[SQUEAKING]

[RUSTLING]

[CLICKING]

GLENN ELLISON: On bounded rationality, I told you there are several different ways people do boundedly rational models last class. And the one I emphasized was this sort of psychology and economics style, where you take one of these rationality-like models you see in behavioral economics classes, and you think about what happens when consumers behave like that and firms exploit them.

But we say, to motivate the rule of thumb approach, I guess one philosophical question of like, why is it we ever even do economics with rational models to begin with? When we have-- you write down some rational model, and then you have some very complicated-- if you're going to publish a paper in a top journal, you've got very complicated mathematical argument of eight pages of calculations proving what happens under this set of assumptions.

And then you say, therefore this happens in this industry. And your question is like, why do we believe that in the first place? There's some parts of rationality that we really do believe, that if people can do something that's better for themselves, they'll do it. And they have some preferences. They have some goals they're trying to achieve.

You can't keep fooling them over and over again. They will catch on to things. But the full rationality, I think, we don't really believe it. And the same with these behavioral models. Like, OK, we believe people are short sighted, but do we really believe the beta, beta hat, delta, quasi-hyperbolic model? Why not any other functional form?

And if there are people are being fooled, is that going to keep working on them in that model, or are they going to change the way they behave when they realize that their life is going badly because of something that they do? And are they going to introspect and change it?

And so, I mean, I think one answer to why we do that is it does prevent you from cheating. It prevents you from just assuming people do this and then getting your conclusion right away. But you can still cheat with rational modeling. And you just, when you're writing down the preferences and you're writing down the cost functions and you're writing down the other things and the information, you slip in a lot of extra assumptions that, in some sense, let you get to your conclusion anyway.

And so then maybe, sociology of economics, one of the reasons we do that is because to publish a paper, you need to have eight pages of complicated-seeming complications to show that you've accomplished something hard in your paper, and then that rationality lets you do that. But then the question is, is it really a good model if it takes a lot of pages of complicated calculations to find out what happens? Because that's when we don't really-- all the minor things that depart from rationality don't matter or could have an effect.

And so what the rule-of-thumb approach tends to do is say that, rather than writing down a rational model and sticking with it, even though it's really hard to solve and we may not believe it, why don't we just write down a model that we think is more like what-- we think is kind of like what consumers do, and we're happy to buy that as an approximation that's easy to solve. And so instead of thinking of the assumptions on information and costs and whatever and epsilon ijk as being logit distributed, let's just go straight to the behavior, write down a behavior that makes the model tractable, and let us describe what happens.

So I'll start with an example. Smallwood and Conlisk is one of my favorite classic economics papers that didn't get taught when I took IO. But Smallwood and Conlisk start with-- they have a question they're starting from of, are superior products going to win out in the long run in the market? And do superior products take over markets and outcompete inferior products in a world where people don't know the quality of every product when it's first invented?

And so they use the example of, suppose you have someone buying a box of cereal. And they're standing here looking at the supermarket shelf, and they're trying to decide what to buy. Obviously, they've only tasted some small fraction of those cereals. And so they have to go there and look at the prices of some subset of them and think about how much they're going to like them and will they like it or not.

If I buy it, what's the probability I'll like it? Will I like it enough that I'll want to buy it again? And then how much extra utility will I get from the future purchase of that cereal now that I have all the extra information? And they said, well, what's even much worse is imagining that, when you're going to buy the box of cereal, you're just walking up to the shelf, and you see that person in front of you buying a box of cereal.

And then what are you supposed to do? You're supposed to say, OK, well, that person is buying this box of cereal. If they bought that box of cereal, they must like that cereal. And so if they like that cereal and then our epsilon-- these things are-- there's a xi j and an epsilon ijk. The xi j must be high for that cereal conditional on that person buying it.

But then you think, well, wait, no, this person might be buying it because he's tasted before, or that person might be buying it because, two days earlier, that person was sitting on their couch at home, and there was a TV, and they were watching a TV, and there was a commercial. And they saw the TV commercial for that cereal.

But if they did see the TV commercial, what's the signaling value of the TV commercial that they saw? Because the firm wouldn't have advertised the cereal unless it knew people were going to repeat purchase it. So maybe I have to have a prior over the cost of that cereal and the reasons why the TV commercial would have been on. But then again, maybe this person is buying it because the person who was two people before them in the aisle bought it and they saw the person before take it off that aisle, or maybe they're just buying it because it's at eye level.

But if it's at eye level, there's this signaling problem of, why is it at eye level? Is it eye level because it's really popular, or is it at eye level because the firm paid the store to put it at eye level? And would they have paid more or less to put it at eye level if it was good or bad? But especially if this person saw the person before, it's almost like I need this prior over all possible purchases by every person since the invention of cereal in the 19th century and who saw what and what they did and what that did to the popularity.

And then I've got to have that prior over that enormous network of a century's worth of cereal purchases, and then invert it and figure out, now what's my posterior on that cereal box? And Smallwood and Conlisk's view is that this is insane. This is not what people do. People walk up. They look at the cereal box. They're buying 50 different things at the supermarket. Most of the time, they just take off the shelf whatever they normally take off the shelf.

And then every once in a while, they've got a little extra time, and they put a little bit of time into thinking, which one do I buy? And so why don't we think about consumers more like that and ask about, then what would happen? Would good cereal brands take over, or do we think that Cheerios is really popular even though Cheerios isn't very good or whatever it is?

So here's what Smallwood and Conlisk did and this idea of just trying to write down-- let's write down a dynamic model that seems reasonable to us and talk about, what's the long-run implications in terms of, do superior brands take over when consumers are fairly naive in their shopping? And I think they very much do believe that consumers are naive when they shop and that this is a better model than the rational one.

So anyway, here's their model. Products differ in quality. A consumer of product k has a bad experience, which they'll call a breakdown with probability bk. So if you're talking about people buying light bulbs, that could be literally the probability that the light bulb burns out in any given-- between any given shopping trip.

If it's a cereal, it just could be that every once in a while, you're eating your cereal, and it just occurs to you in the morning, you know, this really isn't all that good. Maybe I should buy a different cereal. But products differ in the probability of giving a bad experience. And so the highest quality products are the ones with the lowest bk.

What consumers do is they always continue to buy the same product they bought previously until one of these breakdowns occur. When the breakdown occurs, they decide to buy product k with some probability mk of t raised to the sigma power, where mk of t is the market share of the product. So they just buy it with some probability that's related to the product's current market share.

OK. So one example, sigma equals 0. Sigma equals 0 would just be when a breakdown occurs, I just buy a product in proportion to its popularity. So it could be that the supermarket stock shelves so that the number of boxes of each cereal is related to how many people buy it, and therefore I just pick one off the shelf entirely at random. OK. Sorry.

