

[SQUEAKING]

[RUSTLING]

[CLICKING]

**FRANK**

All right, I'm going to get started. This is lecture 14 of 14.13 The lecture today is about attention. OK. Good.

**SCHILBACH:**

So where are we at in this class? In this class, we talked a lot about preferences so far. Think about preference describing in economics. What do people want? What makes them happy in one way or the other? We talked about time preferences, risk preferences, and social preferences, time preferences. How do people make choices over time?

Risk preferences are, how do people deal with risky situations when some things are uncertain and others are more certain? Or sometimes good things happen, sometimes not. And there's a probability distribution over those events. And how does that affect how you make choices?

And social preferences is what we talked about during the last few weeks, which is how does the presence of others affect our choices? How much do we care about others? And do we decide or make choices differently if others are involved?

So now what we've done with preferences is sometimes saying we have our models to think of ways to think about what makes people happy. The second part when you try to make decisions is when you look at your environment, what are prices like? What do we know about others? What are our beliefs? And what do we attend to? And how do we learn when we get new information? So what information do we have? And what information do we attend to? And how do we learn when we get new information given to us?

And so we're going to talk first about attention in this lecture. And then the next couple of times, we're going to talk about beliefs and learning. So what we're going to do today is we're going to talk about three broad issues. One is going to provide you some motivating evidence about attention. We had a little bit on this in lectures 1 and 2. But I wanted to remind you of that a little bit.

Then we're going to talk about two specific papers. One is a pretty famous paper by Raj Chetty and co-authors on inattention to taxes. And then you're going to talk about a second paper, which is learning by noticing. In particular, we're going to ask the question of, how is it possible that people do not attend to things that are important in their lives?

You might say, well, if attention is limited, that's all good. But then we should expect that people attend to really important stuff. So some sort of rational inattention models would say, well, it might well be that people are inattentive in some ways. But really, it can't be possibly that the welfare losses from inattention are large because if they were large, if there's things that people missed, then they would pay attention to those things and shift their attention optimally. The learning by noticing paper is one theory, at least potentially, that addresses this question.

The way these movie perception tests work is, essentially, they look for what's called change blindness. So the way this works is, often, there's one actor, which is this guy here. I think you can see probably right now there's some guy who does some activity. You're supposed to watch the guy.

And then there's a second scene in some ways where the person does something else. In this case, the guy-- sorry, where is he? The guy is taking a phone call. And so what happens then is-- essentially, then, this guy that you see here that takes a phone call is a different person altogether. And so that's called change blindness where, essentially, people then tend to not notice at all that there's two different people in a scene.

There's other types of experiments that do this, as well. One version of that would be people who go to a bank. There's a bank teller. And this is like all psych experiments. There's a bank teller. Here's the other guy. Here's the other guy.

So you see this guy. He looks actually very different to the other guy that you had seen before. So people go to the bank. There's a bank teller. The bank teller says, oh, yeah, I'll have to just get something from the back. And then another person shows up who's a totally different person, often looks very different, sometimes wears different things. And people just don't notice at all. And that's sort of type of experiment is called change blindness.

I've already shown you the gorilla test before. I'm going to not play it again because I've already seen it. But essentially, it's the same where people are asked to pay attention to how many passes people play in a basketball game. And then there's a guy who's in a gorilla suit who shows up and walks through the whole screen. And many people just miss the gorilla suit.

Now, what's going on here? There is a large number of these types of experiments. They're called change blindness or inattention experiments. And there's another version of that, which is called the dichotic listening experiment. This is going back to broadband, 1958.

They essentially look something like this where there's people who wear headphones with messages. On the one hand, on our left, there's supposed to be ignored inputs. And on the other hand, there's supposed to be attended inputs.

So the person in the experiment is asked, only listen to stuff that's, I guess, in your left ear. These are the attended inputs. I'm going to ask you some questions about that. And then what you're supposed to then do is-- are you still stuck? This is bad. One second. I'm going to unshare my screen. Exactly. Give me a second.

So the way this works is you would have headphones on. And on one hand, you hear inputs that you're supposed to attend to. On the other ear, you hear inputs that you're not supposed to attend to. And usually, then, you get explicit instructions to attend to the message in one ear and then asked later about message in the other ear. People can't remember it.

And then what these experiments often are-- when asked to keep a number in their head, people remember the played message much less. So essentially, what this shows is that, A, I guess, people are pretty good at shifting attention in some way. But attention is limited to the extent of when you're also supposed to attend to something else at the same time. It's essentially too much going on, and it's really hard for you to remember everything.

And so to some degree, it tells us, A, I guess, attention is limited. We tend to focus-- when instructed, at least, we tend to focus on stuff that's important, right? So like in the broadband experiment, if I tell you listen to the stuff in the right ear, you actually focus your attention to the stuff in the right ear. You miss the stuff on the left ear, so you can sort of direct your attention to stuff that you think is important or you've been instructed to focus on. So that kind of works.

But at the same time, attention is limited in the sense of you can actually not remember what's in the other ear because you didn't attend to it. So in some sense, it's not like you automatically remember stuff. You have to pay attention to things.

And then at the same time, when we now exogenously reduce people's attention by telling people to keep a number in their head, distracting them in some way, people are much more-- even playing the played message, the message that you're supposed to remember. I'm sure they don't remember the other message, either.

OK. So then in addition, which I already alluded to, attention is limited. There might be different factors that might affect people's attention. So what factors might be important for attention? Could be your physical environment in some ways. It could be it's really hot. It could be it's really cold. It could be other threats in the room, so if you're in some place where you're worried that there will be wild animals coming anytime soon.

