can sort of apply the insights we have learned those to the market of privacy and data security.

GARRETT GLASGOW: And I think an important topic that isn't well understood yet that we should push forward on is maybe a theoretical or definitional issue of what we mean by informational injury. Does informational injury spring from the content of the information itself? Or is it how it's treated by, say, a company that you're doing business with?

And so here's a thought experiment. Suppose you're doing business with a company and they have a data breach, and some of your personal information is stolen and is now out there on the black market. But then you find out the previous week, there was another data breach with a different company, and all that same information was already out there a week ago. Were you harmed by that second data breach or not?

And whether or not you were harmed might depend on what you think of is informational injury. Is that the fact that this information is now out there? If that's the harm, the harm's already done. Or is it, this company is mistreating customer data and is not being fair to customers, it's dealing unfairly? Maybe that's a different kind of informational injury, and in that case, you could be harmed twice by the leak of the same information.

Both of those things could be sources of harm. It's entirely valid to believe they're both sources of harm. But I think that's something that is important to distinguish when we think about what kinds of enforcement we want to do and what kinds of harm we're trying to measure.

JACQUELINE CONNOR: Thank you, everyone. And it looks like we're pretty much out of time, unfortunately. But thank you to our panelists for such a great discussion, and a wonderful way to wrap up the day. Andrew Stivers will be coming up shortly to give the closing remarks, so thank you.

[APPLAUSE]