Sigma 0 would be that there's one box of each cereal. I pick the box of the cereal off at random when there's one box of each. Sigma equals 1 is, I buy it in proportion to the market share. So that could be they stock more boxes of-- the more popular cereals get more shelf space, and you pick according to shelf space. Or it could be that I watch the person before me, I see them pick off a cereal, and I say, OK, that person likes it, maybe I will try it. So that would be a sigma equals 1 model.

A larger sigma could be some explicit reasoning about popularity, saying, this one seems to be the most popular, maybe I will try that, that I think popularity is probably correlated with fewer breakdowns, therefore, I will buy a more popular brand. And so, sigma going to infinity would be, you buy the most popular brand with probability 1.

And then the main result of Smallwood and Conlisk is that, with fairly naive consumers, social learning can work very well. And what it has to be is they have to use the right amount of weight on popularity. So the theorem is, if sigma equals 1, the most popular product dominates in the t goes to infinity limit. That is, the product that has the smallest bk-- so it's that mk of t goes to 1. So limit as t goes to infinity of mk of t goes to 1 for k equals the argmin of bk.

The product that has the lowest breakdown probability has its market share converge to 1 in the long run. So the best product takes over. All other products die out. And that happens when people put the sigma equals 1 weight on probability. So they're not really explicit. It's as if you just do the "watch one person in front of you" model. So it's using very minimal information about probability. That converges to 1.

When sigma is less than 1, we converge to a state with all k products active. So each product has then a market share between 0 and 1 in the long run. Inferior products stay alive in the market. Superior products have bigger market shares. They don't go to 1.

And then, anything like the sigma equals infinity, but any sigma greater than 1 where you pick the most popular product, then inferior products can come to dominate the market, where if you just get some initial popularity, you're initially popular, people buy you, your popularity just goes to 1, and superior products can't outcompete you because you're more popular.

So Smallwood and Conlisk's argument was that markets can select things, but there's this awkward part of it, that consumers have to be fairly naive to get the market to select the best products. They have to put fairly little weight on popularity. They put some weight on popularity, but it's got to be the equivalent to the sigma equals 1 weight on popularity to get social learning to work out.

OK. Anyway, so I'm just going to then just talk about a couple of other papers that have built on this and done other things. So I have a paper with Drew, 1995, where we try to take the Smallwood and Conlisk idea and make it a bit more like a classic economics utility framework. So here now, when consumer i's consume products, they receive utility, and the utility you get from consuming a product is a product specific mean. This is like the xi j or whatever.

So you get a product specific mean utility. And you get an epsilon ikt, which is your own idiosyncratic preference. And we're going to assume that the utility difference between product 1 and probability 2 is a random variable, where sometimes product 1 is better, or sometimes product 2 is better. So this is something about these products that there is a period-specific choice to them that makes one of them better than the other.

So for instance, it could be that these two products are not actually cereal boxes. It could be those two products are ways to drive to work. Do I drive to work on the Mass Pike? Do I drive to work on Storrow Drive? And every day, one of those two ways to drive to work is better, but it changes from day to day.

It could be that these are products that you're buying, where there are two restaurants, and because the food quality varies from day to day due to the quality of the ingredients or the chef that's working, on some days shop A, the food is better than shop B, and then some days shop B is better. Or it could be that these are different-- you're buying from different insurance companies. And if there's a bad event, then one of them was better to have bought from. And if there was no bad event, then the other one is better to have bought from.

So we have something that makes the quality of the two products differ. Product 1 is better with probability p. So you would like to buy the product 1 if p is bigger than a half and buy product 2 if p is less than a half.

And then we have Smallwood and Conlisk-like consumers. We don't have breakdowns. So the inertia part is, consumers buy the same product they bought previously with probability 1 minus alpha. And with probability alpha, they try to decide what-- they consider changing and think about what they should buy.

And when they consider changing, what they do is ask N of their friends about their most recent experience. I form u1t bar to be the sum-- or I form u1t bar and u2t bar. This is the average of the utilities received by all my friends who use product 1. This is the average of the utilities of all my friends who use product 2. And then I pick whichever one has the better average when I've surveyed them. Again, I asked the friends who-- I asked the friends who took Storrow Drive which is better for them, and vice versa.

And then if all friends-- if everyone I ask bought the same product, then I just buy that one. So I take an average of 0 over 0 is negative infinity, so I don't buy that one. And so the observation is, in this model, again, we can sometimes get optimal social learning from just asking friends about products. Typically, this occurs when N is fairly small.

So this is a graph showing, as a function of alpha, what's the range of Ns for which-- what's the range of Ns for which this model gives efficient social learning? And this graph seems to be showing us that if alpha is close to 1, almost everybody is asking their friends every period. Then you need to be asking three, four, five, or six friends.

Most of the time, when alpha is small, this model with the "ask one friend what they bought" works well. So in some sense, asking one friend what they bought, that's a lot like Smallwood and Conlisk because that's like choosing in proportion to market share. But then when alpha is bigger, what you want to do is gather more information. And the more information helps you find the better products.

And again, what's going on here is that there's this more-- better products do have higher mean utilities. And then there's also some effect of popularity weighting in that if something is more popular, you're more likely to hear about it and more likely to not hear about the popularity. And the combination of those two things can give us social learning but doesn't give us social learning if N is too big.

And then one other paper along these lines. Ran Spiegler has a paper, another follow-on paper, he calls "The Market for Quacks," which is the-- my paper with Drew, we were following Smallwood and Conlisk in not having any prices at all in the model, which works well for the people learning which way to drive to work or learning things like that. Doesn't work as well necessarily for buying products.

Spiegler thinks about what happens in word-of-mouth models if you let the firms set prices. So he calls it "The Market for Quacks" because he's thinking about the application-- his application he's thinking about is people selling the many medications and health things or whatever that you see advertised online that are probably no better than not buying some kind of health supplement that's supposed to improve your whatever.

And so what he imagines is that he has this model where you have this large number of products, and whenever you consume one of these products, with probability alpha, you do get a good outcome after having consumed it. But of course, there's an outside option that gives you utility that costs 0, and you also get a positive experience sometimes after not consuming any of these products.

And in particular, he's got that the outside option-- there's an outside option which also gives you a good experience with probability alpha and costs 0. So the outside option is actually just better than all of these products.

So what happens is, the utility that a consumer i gets in period t if they buy product k is 1 minus pkt with probability alpha and minus pkt with probability 1 minus alpha. And what consumers do in each period is, again, like Smallwood and Conlisk, we have this continuum of consumers. Some consumers are using every product. What consumers do in each period is, ask one user of each product about their experience and buy the product that would give the highest utility assuming the same experience if it gives positive utility.

So I ask-- there are k different products. I find one person using each product. Well, how did you like it? Did it work for you? And three people tell you that, yeah, the product worked for me. It gave me utility 1, but I paid this for it. Seven people tell you I didn't like it. It gave me utility 0. And I paid money for it, so I got negative utility.