Physical environment could be really important. It could be really loud. Distractions could be social media. It could be other stuff that you're just trying to do at the same time. Maybe you're trying to solve some problems right now. Maybe you're trying to prepare for some-- for example, I had a phone call just before our class. And I was also trying to think about what to do in class, and it's just really distracting to have other things you're thinking about at the same time.

Could be worries about your own or others' well being. It could be sleep or what people were saying already, people's physical environment. We're going to talk about some of this in the poverty lecture. So there in particular, one thing that poverty tends to come along with is sort of a bunch of different ways of-- in addition to not having money and many other deprivations, poverty tends to come along with lots of other factors that might reduce people's attention.

And that's to do with sleep. It's to do with exposure to noise. It's coming with worries about people's health and so on and so forth.

There's some question about practice. Can you actually practice attention? And there's some people who would argue, at least in education, maybe at younger ages-- suppose you sit kids down and have them practice focusing on tasks like multiplications or whatever and do math problems. Some of this, some people would argue, in addition to just helping people learn math, actually the numbers and so on and how to multiply, you might also actually just help students learn how to focus and how to pay attention carefully to things and particularly sustained attention and not get distracted by other things.

Other distractions, by the way, also, like temptations, for example. So if I put a delicious chocolate cake in front of you and have you-- and if you're really hungry, that might be actually really hard to pay attention to whatever else you're doing just because that's really distracting to you.

OK. And so we're going to, for now, keep thinking about attention being fixed. We're going to get back to the issue of attention being malleable and potentially being depleted by issues such as poverty at the end in the very last lecture. But for now, we're going to look at-- think about attention as fixed but limited. And now we're going to look at, what are the consequences of that?

OK. So then one question you might ask is, well, this listening task that I just showed you or when a monkey walks through some video or the like, it's a pretty low-stakes situation. Who really cares? And one Chicago view of economics would be to say, well, these are just low-stakes situations. But once stakes are really high, then people will surely pay attention.

And so there's a number of different papers that look at this issue. But one quite nice example here is from the stock market, arguably an event of high economic importance because people make a lot of money, potentially, from it. So there's this very nice paper by Huberman and Regev from 2001 which has-- it's essentially an event study and a case study of stock market development of a company called EntreMed.

And this company in 1997 on October, November has very positive early results on a cure for cancer. Now, for any of these types of companies, having new cures and new medications or vaccines or the like are hugely valuable because they can sell a lot of their product in the future, which means, essentially, positive news about that means you'll have lots-- at least with high probability or higher probability-- large profits in the future, to the extent that the stock market is an indication of that. You might expect the stock market to go up.

Now then, November 28, the journal *Nature*, which was a very highly prominent scientific journal, features this new very prominently. And *The New York Times*, in fact, reports on it on page A28, which is very much in the back of the paper. And on May 3, this is a whole six months later. *The New York Times* again features essentially the same story as on November 28 on the front page. OK.

Now, then on November 12, 1998, the *Wall Street Journal* front page then reports about a failed replication. It's essentially to say, well, these are promising new early results. These early results are often small samples. And it's not quite clear, does it replicate really with larger samples and so on. Is this really a good cure? And so you kind of won't have replications. And really, only if you're able to replicate your results is this actually something that's useful and that you can make money off.

Now, what happened to the EntreMed stock prices? And so the first question you might say is, well, in a world of full attention with unlimited arbitrage, as in if people took full advantage of the news and reacted perfectly, what would we expect? And second, in reality, what did actually happen?

So let's first ask the first question about, with full attention, what would we expect? I guess there's a bit of a question what happens in the first one. There's a bit of a question, is there insider trading? Or who gets the news in the first place? So one thing you might say is, well, maybe the first one, only the company knows. And maybe that's not public knowledge. And then if there's insider trading, maybe the stock market might go up.

But surely, like in number two, if *Nature* prominently features it, there's going to be a bunch of traders, a bunch of people who invest in this company or not. If they pay attention, they should really know what's going on, and they should really-- the stock market should go up.

Conditional on that, on the November 28 news, on May 3, there's really no news here, right? It's kind of like we knew this before. If everybody was paying attention before carefully, the stock market should have incorporated. The news is out. Traders know. You should buy this, or stocks at this company is going to be more profitable. So stocks should go up.

So really, if there's full attention, you expect nothing to happen on May 3, '98. And then if there's a failed replication, well, to the extent that this means this cancer cure is entirely useless, the stock market should essentially just go down to where it was before.

Now, but what about in reality? So let me actually just show you. So here's the reality of this. So there's not even - there's the previous events done in here. This is essentially the first story that comes out. You see a little bit of an increase here. But actually, quite a bit in relative terms with something like 5%. It seems like people just didn't really quite notice what was going on.

And there's also no trend upwards. It's not like the news is spreading across people and so on. And then there's like a huge jump now in *The New York Times*, right? So this is essentially here, it's in *The New York Times* but in the back of the journal. Up here, it's essentially *The New York Times*, the front page. And suddenly, lots of people pay attention to news that's actually not news because that was already in *The New York Times* before.

To the extent that everybody pays attention to this, unlimited attention and unlimited arbitrage, we should essentially expect no effects whatsoever. So we see a huge spike in the stock market price. Maybe the stock market overreacted a bit and went down a little bit. But then what we see here is then the negative effective of the failed replication.

Notice that here, the stock market is still quite a bit high. The price is still up quite a bit high, maybe twice as high, about 20 compared to 10 as it was before. It's a little bit hard to interpret that because it could just be that maybe there's still a chance that the cure might work out. Maybe there's other news about the company in the meantime that are positive and so on.

But the key part here is the fact that the stock market went up here on May 4, 1998, really is saying it's news for lots of people. And really, it shouldn't be news. And there shouldn't be any reaction because, essentially, it's old news, as you would say, in the sense of people in 1997 in November 28 already reported all of that. Does that make sense?

OK. And there's lots of examples of these kinds of things. But essentially, some of this is-- in attention, some of it is also confusion. I don't know if any of you have heard of what happened to the Zoom stock market price, stock price. I don't know if any of you have heard of this story.

So essentially, there's a company Zoom, which is the company that we're using right now to record this video. The stock market price of Zoom has gone up a lot during the last three months, rightly so because Zoom now has-- Zoom's profits has gone up a lot. For example, MIT now has a corporate license of Zoom, which essentially means thousands of people now are using or can use Zoom, and that's an expensive thing to do.

It turns out there's another company that's also called Zoom except for that it's a different company and has nothing to do with the actual Zoom technology. It turns out that the Zoom stock market price for this other company went up a lot, as well. The ticker for that company is actually ZOOM, so like Zoom, as opposed to the actual Zoom company that we're using right now. Their ticker is ZM.

And so at some point, actually, the trading for that other company was halted, presumably because people were very confused, and they were not really paying close attention to what is the right ticker, what's the right company that they're actually buying. They thought they were buying Zoom, like the actual online video conferencing company, while in fact they were just buying some other company that's actually not even doing anything anymore.

So there's lots of stories about people being inattentive even in situations when stakes are really, really high. Now, then, you might say, well, how do we now measure? So now in some sense we have some evidence here that people are-- in some sense somewhat informal evidence that people are inattentive. Now, how would you measure the impact of inattention overall?

So if you wanted to just measure and demonstrate that people are inattentive, in addition to what I've just shown you, what might you do? How might you do that? There's some information that you usually get in the world and they should have. One example that we can discuss is essentially prices or taxes.

And now if you want to demonstrate that people are actually missing those prices or taxes, what you can do is now you can make it very salient. You make it really apparent to people. And then you can look at, when you look at changes in prices or the change of making things from non-selling to selling, how do people react? Or maybe you make, for example, taxes very salient. What happens to people's demand for certain products?

And if then making things salient changes people's behavior, that then identifies underlying inattention. So if you just-- people don't have access to certain information, in some sense, that's less interesting because that just means, well, I don't have the information that's really available to me.

But in some ways, I could make information that's really very salient-- I essentially could reduce the salience, as you say, of certain types of information-- again, information that should be available to them. For example, suppose I do sales. And I'm trying to, as a company, sell things.

You go to a store. And usually, sales are made very salient to people. Now you could essentially just change the prices without making these sales very salient. And then you could look at, how attentive are people to that?

In a way, that's, in some ways, less natural, right, because when people are-- in a way, what we want to demonstrate is that in people's real lives, there's information available to them. And they're not really paying sufficiently much attention to it. So if you can then change that in some ways and demonstrate that now they're changing their behavior, that must mean that previously, they had been like misoptimizing.

If instead you did something like if you reduce the salience of certain types of information, that's, in a way, somewhat less interesting because it would say, well, it doesn't really mean that they were inattentive to start with. It's more like you're hiding something from people, and now they can't find it. That's, in some ways, somewhat less interesting. But I think in principle, you'd also, exactly as you say, identify inattention.

So there's a few type of things that people have done. And to be clear, there's often a pretty fine line between attention and memory. So one thing we could do is-- and this is what Maya was saying previously-- you could make certain features, like taxes, very salient. That's what we're going to look at next.

In addition, you could provide some information when the correct response is known already. That's one we're going to talk about afterwards is the study by Hanna et al. So that's kind of like when people have certain types of information already but then providing that information again or giving that information in some sort of concise form and potentially notifying people of the fact that this information could be or should be important. And then we now see people react to it. That doesn't mean that that must mean that, before, they were inattentive.

Another thing that people have done a lot is reminders. Reminders often are sort of like-- you can think of the pieces like memory issues, and people just forget things. But in some sense, memory issues and attention issues are very closely linked. If you forget something, then you don't pay attention to something.

So if I remind you of something, now you're paying attention to, for example, your savings or your medical adherence, like you should take certain drugs and so on. There's a number of studies that essentially provide reminders to people. And if these reminders have effects, then that must mean that people were not paying attention to those kinds of things, often presumably because of memory problems. They just forget sometimes.

OK. So now let me start with a very simple model to help us understand the results from the Chetty et al. paper. It's very simple and very sort of just sketched. But I think it's actually helpful to see what people are doing.

So consider a good with a value of  $V$  inclusive of the price, which is the sum of two components. The components are little  $v$  and  $o$ . So there's visible and salient component  $v$ , and there's an opaque component  $o$ . OK.

And so if you're now inattentive, the consumer, inattentive consumer perceived value-- instead of perceiving the true value  $V$ , the consumer perceives the value of  $\hat{V}$ , which is little  $v$ , which is a salient component. So even if you're inattentive, the assumption is the salient part of the good you always see. For example, what color is the good or the like? You always see that.

And in addition, the opaque component, the stuff that's not salient, you're only going to pay attention to the fraction of  $1 - \theta$ . So the degree of inattention-- so  $\theta$  is our inattention parameter. And  $\theta$  measures how much are you paying attention to stuff that's not particularly salient.

So for  $\theta$  equals 0, you're back to the cases before. For  $\theta$  equals 0,  $\hat{V}$  equals  $V$ , which is essentially just a standard case. If  $\theta$  were 1-- so  $\theta$  is supposed to be between 0 and 1. If  $\theta$  is 1, then essentially you're not paying attention at all to the opaque component. And anywhere in between, essentially, then, you're paying only partial attention to the opaque component.

So the interpretation is each individual essentially sees  $v$ -- sorry, sees  $o$  to some degree but processes it only partially to the degree of  $\theta$ . Of course, whether you actually see it and not process it or whether you miss it entirely is, in some sense, a philosophical question. But think about like, everybody has, at least in principle, access to the component  $o$ . But they only process and pay attention to it only to some degree.