And so then what consumers do is they pick the-- in some sense, they're going to choose among the products that gave utility 1 to their friends the one that's offered at the lowest price this period. And if every one of their friends got 0, got negative utility, then they just buy the outside option.

OK. What happens here? This model is-- if you think back to the Stahl model of price dispersion, this model is going to be kind of like that Stahl model of price dispersion, because imagine you're a firm. You're a firm, and you know that there's going to be some consumers out there who heard positive reviews of your product and negative reviews of every other product.

And so if they heard a positive news about your product and negative about everybody else's, you'd want to set p equals 1. But then there can be other consumers who heard positive things about your product and about exactly one other product. And so if they heard positive things about two products, including yours, you want to be epsilon less than the product that they heard-- the other product they heard good things about.

And so it's kind of like that kind of Stahl model, where there's multiple types of consumers, the consumers who heard good things only about yours, where they're like the non-shoppers, and the consumers who heard good things about you and several others. And there, you want to be undercutting the other prices. So what you're going to get is some kind of mixed strategy equilibrium, where you're mixing over some interval, and you're sometimes charging a high price and sometimes charging a low price to try to get the people who've heard about multiple products.

It's not as extreme as that Stahl model in that there's-- it's going to be very rare for consumers who have heard about all N products. So you don't need to be the single lowest price product in the market to get people. They've probably only heard about two or three or something like that. And so anyway, you'll have some distribution-- I imagine it probably looks like this-- of what prices look like.

An interesting observation is that in this model, price is inversely related to product quality. So if you imagine, if the products are all-- this is common product quality. If the products are all good and people actually have positive experiences with them, then you as a firm selling are going to be like, OK, if someone heard about my-- their friend had a good experience with my product, they probably also talked to friends who had good experiences with the other products. So I'm probably competing with three or four or five or six other products for their business. Therefore, I need to price aggressively to be-- beat those four or five or six products.

But if the products are terrible and there's a 1 in 100 chance that they actually do anything, then you would think, OK, well, if someone had a positive experience with my product, then when they're asking-- they found a friend who liked my product. They probably didn't find a friend who liked any of the other products. So I'm probably a monopolist, so I'm probably safe pricing close to 1.

So in equilibrium, competition between low quality products, we get high prices. Competition between high quality products, we get low prices. And the reason is that the competition between high quality products is actually more real competition, whereas the competition for the low quality products is just getting the people who have the rare mistaken impression.

So this is not a fully rational model. The consumers are not somehow inverting the equilibrium price distribution or inverting the market shares and realizing, OK, if prices are high, it must be that the quality is low. Therefore, I'm learning something about-- I had a prior about alpha, and now I'm learning something about alpha, or I'm not updating alpha from realizing that I'm just getting i.i.d.

These are all-- these seem to be i.i.d draws as I draw them every period. They're just every period taking this person liked it, therefore this is quality 1. This person didn't like it. These are all quality 0. But I think Spiegler's view is that this seems like a reasonable force and that the naive model may be a better way to think about, that this force of competition or quality drives down prices could be a real thing.

OK. Second thing I want to do is-- and I figured I would do something fun today in honor of being here before the holidays. So anyway, I'm going to talk about-- this is another rule-of-thumb paper I've written. It's not actually really an IO paper, but I feel it's useful for graduate students who have thought about these issues.

So this is a paper thinking about the journal refereeing process. So 6 or 7 years from now, many of you will be professors, and you'll be sitting in offices somewhere, and this is what you'll be complaining about, is the journal refereeing process. So in particular, what's very difficult, as people find out, is somehow it's very, very hard to publish an economics paper, and it takes years and years and years and years.

And you try to write about something topical, and then it takes five years to get your paper through the journal refereeing process. And after five years getting through the journal refereeing process, your wonderfully topical thing is now completely irrelevant because the things you're writing about are no longer topical. Actually, the experience-- I'm a department head. I'm also on Promotions and Tenure Committee for people in other departments.

And in other social science fields where people write books, this problem also just appears in spades. We're constantly reading cases of, here's this person. They've been a faculty member for five years. Their book, which was based on their graduate student thesis, is still not out yet because they've just been revising their book for the past five years. And the book is five years less timely than it was when they wrote it originally.

So anyway, how do I think about why publication works this way? Here's a theory. So in the model, papers have a two-dimensional quality q and r. q I think of as the fundamental idea in the paper or the quality of the unchangeable insight that this paper has.

The q is what you would teach when you teach that paper in your graduate course. This is the paper's idea. This is how that paper changes the way we think about something. This was their interesting idea they came up with.

And then the r is the other dimensions of quality that you can improve in a revision. So rather than having linear functional forms, you can do everything with arbitrary functional forms. You can go through the seven most recent econometric techniques published in *Econometrica* and re-estimate it using these seven different estimators and talk about how all those estimators do things.

Obviously, some parts of the things we improve in revision are making us more confident in the results. They are valuable. Many of them are just doing everything-- addressing every potential criticism that anyone could have come up with and just continuing to just make your paper harder and harder to read. So we value both aspects of quality, the idea part and the craftsmanship part.

And referees know that they're supposed-- that there's a social norm in the profession that we put alpha weight on the q and we put 1 minus alpha weight on the r. And they know that they're supposed to recommend acceptance if this weighted average of q and r is at least z. But referees being people who are, in general, 32 years old or whatever, not in the profession for all that long, don't necessarily know what alpha and z are.

And they're trying to learn. They're just trying to be good referees and evaluate papers appropriately when asked to evaluate a paper. And so at time t, the profession ends up using these weights, alpha t and zt. And all the referees in the profession are just trying to update all the time and learn, what is alpha t, what's the weight on these two, and what's the quality threshold where I'm supposed to say yes versus no?

I assume that authors rationally divide their time between producing the two types of quality. They start by picking some amount of time tq. So 1 is the total amount of time you have available. You start by spending some time tq in 0,1 to get a random draw on an idea.

And at the end of that time tq that you spent developing random draws on the ideas, then you get an idea of quality q. And there's some CDF F of q given tq that's increasing. So for instance, it could be that just q is uniform on 0 to tq, is one. But you could also make it normal mean tq or whatever.

And then after you get your idea, you send your paper in to the journal, and they evaluate your paper. And then they tell you how good-- they give you some answer, r of q, which is like, if you can make your paper-- get your craftsmanship up to quality r of q, we would take your paper. Otherwise, no, it's not good enough.

So here's your paper. In general, actually, you send a paper to a lot of top economics journals today, you're going to get five referee reports that are going to be a total of like 18 pages long and have 67 comments of things you could improve in your paper. And you then decide-- they're telling you how much you have to do to get the paper into the journal.