OK. So Chetty et al. applies this model essentially to taxes. So one very interesting feature about some taxes is the fact that sales taxes are only added at the register, right? At least in most cases, you would shop in the store. And then the sales tax, only at the end once you actually go to the store, will be added.

Now, people are not-- people know. When you ask people directly, what is the sales tax in your state, people actually on average are pretty good at knowing this. So it's not like people don't know that there's a sales tax. But they might just forget that there are sales taxes for goods.

So now what you can do now is you can compare the demand response to sales tax changes versus the demand response to other price changes, right? So if you see prices fluctuating in a store over time or across goods, presumably, everybody sees those price changes overall, if these are price changes that are not related to sales taxes, the reason being once you're in the store, you're going to look at prices. You're going to see is the price high or low even before you go to the register. So presumably, you see the price tag pretty much most of the time.

In contrast, when you look at sales tax changes, if the sales taxes go up for some reason or if there's a sales tax for some items versus others, you might just miss that at the time when you're choosing your goods before you go to the register. And once you're then at the register, you might not even notice that you have paid the sales tax. Or you might be surprised and just don't change your behavior any more but at least notice it.

So now what did Chetty et al. do? They have data on the demand for items in a grocery store. And they have essentially the demand,  $D$  of  $V$  hat. Remember,  $V$  hat is the function of the perceived value, which I showed you here before. Let me go back for a second.

So you see  $V$  hat here is  $v$ , which is the salient component, plus  $1$  minus  $\theta$  times  $o$  as the opaque component. So the demand for goods depends on the perceived value,  $V$  hat, that people have. And there's a visible part as the value  $V$ , which is inclusive of the price. So when you look at when you're in a store, you look at the good. You see the good. You see whether you like it. You see the brand. You see the color and so on and so forth.

You also see the price, so the valuation that you give or the value  $V$  that you draw from that is essentially telling how much they want to have this thing minus the price that you see, which goes into the visible part of your valuation. And then there's a less visible part, which is just stuff that's essentially hidden. And there's a bunch of different parts to that that you might not pay attention to, which could be ingredients or the like, or was it organic or not and the like.

But in particular, the sales tax, right? So if you think about your valuation, you would have to essentially, however much you like the good-- you have to subtract the sales tax because you have to pay taxes on that good. And that's less visible to some people. So you might not pay attention to it.

So in this case, it'll be  $V$  hat is  $v$ , the visible component, plus  $1$  minus  $\theta$  times  $o$ . And that's just repeating what I showed you before. And that equals  $v$  minus  $1$  minus  $\theta$  times  $tp$ , which is the sales tax. And now to the extent that people pay attention to  $\theta$ , the sales tax might matter more or less.

There might be other opaque components that are unobservable both to the experimenter but also to the person. They can abstract from that. So for now, they can essentially say the less visible part is only the state tax. OK. Any questions on this?

OK. So now, and this is sort of like-- so now, essentially, what this gives us now, we can essentially try to see if we have this-- how can we identify this parameter  $\theta$  by making things very salient? And so the idea is now to make the tax fully salient. Let me show you how this looks like in practice.



Here's the math. But I'll go back to this in a second. Let me just show you what this looks like in the experiment. So the idea in the experiment is to say-- you see on the lower part here, you see a normal price. You see an original price tag here. This is what the price tag usually would look like.

So this would be an item. I guess this is brushes. This item would cost \$5.79 usually. And then what they did in the experiment, they essentially added a tag where they said, now here, \$5.79 plus sales tax is \$6.22. And now they make it entirely salient that there's a sales tax, and you make the true price very salient.

OK. And now the question is if you do that, if we go from theta being, for example, say 5.5, you only partially observe the sales tax. If you go from that to making the sales tax fully salient, what happens to people's demand? And can you infer from that theta, what theta actually looks like?

So let me go back to this. So essentially, what we're trying to look at, we're trying to look at a change in demand when theta falls to 0. Why does theta fall to 0? The assumption here is that whatever is theta people have before, so you might partially or even not at all-- theta might be 0.5 or 1 or whatever for people.

But now the intervention that I just showed you brings theta down to 0, right? Because now it's fully salient to people what the sales tax is. And so the question is now, can we have an expression for the change in log demand when theta falls from whatever it is before down to 0?

So now what we have is the change in log demand is the demand that we had before. This is what I showed you before. This was essentially demand as a function of  $V$ . Remember what I showed you here before. Give me one second. I'll go back.

So what I showed you here, demand is a function of  $V$  hat. And  $V$  hat is this animal here. It's  $v$  minus 1 minus theta times  $t_p$ . Now, the question is, what happens to demand when or log demand here-- forget about the log. It's not that interesting. That's just for math. So here's the demand as a function of what it was before. And here's demand when I now make things very salient, right?

So now here in this case, theta equals 0. And here, theta is whatever it had before. Now when you have this expression, what you can do is you can linearize that, take the derivative and say, what is this difference? Well, this difference is-- and you can do the math. It's essentially theta times  $t_p$ . And then here, this is the derivative with respect to theta.

And what this gives you essentially is that the change in log demand is a function of theta, a function of tau, which is like the tax, and eta, which is the price elasticity. So essentially, that's to say if you have now an intervention, we can look at, how does demand or log demand change when we do this intervention, assuming that this intervention brings theta from whatever it was before down to 0?

We can look at now what happens to demand. That's delta log demand here. That's just a definition. And you can have that. That's an expression of how we can relate that to theta to  $t$ , which is like the taxes-- it is at something like 7%-- and then eta, which is the price elasticity of demand.

So that is to say, if we have this change in demand that we observe from changing theta down to 0 and if we had the price elasticity from just other variation in prices that we see in the world and if you have  $t$ , which is just how much the tax is, that would allow us to identify theta. And that's what the experiment is trying to do.

So that implies essentially-- so we can just rearrange this. We can just say  $\theta$  is, just sort of rearranging things, minus the change in log demand from this intervention divided by  $t$  times the price elasticity. OK, and that is just to say essentially what we're trying to see is, intuitively, that's just to say once we make things very salient going from things not being salient to start with to making things very salient, how much does demand change there?

And then we can look at, well, for-- in this case, I guess, this is the taxes made very salient. And then we can look at, well, how does this compare to other changes in prices, assuming that essentially people are paying attention to prices anyway? So if you look at things becoming more expensive by 1%, how much are people changing their behavior there?

And now we have to take into account that what's being made salient here is the tax  $t$ . And from backing out the relative changes in demand for usual price changes versus making things salient, we can now try to measure  $\theta$  as how much did they pay attention before.

So one way to say this is for example, if we make things very salient and the change in log demand did not-- there's no change in log demand, what does that imply for  $\theta$ ? What would we learn from that? So what was the change here in demand at 0? People did not change the demand at all when you make things very salient versus what it was before.

Or put differently, if  $\theta$  equals 0 here, then these two expressions are the same, which essentially means that this thing here is also going to be 0. So essentially, to the extent that our intervention induces large changes in demand, that must mean that essentially, people were very attentive to start with. Essentially, if I'm making the prices very salient and nothing happens, well, that means people were already paying attention before.

Now, in the maximum case, if people were paying no attention to prices before, what would happen then? Well, in that case, I guess-- oops, sorry. In that case,  $\theta$  would be 1. So essentially, it would just be going from-- we would look at what's the demand or change in demand from going  $v$  minus  $tp$  to  $v$ , which is essentially just a change in  $tp$ . So that would be essentially say if the tax was 7%, now we just have to look at what is the equivalent effect of a 7% increase in the price, which is exactly what we have here.

So essentially, just to summarize-- and I think the math is, in some sense, not that interesting. But essentially, to summarize, it's like the larger our reaction is to this intervention, the more inattentive most people have been before because if they're inattentive to start with, they react more to making things more salient. So let me show you what this looks like in practice.

Again, the goal here is to estimate a change in demand from making taxes fully salient. Again, here's the original tag, and here's the experimental tag. It's very, very salient, very, very clear, and very hard to miss for people. And the assumption essentially is that people pay now full attention to this new tag.

They have a three-week period in which-- and this is an experiment in California. They have a three-week period in which they modify price tags of certain items so that they go essentially to different stores. They pick one store. And then essentially, they modify the price tag only of the subset of randomized goods. And they make the after-tax price salient in addition to the pre-tax price, right? The pre-tax price is already very salient here. In addition, they make the after-tax price very salient.

And now they're going to do what's called a triple-diff estimate, which is, in some sense, a very funky or fancy word just to say what we're going to do is we compare the sales or the demand during the treatment periods to the following. So there's a treatment period. We compare how much of certain goods have been treated. How much do the sales change compared to the previous week sales for the same items?

Second, we're going to compare them to sales for items in which the tax was not made salient. So there's some goods in the store for which the tax was not made salient. And then finally, there's going to be control stores. So we're going to look at essentially, in the stores that receive the treatment or in the store that receives the treatment, what happens to the goods that receive the treatment compared to the goods that don't receive the treatment?

We compare that to demand before the treatment was enacted for both of those types of goods. Essentially, what's the relative change in demand for the treated versus the non-treated items before versus after? So that's essentially what people would call a diff in diff, like difference in differences. And in addition to that, they compare that difference in differences to the same difference in differences in the control stores. So let me show you exactly what this looks like.

So this is a bit of a messy table or a table with a lot of information. But it's in fact quite simple. So what we have here is at the top is a treatment store. And at the bottom, there's some control stores. So in the control stores, nothing has been changed. So there's no experiment. There's no additional tax and so on being added.

In the treatment stores, there's going to be control categories and treated categories. Notice that they didn't fully randomize all the goods. So it's not like if there's one brush next to another, one brush has taxes versus not. Presumably, because they're trying to minimize spillovers in some sense, if you only make this tax salient for one good that's next to another good, people would get very confused because they're like, well, why does this one have taxes made salient and then the other one not?

So now this is across categories where people maybe only pay attention to certain goods in certain categories to start with. So what we're seeing here is now here, these are the control categories and the treated categories. And we can look at what's called the difference in differences. We can look at the change in demand in the treatment stores.

What we're going to look at is essentially the change in demand in the control categories and the treatment categories. So what we see here is we can compare how much is sold in the experimental period, which is this row here, compared to our baseline. What we see is essentially demand in the control categories was going up a little bit by 0.84. That's just because of seasonalities. Maybe people were just buying more anyway for whatever reason.

And in the treated categories over the same period, comparing essentially the experimental periods to the baseline period, demand went down. So now the difference in differences estimate is essentially the change in demand in the treatment stores, comparing the treated categories compared to the control categories, which is just the difference between this column and this column, which is minus 2.14 units. OK.

So that's essentially, how much did the demand change before versus after in the treated categories compared to the same change in the control categories? So that's essentially a difference in differences estimator. Now, they also have the control stores. Why would we want the control stores, as well? Why is that helpful to have?

Suppose that improved sunscreen, and now it's getting warmer. It's really hot, and everybody was out buying sunscreen. Then you see essentially people's demand in the control categories. The treatment category goes up by a lot or goes down by a lot for whatever reason. Suppose it's umbrellas, and one week was really rainy, and the other one was not rainy. And we see now essentially seasonal or other differences over time in some of the categories that would, even if it's randomized, it could just happen to be the case.