And then authors decide how much time of their remaining time between 0 and 1 minus tq to spend improving the paper. And time spent improving the craftsmanship increases r to h of tr plus eta, where eta is again some random variable. So sometimes you're able to do the things well, sometimes you're not able to do the things well. And then I make eta, again, a uniform distribution.

And then what editors do is editors accept the highest quality papers that are resubmitted. The editors all also know alpha t. They all know the current market weights on these things. The editors, though, don't use zt. The editors, they know how many papers they're supposed to accept, so they just accept the best.

Even if the referee has told people to do more than was feasible, the editors have to go against the referee and say, I have to accept something. And so the editors just accept some fraction tau of papers that are highest on this metric alpha q plus 1 minus alpha r. So anyway, that's my model of academics writing and revising papers.

First observation. This model has a continuum of consistent social norms. If you think about it, if I pick any alpha, if I pick alpha equals 1 so only q counts, if only q counts, what people are going to do is just spend 100% of their time producing ideas, have zero craftsmanship, and there's going to be some level z that clears the market. So if you spend 100% of your time on q quality, a fraction tau of people are going to get qualities above this level, and 1 minus tau are going to get qualities below that level. And that clears the market.

If you choose alpha equals 0, then people are going to spend zero time thinking up ideas, and they're just going to have very, very complicated estimations or proofs of their non-ideas. They're going to have incredibly high levels of r quality, but they're going to have no ideas in the paper. And again, the people who get the best draws-- everyone spends time 1. The people who get the best draws here, the tau fraction, get the best draws, get their papers published. The other people get the bad draws on r and don't get their papers published.

And then, for every weight in between, people are going to spend some time in q and some time in r, and there's some quality threshold that clears the market. So in some sense, there's nothing in this model that pins down whether we care about main ideas or whether we care about craftsmanship. Profession could have any set of norms. They're all stable.

And this would be an example of what this equilibrium would look like. This is a weight where people are putting weight on both. So people choose some time q. It must be a pretty high. People put some weight on q. This is equal weights, looks like. People put some weight on q that's between 0 and 1. They come up with ideas.

If at the end of your time q you end up with a bad idea, you send it to the journal, and the journal is like, in order to get your paper accepted, you're going to have to reach this super high threshold on craftsmanship because you don't have much of an idea, you look at it and you're like, given that I made the distribution that r is h of tr plus eta, where eta is uniform 0 sigma, there's an upper bound to how good my craftsmanship can be even if I spend all the time on it.

So you've got a bad idea. They tell you what to do. You're like, there's no way I can do that. You just give up. You don't revise your paper, and you go home. If your paper is better, you revise it to death. You set tr to be 1 minus tq. You do all the revisions you can. You get this draw on r quality. And if you get a good draw, you get accepted. You get a bad draw, you get rejected.

And then, if your quality is super high, then they recognize your quality is super high and they tell you, you only need to do this-- reach this level of r quality to get accepted. So what you do is there, you're able to get it accepted with probability 1 here. And up here, you can get it accepted with probability 1 and even take a vacation, some vacation time.

So here, you take some time off. But any of these papers that are super high quality, they get accepted with probability 1, and you just do the-- you set-- you choose h of tr star is just equal to the thing that you're being asked to do so that even if you get the worst data draw possible, the paper gets accepted with probability 1. So you ensure that your paper gets in.

In my model, people have these lexicographic preferences over paper acceptance and leisure. So anything that adds any probability to your paper getting accepted, you do it, and you only take leisure when your probability of getting in is 0 or when it's 0 marginal impact. Anyway. So this is what we think would happen.

And so that's what an equilibrium would look like. And then the question about, where do these dynamics come from, because it's gotten harder and harder to publish papers over the years. Why do I think that happens? What I then add is an overconfidence bias. So everybody thinks their papers are better than they actually are.

So either you think your q is better than it really is, your idea is more interesting, or you think your craftsmanship is better than your craftsmanship really is because you're using the best techniques you know about and you think, therefore, I'm state of the art, when you're not. But everybody just thinks their paper is epsilon better than other people think their paper is. OK?

And so what's going to happen when people are learning is, this norm with an overconfidence bias is going to confuse people because they're going to be-- two things are going to be happening to them. One is, they're going to be submitting papers, and they're going to be told that you've got to reach this really, really high-- well, what seems to you to be a really, really high line because you think your paper is better than it is, and you're like, wow, it's really, really hard to get a paper in. For my paper to get accepted, I have to do this laundry list of 67 things better than I'm doing today.

And yet you also referee papers and you say, this is not as good as mine. You should reject it. And then you see the journal taking the paper. And so you have this sort of cognitive dissonance. It's like, why is it that I can't get my papers in and these mediocre things either that I refereed or I see in the journal are getting published? And so you have to try to reconcile those two things in your mind and think, maybe I misunderstood the alpha z process.

Again, here, I could do something rational where-- something fully rational where people have this prior overall alphas and z's that the world might have and a prior over the whole process that is being used to generate acceptance and rejections. And think about the rational Bayesian updating. When you're getting these two conflicting pieces of evidence, how do you best-- in some sense, it's a Bayesian updating where the true state of the world is outside the support of your prior because you don't think that you're biased.

But anyway, what I did instead was think about just a different objective function. That is, I'm basically getting two types of evidence. One is, I'm being asked, these are the r of q's on my paper. And I'm trying to figure out, what is the correct-- I'm trying to figure out, what is the correct line? And this is what I think r of q is on the observations on my paper.

And I'm going to say that people are in some-- one thing they're trying to do is fit a line to the data that they're getting on their own papers with some least squares procedure. But then the other thing they're doing is they're seeing other papers are accepted or rejected. And every time a paper-- if this paper is accepted but this paper is also accepted, and this one is rejected and this one is rejected and this one is rejected, and this one's accepted, they're trying to, in some sense, also fit a line to this data that separates the accepted papers from the rejected papers.

And what I do, entirely to make the model tractable, is say that you're using a least squares loss function for the things you're told about your own papers, and you're using a sum of absolute deviations loss function for the acceptance and rejections and trying to fit the two data together with the same line. And what's going on is that they're-- actually, basically, there are-- what's happening is in reality is there are two lines.

This is the r of q line, what you perceive you're being asked to do, and this is the acceptance-rejection line, which is lower. And you think you're being asked to do more because you think your quality is better than it is. So the data is really coming in on these two lines, and you're trying to fit a single line to this data, not recognizing that the observations from your papers are different from the observations on the other papers.

And the observation of the paper is if people have views on this, what happens? What happens is, people perceive this line to be the line. And everyone thinks it's harder to publish than it is. They always tell the authors when they're refereeing, you need to get really, really good to get in, and you need to reach this high r of q line. But then the editors can't fill the journal with papers who reach the high r of q line.