So that would not have to do anything with a sales tax being salient. That's just by chance. So now it's very nice to have some control stores where we can look at like the same difference in difference estimates and some of the control stores, in some sense, like a placebo. They can do the same things. There's going to be the control categories and the treatment categories.

Notice, to be very clear, in the control stores, the treated categories do not receive any actual treatment. So we can do the same difference in differences. And what we see here is now that there's essentially no difference here. So the demand in the control categories went up a little bit for whatever reason in the treated.

And this is a little bit smaller than in this other store. This difference is not statistically significant. But anyway, the control categories go up a little bit. The treated categories go also up a little bit. When you do this difference in differences estimate, you find essentially no difference.

In the control categories, there's no change in the treated categories compared to the baseline compared to the control categories. So difference in difference estimate in the control category is essentially 0. And now what you can do is essentially the difference in differences estimate. You can compare. You can subtract this number here minus this number here, which gives you the triple-diff estimate, which essentially is the relative change of treatment versus control before and after and treatment versus control stores.

It's a little bit lots of differences in your head. But I hope this comparison from the table makes it clear. Any questions on the estimate or the procedure that they're doing?

OK. So what we see here now is essentially now making these taxes salient in the treated categories in the treatment stores reduces demand, right? Because essentially, demand goes down. Why does the demand go down? Well, the perceived price that people face has now gone up. And now they buy less of these goods, presumably because before, they were not paying attention. Now the price goes up. And now, essentially, demand falls, as it should, because now things are more expensive.

And this is here the result. The average quantity sold decreases by 2.2 units-- this is units of the goods-- relative to baseline level of 25. That's an 8.8% decline. And now we can use our own formula and compute the degree of inattention. Essentially, now we can say, OK, now we know what the tax was. The tax was 7.375%.

We can estimate the price elasticity in some other ways. Essentially, we can say, what happens if prices go up by 1% for other reasons? What if the salient component that's already salient goes up? This is week-by-week variation across stores.

So if the prices go up, what's the price elasticity? And the estimate of those price elasticity is also given from other variation. That's 1.59. And now we can essentially back out theta. Essentially, what we're trying to ask is we have a change in 2.2 units that is making prices salient. Now the question is, well, how large of a change is it relative to a change in prices of 1%, 2%, and so on?

We know that essentially the price elasticity is 1.59. So we can essentially just back out theta. And so what we get from this formula now is that theta is about 0.75, which says that consumers react to price changes due to sales tax changes only a quarter as much as to other price changes. So essentially, consumers are very inattentive, in fact, to sales taxes, to sales tax changes. In fact, they only perceive a quarter as much as any other price changes in the world.

And that's really saying, look, people really seem to be missing these things that are not salient. Now if you make them salient, people change their behavior. Therefore, they must have been inattentive to start with.

In addition to that kind of evidence, Chetty et al. also have some evidence on using non-experimental variation. And in some sense, you could say the evidence that I just showed you from the store is a little bit weird because, in some sense, you see these different tags. And maybe the tags raise attention to the price even regardless of the tax. And it's a little bit weird in some ways.

So it's quite nice to have some complementary evidence of a non-experimental panel deliberation. And so what they do is they look at beer consumption. And it turns out beer, tobacco, and other types of goods have two types of taxes levied on them. They have an excise tax, which is included in the price. This is highly salient during choice the process because essentially it's in the price already. So if taxes go up, essentially, the price goes up. And people will notice if they pay attention or since they pay attention to prices, since they like the price tag already.

And then there's a sales tax. This is the tax that we were just talking about before. That's essentially very opaque during the choice process, as we just showed. And now we can look at variation across states and over time of changes in excise taxes and sales taxes.

And we can look at, how much do people react to excise taxes relative to how much people react to sales taxes. And that relative reaction will tell us, again, something about theta. And what Chetty et al. finds in here is a table for that. Essentially, when you look at changes in excise taxes, this tax is very salient. The price elasticity is above 0.8 or 0.9.

So essentially, when the price goes up by 1%, people react by about almost 1%. In contrast, there's some differences in estimates. But once you control for things properly, people essentially do not react to the sales tax at all. Maybe a little bit, but overall, the number's very, very low.

And now, essentially, the ratio of these types of reaction tells you how attentive people are. Essentially, if you divide, say, for example-- if you took column number three, if you divide 0.003 by 0.86, and you have to do a little bit of additional math that's uninteresting-- you essentially will get a result that, essentially, people do not pay attention to sales taxes.  $1 - \theta$  is eventually close to 0.

Put differently, the degree of inattention is almost 1. Theta is almost 1. So essentially, people do not pay attention to these types of taxes. Inattention is almost complete. So there's substantial consumer inattention to non-transparent taxes.

So what I've shown you so far is, essentially, the existence of inattention. It doesn't say why people are inattentive. It could be people just can't do the math in their head. It could be people don't want to bother with it. They're sort of saying, well, it's kind of a small change. So why do we really care?

It could be that people who buy alcohol half of the time are drunk, and they don't pay any attention. There's no modeling in any way of, why are people doing this? This is just saying this shows the existence of inattention in those kinds of consumer choices. It doesn't say why people are inattentive.

And you might say, well, 7% of a tax is kind of a small distortion. So it might be actually optimal to miss that, which is a reasonable approach. It is probably quite a bit of money and quite a bit of distortion if you think about it and you actually-- 7% is not nothing. But if you have a lot of money anyway, you might as well just not pay any attention to that stuff and if that helps you pay attention to other things.

Notice that even then, that shows the existence of inattention. It's just a bit of a question, like, how important is inattention? And this actually gets me to the next point, which is a question of, why should we care about inattention to taxes? Why is that important or interesting potentially?

Question exactly is, who's paying attention? Who's not paying attention? And then who's sort of distorting their behavior potentially? There's lots of interesting issues about heterogeneity and who's paying attention. Let me talk about this in a second.

Let me ask differently. Why is it that some taxes, the excise taxes, are very salient? And why is it that other taxes are not salient? What do you think about that? Is that an accident? Or why is it that the taxes for alcohol and tobacco, et cetera, are very salient? And why is it that other taxes for shampoo or whatever are actually not salient?

There's two simple objectives if you took a public economics class, if you took John Gruber's class, but also if you took 1401 or 1403. The reason why the government levies taxes is-- one big reason is to generate revenue, right? The government needs money. Somehow, we need to tax people.

And the issue, of course, is that if you levy taxes on some goods versus others, that can distort consumer choices. And there's essentially some theorems and some considerations about which goods should you tax. And ideally, you would tax goods that people are consuming anyway.

Essentially, what you want is people make-- so essentially, when you have taxes on some goods but not on others, if people have different price elasticities on some goods versus others and if you increase one good or all goods by a certain amount, if people have different price elasticities, then you might distort their behavior. So if you increase and you start taxing one thing and not another, people might shift to the other good. And we don't want that because, presumably, people are already optimizing to start with.

Now, if people are inelastic to certain goods, if people need certain goods anyway and they buy them anyway-- and the last thing, essentially, there's public goods. Public economics would tell you or some theories would tell you you should tax these goods more. Now, distortion, when people are-- so the distortions are lower if people don't fully react to taxes.

That could be if, for example, you started taxing bread, for example, while everybody-- or certain goods like potatoes, where everybody needs to buy these goods anyway because people need to eat-- well, then, there would be very low-- these goods would be very inelastic. And that's because they need those things. Not because they're inattentive, but just because they need these types of goods.

Now, there's obviously issues with poverty. We don't want to tax the poor too much. But in general, they want to tax things that are inelastic. And so in a way, if people are inattentive to the taxes, in a way, that's great because now they're not distorting their behavior. And we can still get money from them. So in some sense, actually, some of the inattention for that purpose might be actually good because now people are not distorting their behavior by looking at those prices. So being inattentive in some ways is good.

But exactly as Maya was saying, for some goods-- and these are often labeled as sin goods or the like, where there's either externalities or internalities. For example, if you think about tobacco or alcohol, there are externalities. Tobacco, for example, secondhand smoking is bad for others. For alcohol, there's often externalities in terms of drunk driving or violence or the like. So the government wants to increase the price of alcohol because when people are making choices of how much alcohol to buy, they might not take into account the effects of alcohol on others.

By internalities, I mean people might not necessarily take into account the effect of alcohol or tobacco on themselves. This is what we discussed previously. This could be because of present bias. This could be because of biased beliefs or the like. Essentially, people might hurt their future selves. And the government might say, well, let's increase taxes because their future sales-- people might like that there might be a way of helping people, for instance, deal with self control problems.

So here, the government explicitly wants consumers to react to the taxes. The point is not to make money from people. And in fact, sometimes you're worried about making too much money from people. The point here is really to change behavior to reduce the behaviors that are being taxed. So you want to make such taxes particularly salient.

Ready? Now, look at the taxes on alcohol versus the taxes on-- some of the excise taxes versus the sales taxes. The excise taxes are precisely salient because we want people to react to it while the sales taxes are not. Of course, there's variation across places. But it's not an accident that we see this in the world the way that is.

Now, one thing that's quite interesting now is-- and I think José or [INAUDIBLE] was saying that. It's like when consumers are heterogeneous, then there's a lot of quite interesting issues. For example, if you have the poor paying more or less attention compared to the rich or if only a subset of people pay attention to those goods, then in a way, you would find that the average elasticity is still quite low. But some people are actually distorting their behavior quite a bit, and others do not.

And that gets relatively quickly quite complicated. And you could then get into situations where you make some people a lot worse off and some people better off when making something salient versus not. Any questions on all of this?

OK. So then we get to the next question. This is getting a little bit at the question that was asked previously about, can inattention or attention have large effects? And so on the one hand, you might say, well, people's choices-- for example, consumption patterns-- are distorted due to limited attention.

But you might ask, well, if it's really true that people's consumption patterns have really distorted that much, if I'm so much worse off by not paying attention to things, well, eventually, I'm going to notice this. Somebody is going to tell me. Or in a way, I notice myself eventually that I'm doing things really worse compared to my friends. And eventually, if things are really distorted a lot, I might pay attention to it.

So the underlying idea is that, potentially, people are actually inattentive in the sense of saying, look, people's attention is limited in some ways. But it's optimal for people to pay attention to things that are really important. And there, the underlying question is, what is salient to people? And how do people decide what to focus on? And won't people pay attention to important things anyway if they have to?

And so then the question is, well, how is it possible that inattention could have large effects anyway? And so the next paper is a good example of that. And here's a very intuitive example of why that might be.

So consider the following situation. You have been getting headaches. You go to the doctor. And the doctor asks you whether it gets worse after eating certain foods. And so what are you going to say? Well, the answer will probably be like, I don't know. And why is that? Why might you just not even know the answer to that question?

An underlying theory in your mind of why you might be getting headaches. And you might just not have a theory that it's related to your food consumption. So while you have lots of data in front of you-- you eat every day, you get headaches on some days and not on others-- you could surely pay attention to it if you wanted to. But it's not even a theory in your mind, so you won't even encode the relevant information if you don't expect to have food allergies or gluten allergy or the like if you don't expect that to be a likely cause.

Then you're never going to actually pay attention, and you get infinite amounts of data on every day and headaches and food consumption. But you're never going to notice because you don't have a theory in your mind that that might be the case. And so there's a relationship between attention and memory. You only remember or pay attention to stuff that actually are theories of yours. And then you'll remember those things that you actually paid attention to in the first place.