So there are these people down here who had the medium quality ideas that revised them to death, who fell just short of what the referees told them to do. And then the editors take the paper anyway because they need to fill the journal. And so they take the papers that were slightly lower quality on the r dimension than what they were told to do. So that's what the data actually looks like.

And then when people are trying to fit these two data sources together, the r of q data and the acceptance-rejection data, the best way to do that is by tilting the line down because the surprise acceptances are papers that are low q, high r, and get in. And so the best way that you can explain that on these really great papers, I'm being asked to do a lot, and yet down here I see these low q, high r papers getting in, is to say, oh, I think it must be that the best way to fit this data is to tilt the line down and think that r quality is more important than I thought. And these high r papers are getting in because r quality matters more.

And so the argument is that what's going to happen is, people are just going to keep being surprised that papers get accepted that they refereed. The surprise acceptances are overwhelmingly low q, high r papers. And so the best way to explain that those surprise acceptances and yet the high demands being made for the high q papers are tilting the line and thinking r quality is a little bit more important than I am.

And so we end up with this situation where just perpetually, everyone always is holding people to standards that are slightly too high, and then slightly twisting their views to think about the importance of craftsmanship versus the importance of ideas. And this graph shows you-- if you think of this as the set of social norms that would all be consistent with alpha on this axis and z on this axis, we're always staying slightly above that set of social norms. And then this dynamics of this model would just slowly drift and drift and drift and drift and drift.

And the argument is that-- I'm going to argue that this was a 30-year disequilibrium process. For 30 years, we've been out of equilibrium. Referees are always slightly too high. I'm just trying to reconcile the difference in their mind between the slightly too high standards they think are being applied with the acceptances they see. They just keep thinking that it must be the craftsmanship that got that one in. And then they keep just changing their beliefs about the importance of craftsmanship.

Any questions there? So again, the rule-of-thumb approach here is just I'm trying to say what the referees do. And I've actually-- some chance I will rewrite this paper doing the fully Bayesian version, which I think there's been big advances in Bayesian learning with misspecified priors in the last 20 years, where I think maybe now I could solve this Bayesian model with misspecified priors and show that kind of what I say happens here, happens in a fully Bayesian model with misspecified priors.

But I think that the idea of the rule-of-thumb approach is just to say, think about what you think people would do. Think about what you think they would do that's reasonable. I don't think it's reasonable-- I think it seems like referees are fairly naive, and this is how they're trying to reconcile the two things. This could happen. And then write down a model that you can actually solve.

And like, why do I have a quadratic loss function in one part and a least absolute deviations loss function in one part? And it's exactly because that was the model I could solve. We might as well solve-- write down models we can solve and say what we can about them.

The, let me say, empirics on this, I have another kind of depressing paper in the *JPE,* looking at how long it takes a paper to get accepted from the time it's first submitted historically. And in the 1970s, papers were typically getting accepted at all the top journals 6 to 8 months after they were first accepted. And it's kind of remarkable when you realize, at the time, the Xerox machine had not been invented.

And so how did refereeing a paper work? You would type up two or three copies of your paper. You would send-- or actually, in the 1960s, you would send in a single draft of the manuscript that your secretary typed. But you would send the paper in-- so '60s, certainly, your secretary typed one copy of the paper, you sent it to the journal, they sent it to the referee.

The referee read the paper, made comments, sent it back to the journal office because that was the only copy. The journal office then sent the paper to a second referee. That second referee made comments, sent it back to the journal office. The journal office made comments and mailed you a decision. Somehow, at the time, papers were getting published 5, 6 months after they were submitted.

Starting in the '70s, when the revise and resubmit was invented, things just got longer and longer and longer and longer. And by the year 2000, some journals were averaging 2 and 1/2 years from the time a paper was submitted to get it published, some were averaging two years, and so on. And this is the time that's conditional on a good outcome.

The way the process tends to work, though, is the top journals have a 96% rejection rate. So you send it to journal one, they screw around with it for a year and a half, and then they reject it. And then the second one, and then they mess around with it for six months, and they reject it. And the third one, it's a year, and then they reject it. And then finally, you get someone to spend three years revising it before they then-- two or three years, and then they publish it.

So actually, I have this very depressing statistic that if you look at the absolute stars on the IO market who do empirical IO and how long it takes them to get their job market papers published, the median time is now 7 years, which, you know, it's a long time, because then you spent also three years writing that paper before you first graduated from school. So.

Anyway, I hope future generations will recognize what they're doing and solve this problem. And I've made some attempts with limited success. Any questions? Yeah.

AUDIENCE: How important is it in the model that the authors perfectly observe the quality of their ideas?

GLENN ELLISON: So I don't think it's at all important. I think what's going on is I'm having the authors do this sort of least squares fit of this. So they're being asked to do-- every time you submit a paper with quality q, you're being asked to r of q. And you just fit a least squares line to that.

So if there was sort of-- if your paper qualities were like this, fitting a line to it, I think, would do roughly the same thing. What's important is the argument that the very high quality papers, you're asked to do a lot, and they do all get accepted because you do what you're being asked. And the low quality ones are the ones that get in not doing what they were asked.

So the high quality ones don't get in not doing what they were asked because they always do what they were asked enough. That's, I think, the main thing. But the randomness I don't think would affect-- on either one shouldn't affect it.

Let me say, advertising is not something I can do any-- do justice in this paper. There is a field called marketing. I don't know if any of you are from the Sloan Marketing Department, but most business schools have marketing departments. Marketing departments have lots and lots of professors. There are like a thousand marketing professors in the United States.

There are many, many marketing journals, just like there are economics journals. Those marketing journals are full of papers. And obviously, people in marketing do advertising. They teach entire-- all of their courses are entirely marketing, not one week of my class or something like that. So there's an enormous literature out there.

Marketing is a field where-- I put here influences on marketing. So if you look at what's being cited by marketing papers, the largest-- well, obviously, mostly they cite other marketing papers. So this is, what non-business things are they citing? They do cite a lot of psychology papers, and there's been of a rise of psychology in marketing over the last 40-year time span, where people in marketing are paying more and more attention to what psychologists are saying and citing what psychologists are saying.

There's also been an increasing trend in citing what economists are doing. And so now it's 6% of the citations in marketing papers are to economics papers, rather than to marketing papers. So there's a substantial influence of economics on marketing. And if you look at what within economics are marketing people citing, IO, which is the yellow line here, is a big part of what's going on in marketing.

And so there's just-- there are marketing papers that look like psychology papers or look very different from economics. But there are a lot of marketing papers that look like IO theory papers and that look like empirical IO papers. So they're just-- marketing people write all about the way that firms can price or make their products appealing to customers. And they write a lot about empirical demand estimation and how advertising campaigns affect demand and so on. And both of those parts look a lot like IO.

So what I'm going to do today is, today, I'm just going to do a few classic models of advertising. And then it's, again, a week from Monday, I'm going to do some empirics on advertising. So models of advertising. Generally, I think you can see the way people in the economics literature break down three different-- obviously, to explain advertising, you've got to say, why is it that advertising changes what people purchase?

If you start with our basic model of consumers have preferences ui jt. If they buy product j at time t and it's xjt beta minus alpha pjt and they buy the best thing, then what does an advertisement do if it doesn't change your-- if that's your utility function, you're always maximizing your utility. Advertising shouldn't affect anything that goes on in the world.

And yet we see firms spend lots and lots of money on advertising. It must be advertising changes behavior or changes something that increases profits. What is it? One thing, one class of models is advertising can have informational content. It can tell you that a product exists, what the product's price is, what the product's attributes are. But in some sense, it's like a search model. You don't know everything, but then the advertising tells you things that you would otherwise have to invest search effort in figuring out, and that increases the number of people who buy the product.

And classic examples like this. Informative advertising. You just get an ad in the mail or in your newspaper for Target. And the advertisement for Target tells you you can get these boots for $10 and those jackets for $15 or $20 and these clothes for $4 or $5. They show you the pictures of things to tell you what Target is selling and what it costs. And this would be informative advertising.

A second thing that advertising can do is not necessarily tell you anything about the product itself, but just signal to you something about the quality. So for instance, let me look at these two ads. You guys are probably all too young to remember pets.com. Pets.com in the first internet boom-- just like when there was amazon.com, pets.com started up around the same time.

Pets.com, just a miraculous rise and fall of a company in the space of a year or two, where pets.com just burned through hundreds of millions of dollars. They had this sock puppet. And they put the sock puppet on the Super Bowl and spent millions of dollars just having the sock puppet just appear on the Super Bowl. The sock puppet did not tell you anything about-- obviously, what is pets.com? They sell stuff for pets.

You know what's in a Petco store. You probably know what's on pets.com. But the sock puppet, in some sense, maybe was sort of saying that, wow, they just spent $2.4 million to have the sock puppet just interviewing somebody during the Super Bowl. It must be they think they have a good product. And if I actually go to pets.com, I'll buy stuff because otherwise, how could they have burned through $2 million in 30 seconds like that? And so you get this sort of signaling idea.

I would say the same thing about-- you look at the Patagonia ad here. Again, I don't know. I mean, maybe it's all Photoshopped, but it must have been expensive to send some guy with a camera up to the top of some mountain wearing those Patagonia stuff and taking the pictures. And it must be that their clothing-- either this person is really not uncomfortable up there on that mountain wearing the Patagonia stuff and so the clothing is good, or they're just spending a lot of money on the ads because they know that if I bought the clothing, I would like it or something like that.

And then I guess the third thing we think about advertising is advertising can be changing your preferences or directly affecting purchase decisions of non-rational things. So if I look at this commercial, "Let's Grab a Beer." And those people are having just a great time having those beers outdoors at some thing, and they're all relatively attractive or whatever, it must be I think that somehow, if only I was having that beer, my life would be like that instead of what my life is like today.

And I would like to be those people. And somehow, if I just had that beer, this is me, or something. It changes your preferences from being a person who wants to work on a problem set to being a person who wants to be out on the street drinking beer. And your tastes change, and therefore your purchases change. OK.

So how do-- let me see-- how do I model these, and what do we get out of the model? So first classic model of informative advertising is Butters, 1977. Butters was thinking about N firms selling identical goods. There's this unit mass of identical consumers, and they're all going to get utility v minus p if they buy this good at price p. All the firms have some cost c of production.

And this is a problem where the consumers don't know either that the product exists or where they would buy it. So what the firms can do is they can spend amount of money-- so if they spend A of x on advertising, their ads will reach a fraction x of consumers. So it's, I guess, good to think about this as putting ads out there on the radio or print airwaves. And just every ad you put out there, you're going to hit more consumers.

There's decreasing returns because the first ads I send out all hit somebody. But then the later ads that I send out are mostly hitting people who already heard about my product. And so by having them hear about my product, I'm just wasting duplicated messages on them.

And so I'm sending out these messages. And the more people I want to reach, the higher is the A of x, the money I'm going to have to spend. And it could be that even getting to x equals 1. This thing could almost asymptote to infinity. If I'm trying to reach every person in the world, it's hard.

And so they consider a game where firms simultaneously choose two things, the amount of advertising x they're going to do, and the price is p. And then consumers who buy at least one ad buy from the cheapest firm that they heard about. And again, the goods are identical. So it's just, if v minus pj is positive, then I'll buy it from the cheapest firm I found out there.

And if I don't hear about a firm, I can't buy from them. This could be a good model for things advertised through infomercials on TV also, or something like that. It's like, if you see the infomercial, you can then buy the product and write down the phone number.

OK. Observations. Again, this was one of the early papers that had an equilibrium with price dispersion. This is like the Spiegler paper I talked about earlier. This is like Stahl from a few weeks ago. Again, some consumers are going to only get one ad. If the consumer only got your ad and nobody else's, then you want to set p equals v, but if they've got multiple ads, we're again in this situation where, in equilibrium, we're going to get some distribution on prices that has an upper bound of v and then goes down from there.

And it's just determined by, I want to undercut, but if they only see one ad, I'd like to be the highest-- I'd like to set at price equals v. But then they might see two or three or four or five ads, and so I undercut. We get this mixed strategy pricing equilibrium.

Advertising levels. A prime of 0 is small and A prime of 1 is really big. So if I make it even more stark, if this thing starts out like in advertising, the first people is really cheap, and then reaching the last people is really expensive, there's going to be some internal level x star of advertising that I pick. And I'm going to pick it x star so that the marginal cost of reaching one more consumer equals the marginal profit I get off the last consumer I reach. So the first order condition is going to be the profit on the marginal consumer reached equals the cost of the marginal consumer reached. OK?

Interesting observation that Butters has is that in this model, advertising levels are socially efficient. You might have thought this is like entry where there's business stealing and because of business stealing, you could have too much entry. Or it could be because you don't internalize consumer surplus gains, there could be too little entry or entry being advertising. But in this model, with the technologies he's got, advertising is socially efficient.

And we think of that like-- it's an argument that here's a special case where there's not a distortion, and the informative advertising in the world should exactly match what a social planner would want it to be. And why is that? So first order condition for x star gives the marginal cost of reaching the consumer equals the marginal benefit to the firm of reaching the consumer. Marginal cost of reaching the consumer is A prime of x star.

What's the marginal benefit? The marginal benefit, I'm indifferent. I'm mixing over these prices. So I'm both choosing a pure strategy x star, and I'm mixing over prices. So it must be I'm indifferent. My profits are the same regardless of which price I choose from my price distribution. And so my profit on the marginal consumer is the same, regardless of which profit price I choose from the price distribution.