And so selective attention, then, may have persistent effects on what we learn. Essentially, you could have unlimited amounts of information in front of you. And you will never learn because you never just had it on your radar that this information is actually relevant for the issue at hand. And Josh Schwartzstein has a very nice model that illustrates that.

Another example is from medicine. Many women died from childbed fever at hospitals in the mid-19th century. The popular theories were bad smells at the hospital, presence of male doctors wounded the modesty of the mothers. That's, of course, nonsense. The true explanation is germs. Doctors didn't wash their hands.

And again, it's something-- that's not a tricky thing to figure out if you have the idea, if you have the hypothesis or the idea that this could be important. And presumably, some doctors are washing their hands anyway and sometimes are not, and they would have a lot fewer deaths. But nobody was noticing because, again, just nobody was paying attention. So once you have the right theory in mind, then you can test it. And then, of course, you can gather the right information and pay attention to the right information and learn.

Now, the basic model from Josh Schwartzstein-- beliefs today matter for what is being attended to today, which then affect people's beliefs tomorrow. And now if you don't attend to the right things, if your model of the world is the wrong model, you might never learn to attend to important aspects of the world. And so then forecasts and beliefs may be biased and persistently biased in a systematic fashion. And you might persistently misreact and misattribute causes to unimportant variables because you have never even considered the idea of the correct model of the world.



And so now the paper I'm going to show you is an example of that. This is seaweed farming in Indonesia, which is not immediately obvious why that's related. But I'll show you in a second. So this is what seaweed farming looks like. I'll give you this relatively briefly.

But essentially, it's like, at some beaches, you can grow seaweeds in these rows. And you see these round things here, which are essentially pods. And so there are many factors that are important in seaweed farming. Essentially, it's like these pods, which are essentially these round things that you see here, they're essentially grown in different lines. And so if you're seaweed farmers, what's really important is line spacing-- or potentially important is line spacing, pod spacing, and pod size, how we got these pods.

Now, how does seaweed farming actually work? I've never done seaweed farming myself, but apparently this is how it works. So farmers use what's called the bottom method, which is they drive wooden stakes in shallow bottom near shore, and they attach lines to these stakes. And they take raw seaweed from the last harvest and cut it into pods, like these kind of roundish things.

And the pods are planted by attaching them at a given interval on the line from the sea. And at low tides, farmers tend the plots. And so the seaweed is then harvested after 30 to 40 days.

And so now many dimensions could matter-- the pod size, the distance between the lines, the distance between pods, timing, and so on. One nice thing about seaweed is there's many different pods. You can actually try and learn and estimate the importance of these factors over time if only you paid attention.

But the question is, which of these factors are important? And are people paying attention to those factors? And so what the experiment now is doing is there's lots of farmers in the experiment that are quite experienced. They have been doing this for a long time, 18 years of farming. Many of them or the vast majority are literate. So these are people who should really know what they're doing in their seaweed farming experience.

And so the enumerators went to visit these farmers and to measure and document their farming methods. And when you asked them about current pod size, you just ask them how big are your pods, people just don't know the answer to. They also don't know the answer to, what's the optimum pod size? When you just ask them how big are your pods when you plant them, they don't know. When you ask them what's the best way to do it, they don't know.

Importantly, they know exactly the answer to other questions. They know the length of the typical line. They know the distance between the lines. And they also have very clear ideas of what's optimal to do so. But they seem to essentially neglect the pod size dimension entirely when they're thinking about what to do.

And so in the experimental trial now, they did essentially different treatments. And the treatments were essentially such that farmers were provided experimental variation in the different pod sizes and essentially induced to experiment with pod sizes. And in some cases, farmers were just left on their own. They were just induced to do this experiment. And the question was like, now will farmers learn from their own in this experiment?

And then in addition, they also have an information condition where they then summarized that information to the farmers in addition. So the question is now, once farmers are induced to experiment, you surely would think that farmers would learn now that they're given all the information that they need on, like, here's the different pod sizes. Here's the different profits that you get from the different types of pods. Now they might learn on their own.

But if you don't attend to pod size in the first place, if you think it's irrelevant anyway, even if you are induced to do the experiment, you're not going to learn at all. And so then they did essentially these follow-up surveys. And then we'll just skip.

So now what they do is they also go through an information provision where they essentially tell farmers the [INAUDIBLE] they were doing with everybody. But in addition, they also provided information and did some simple calculations on which is the best combination of pod size and distance. Notice that farmers are literate. They're able to do this on their own. But the experiment was just providing this information to farmers explicitly with a recommendation about what's the pod weight and what is the optimal distance.

And so now what the experiment finds is large gains from changing the farming methodology. And the trial participation itself only has an insignificant effect. That's to say, in addition, then, summarized data for people has much larger effect. Essentially, focusing people's attention to things that they should already know changes their farming behavior a lot.

And it's essentially evidence that people were not paying attention in the first place even if you make them participate in a trial where they should have really or could have, at least, learned things on their own. There's large impacts of the trial if the data from the trials is presented to farmers. There's no impact of the trial on dimensions that farmers had already noticed previously, presumably because they're already paying attention and were optimizing.

So now what did we learn from this? And let me quickly summarize and let you go. So we have systematic learning failures even though all the information was available. So farmers simply did not pay attention because they didn't think that pod size was relevant in any way in their lives.

And that's a potential explanation why people might not pay attention even to important information because if your model of the world is just wrong, why would you even think that this is important? And the problem here is then there's never an opportunity to actually learn. You never even collect information that might reject your model because you're convinced that your model is right in the first place.

And so then lack of attention might generate arbitrarily large welfare losses. You can really screw up big time for a long time and never really change your behavior because, essentially, you never learn. You never even pay attention to any information that you might have available that says your model is wrong.