So I'm going to evaluate it, looking at the case where I set a price at the upper bound of the price distribution v. So what's the marginal value of reaching one more consumer when I set a price of v? Well, I get v minus c is the profit from selling to that consumer. And then I get that multiplied by the probability that the consumer saw no other ad, because when I set price v, I sell to them only if they got no other ad, because if they got any other ad, they'll find a lower price, and they won't buy from me.

But this is the social benefit of an ad, of reaching a consumer who's never been reached before. Because if a consumer has never been reached before, the social surplus from the firms interacting with that consumer goes from 0 to v minus c. So what we're getting is that the private return to reach one more consumer is the social return to reaching one more consumer.

And it's not obvious if you think about any other price. When you set any other price, you are doing some business stealing, and you are stealing consumers away from other firms, and you are increasing the consumer surplus of the consumers. But somehow, this model is such that those two things do exactly offset. And so there's no social inefficiency here.

So it's at least one case where we don't worry about advertising being distortionary. And obviously, this is an extreme. Butters has come up with a special case where this is true. If you had the firm selling imperfect substitutes rather than perfect substitutes, we wouldn't get this-- we could get this kind of mixed strategy equilibrium.

We get the kind of pure strategy equilibrium pricing like we got with search, with differentiated goods. You'd have the business stealing effects, you'd have the consumer surplus effects. But I think, in the entry literature, we have this intuition that one of them was often going to be bigger than the others. Here, I think we don't have that intuition. OK.

Second class of models I want to talk about then is, how do we think about signaling models of advertising, and when would they work? So Milgrom and Roberts is your classic paper doing this in the 1980s. Milgrom and Roberts' model, all firms that are entering the market, they try to come up with some product. They invent some quality q. We're just going to make it a two-point support. Your quality is either low or high.

Obviously, all the firms are trying to invent quality q, but they sometimes come up with quality H or sometimes come up with quality L. There's some common prior that the consumers and the firm have over the probability that the quality is high of a new product. The firm observes its own quality and then chooses a price p for the product and a level of advertising A.

Important in this model is consumers observe A. The consumers observe that they just put that sock puppet on the Super Bowl. That was $2.4 million. In some sense, this is what we tend to refer to as a money burning theory of advertising. It almost doesn't matter what they did with the money.

It's like, yeah, those guys just put $2.4 million in dollar bills in a big pile and lit it on fire. That proves to me that they must have thought it was worth it for them to burn $2.4 million. So they must be the type of firm that wants to make big bonfires of money.

So anyway, consumers observe A. They Bayesian update on what q is, knowing what different firms would do in terms of burning money or not. And so then they form a posterior on what q is, and they buy the product. And then the thought here is that this is like a dynamic model, although they don't really model the dynamics fully. But then after you buy the product, you learn something more about the quality, and then you may buy the product again, or you may buy it a third time, you may continue to buy it for many years, you may tell your friends about it, whatever.

So you observe the A. You decide whether to buy. If you buy, you learn something about the quality. And then there's some repeat purchasing going on later on. OK. They say, whatever's going on in all those later stages, we're just going to model this in a reduced form way. We're going to say that the profits the firm gets are pi of p given q and q hat that depend both on the true quality q and on the quality q hat that consumers expect to receive when making purchase decisions.

And generally, profit is going to be increasing in both of these arguments. If consumers think your product is good, even though it's not good, they're still going to buy it once, and you're going to make money from them thinking it's good. But then conditional on how good they think it is, what makes them buy it or not, if it's actually good, then when they buy it, they're going to enjoy it. They're going to buy it again. And so your profit's increasing in both the true quality q and in q hat, which is their posterior on q after you've had your advertising campaign and set your price.

Observations. This model can have a separating equilibrium, where only firms with high quality advertise. So we get this equilibrium where you have these two things, a low price and zero advertising and a high price and some level of advertising. And all the low quality firms do this. All the high quality firms do this.

When is it that this can happen? I think two things you're going to want to have. First is that repeat purchases have to be important. If there were no repeat purchases in the model, this wouldn't happen because if there were no repeat purchases, low and high quality firms would earn exactly the same profits. So if the high quality firms found it better to set a high price and burn money, the low quality firms would also set a high price, profit, and burn money.

So what you have to have happen is there's got to be repeat purchases so that the high quality firms are willing to burn this money because they know the consumers will keep coming back. The low quality firms are only going to sell to consumers once. So you have to burn an amount of money such that the-- it's got to be that pH minus pL or pH minus c D of pH or D1 of pH minus pH minus c-- pL minus c.

So imagine that this is the extra profits I get in the first period from burning the money. It's got to be that this is less than A because if this is less-- if this was bigger than A, then even the low quality firms would just go ahead and burn the money in order to get the extra demand in the first period. So it's got to be that this is less than A, but this plus the long-run benefit that you get from the repeat purchases at t equals 2 or t equals 3 or whatever does make it worth advertising.

And so we need a repeat purchasing effect. And then we also need some kind of small-- we need word of mouth learning effects to be small because if word of mouth learning effects were going to be small, you could just come in, set a price of pH, do zero advertising, have only a few people buy your product. Then those few people tell everybody and they tell everybody and they tell everybody, and the thing takes off, and the long-run benefits are there, even without having spent the money on the advertising in the first place.

So what we need is something about repeat purchases help-- getting people to buy once, the repeat purchases make it worth spending the advertising. And you can't do it without spending the advertising because you can't do it by just a small release and then let people tell their friends instead.

Technologically, this advertising is money burning. It's not changing the utility functions at all, but it can increase-- it can be socially valuable if it increases the match quality. So if you imagine a model where you have-- there are some people who only really have value for a high quality-- some people who are serious hikers. And so if it's not a really good Patagonia coat that's actually going to keep you warm on a mountain, you don't want it. And then there's some low quality consumers who are just wearing it around town. And if it doesn't actually prevent the wind from getting you, it doesn't really matter because you're just wearing it around the city.

If the advertising helps us with the match quality and lets the people who want high quality buy high quality, people who want low quality, buy low quality, then this sort of money burning can raise utility. And we might not want to ban it, even though it's not technologically valuable. But it can be the only way to signal information here.

And in some models like this, you can actually even do signaling without any advertising at all. So it's possible that there can be-- don't go in advertising. There can be equilibria like this, where some firms set a really high price, and then people know that they're only going to sell to a small number of people. So it must be they're expecting to sell to that small number of people over and over and over again.

Whereas you can set a low price and sell to a lot of people who just buy once or something like that. So you can get you can get price as a signal. You can get price as a signal even without advertising. But advertising makes it easier to get signaling models to work out. OK.

OK. Questions? So then I thought the final one, I will go to things based on my beer ad. And here, I'm covering an unpublished paper that's unpublished for 15 years or whatever. I'm sure Jesse has long since given up on this paper, but I always really liked it, so I'm going to keep covering it anyway. So this is a behavioral economics paper about advertising changing preferences.

And it starts out with noting, there's a lot of very, very good evidence about fallibility of memory and manipulability of memory that makes you think that these things can work. So, for instance, one of these very classic experiments, the famous person experiment is, I give you a list of people. And I write down, I give you a list of kind of famous people, and I ask you, which of these people are famous?

And I put in a bunch of names on there like, Ban Ki-moon and Sebastian Weisdorf and whatever. And these people are people who you've kind of heard of. And you're like, OK, yeah, Ban Ki-moon, that's some guy who did something UN official or something. I don't remember quite who he is.

Sebastian Weisdorf, that just doesn't ring a bell. I'm going to put not famous. And you put names of actresses and whatever that, again, you know, I have no idea what that person's ever been in, but I know that's the name of some actress and whatever. And people are pretty good at this task of saying, is someone famous?

And one of the interesting things is you give people that task, you give them, here's 20 people, is this person famous or not? And some of them are just made up names. And people are very, very good at that task. They can write down famous, famous, not famous, famous, not famous, even though they have no idea who the people actually are.

And then you bring the same students back in the list-- into the lab. One week later, you hand them exactly the same list that they said yes or no on. And they're much worse at the task. And it's because, OK, Sebastian Weisdorf, I've kind of heard that. But is that because Sebastian Weisdorf is some guy I've heard of, or is it because he was a guy who was on the list that I did last week? And people no longer can remember the things as well as they did the week before or whatever.

He also gives these examples of these marketing experiments with orange juice, where you have people taste orange juice. And you prime them before with the correct answers, like here's an orange juice that's got a really tangy taste or whatever. And then here's an orange juice that's got a really sweet taste or whatever. Which one do you like?

And then the tangy taste one is just like, you put vinegar in the orange juice, and you just made it worse. But if you've described it like, oh, here's a really sour orange juice taste, or here's a really tangy one taste, somehow, if you've told them it's a really tangy one, they're like, oh, yeah, I like that one, that's tangy, or something like that.

And so his thought on the beer example is, people have had beer many times in their lives, and they've had parties with their friends many times. And the thing is, you like to go to events where you're having a good time. And you don't remember whether it was me who had a good time drinking the Budweiser or if it was those people I saw on TV having a good time drinking the Budweiser.

And so then when I'm there in the store, I'm trying to remember, oh, should I buy the Budweiser or not? And I'm just trying to remember some previous time when I had a party and I served Budweiser, did people have a lot of fun? And I can't keep track of whether it was my party with the Budweiser where people had the fun or it was that one on TV I saw with people having Budweiser, where everyone had a great time.

And in some sense, they've changed my memory and put all these memories in there that I can't distinguish from my own memories. And therefore I can't-- and so you're changing my preferences by changing what I remember about my prior experiences with the product, confusing my experiences with the people on TV's experiences with the product.

So the way he does this is he has this continuum of consumers who have types theta i. And the type theta consumers get benefit v with probability theta when they consume the good. So could this just be-- depending on your friend group, you could be in a friend group where serving Budweiser at your party, 100% chance the party goes well. You could be in a friend group where serving Budweiser at your party, just 0% chance the party goes well because your friends all hate Budweiser.

Anyway. So each consumer is different in their experiences when they've consumed the good. Consumers have consumed the good on N prior occasions and gotten these draws. But they don't know theta. I can't go, every time in my life that I've bought Budweiser for a party, how did that party go? Let me count them all up and add them all up.

What I do is I just remember often just one. I just call to mind. Let me try to remember some time when I served Budweiser. I remember that experience. I remember how it went. And then I Bayesian update. So I have some-- there's some common prior over the thetas. And then I Bayesian update to form, what's my expected view of my preference for the product, conditional on the events that I'm remembering?

And then what Jesse does, he models advertising as potentially altering memory in one of two ways. One is an ad could just take a memory out of your head and stick a fake memory in instead. So it's just taking one of your N observations out, putting another observation in so that when you do the Bayesian updating, you might recall that memory.

The other thing is a notion of priming memories. That is, we all have all these memories in our heads. If I get you to make one memory, call that memory to the front of your mind, then the next time you're trying to call a memory to the front of your mind, that one will come forward again. It's been more recent and more primed to come forward.

And so here, this is, you show people a positive experience while they're watching the ad. They remember the most similar experience that they've had. And then because that experience has been brought through to the front of their mind, it's more likely to be drawn the next time it comes out. So this is like you're manipulating memories by making people's memories more-- changing the probability of recall of different memories that they have.

And then the paper itself, it mostly focuses on this. It gets into this highly rational case where consumers are aware that their memories are being manipulated, and they are aware of the equilibrium level of memory manipulation that's going on by the firms. And then they try to actually-- if I were him, I might have started with more of the non-fully rational thing, where people are just being manipulated and don't realize it. But he's mostly doing the, they are trying to offset the manipulation of their memories.

And so, in some sense, they downweight it. They're like, I know this firm in equilibrium is advertising. Therefore, I know that this memory might be mine or might be someone else's. Therefore, when I'm doing the Bayesian updating, I'm doing the Bayesian updating of expectation of theta given ri and given A star, where I know what the equilibrium advertising level is in the world. And so I know, and I'm trying to undo this.

But nonetheless, the point is, even in this model, where memories are being artificially manipulated, firms might still want to manipulate the memories. And this is kind of like for information design reasons. So you're not changing the mean-- you're not changing the mean level of people's posterior on theta, but you are changing the distributions on the theta.

And so if you can change the distribution of thetas in a way that lets you extract more of the social surplus, it may make sense to advertise even if the advertising is being undone. So if the advertising is not being undone, it's clear you want to advertise. But even when it's being undone, you might want to do it for information design reasons, where if I can get everyone to have the same posterior, then I can charge that posterior and extract all the social surplus. Or if there's more heterogeneity, I only extract a fraction of the social surplus.

Anyway. So it's just got some interesting observations. And it's also interesting about-- you can think about when advertising is going to be more and less likely because it's going to depend on product quality because if you actually have a high-quality product, you don't need-- you don't get as much marginal benefit from replacing memories in people's minds with fake memories because the true memories are good. And so you can get this sort of advertising as a signal of bad quality and various interesting things.

OK. Anyway, I am out of time. So anyway, Monday, I'm going to-- again, I'm going to do auctions. It's going to be mostly auction theory. I am going to do one empirical paper that I really like. This is the Hendricks and Porter paper about offshore oil drilling auctions.

And then, yeah, that's the plan for next week. I'll do auctions. Tobias comes in Wednesday and does structural empirical auctions. Then we go back to advertising empirics the week after.