[SQUEAKING]

[RUSTLING]

[CLICKING]

GLENN ELLISON: OK, anyway, so online markets. So what I've got several different papers I figured I'd talk about today. All of them are talking about online retail in one form or another and just different issues in online retail trying to understand the industry. So the first paper I was going to talk about, Einav, Kuchler, Levin, Sundaresan, *Learning from Seller Experiments in Online Markets.* This is a paper studying eBay.

And while eBay has been in a long run decline, it's still an important way in which things get sold. And I also just wanted to cover it because I felt the methodology of the paper is interesting as a way to attack problems and these ideas that what they're talking about is an alternative to doing field experiments or setting up agreements to run experiments with online service providers is what they call opportunistic experiments.

There are many, many businesses in this world. Those businesses are trying to operate optimally. They run lots of experiments themselves. And so their idea is to take advantage of all these experiments that sellers are running on eBay and analyze those.

And obviously, you know, you get less controlled experiments than if you were setting up the experiment yourself. Each experiment is not ideal, but, you know, there's thousands and thousands of not ideal experiments may be more informative and cheaper to run than just one experiment you design yourself.

So to start with, they note that, you know, very few things on eBay are actually unique. They said, you know, I guess interesting fact about eBay, for about a half of the items that are sold on eBay, you can find a nearly identical item sold by exactly the same seller. And so when they're talking about an experiment, an experiment meaning you have exactly the same seller selling a product with exactly the same listing title listed in the same category and subtitle. And they find 55 million seller experiments involving a total of 350 million listings.

For much of their analysis, they're restricting attention to experiments that have several different features. There are at least two items auctioned. There's also at least one successful posted price sale. They want a posted price sale because they're going to use that as a reference for what's the market price of this if you just sell it. So not through an auction. And even with those restrictions, they come up with 244,000 experiments that have 7.6 million listings.

So the idea is just that eBay sellers are just experimenting all the time with all kinds of things. And this is something we can take advantage of. The paper itself has multiple different sub papers discussing what to do with this. So one of the sub papers is how do "Buy It Now" prices affect outcomes?

On eBay, if you've never been on an eBay auction, it's like it's classic English ascending bid auction, but there's also this option of just click here to buy it for $125 and end the auction. And so one of the papers they do is how does having a Buy It Now price affect the outcome?

Second is, how does shipping fees affect outcomes? There's this interesting earlier paper by Hossain and Morgan that argued that consumers are fairly naive in dealing with shipping fees and that you make more money by having good with $9.95 shipping than good with $5.95 shipping. Just somehow you offer the same good, different shipping prices, people somehow focus on the item price, not on the shipping price. And so the plus shipping and handling you make more money selling with a high shipping and handling fee.

Anyway, they're going to do an experiment on that or many sellers have done experiments on that. And then they also have this sort of third paper about how and why the reserve prices affect revenues. I'm just going to talk about the third one because it feeds back to things I talked about in the auction literature about optimal reserve prices and auctions.

OK, here's just their motivating figure to explain eBay experiments. Like, so they did a search for "TaylorMade driver." TaylorMade is a brand of golf club manufacturer. You do a search for "TaylorMade driver" on eBay and you just see a screen like this just full of people selling new and used golf clubs. And it looks like there's just a zillion different golf clubs out there on eBay.

But if you do a more specific search. So for this specific search, they typed in "TaylorMade burner 09 driver 2009 golf club new quote 10RH." And so you do this very specific search and then what do you find? You find that there is one seller who uses that as his title when he sells TaylorMade drivers online. And so all the searches say "TaylorMade burner 09 driver 2009 golf club new 10.5 degrees right hand."

And what you'll notice in this thing is that the seller has these many, many auctions. This one is closing in 3 hours and 21 minutes. This one closes in 3 hours and 22 minutes. This one closes in 6 hours and 9 minutes. 6 hours and 9 minutes. 8 hours and 59. 9, 41.

So this seller is selling auctioning off eight of these in the next 12 hours. And you can see this seller is conducting an experiment. This golf club here-- maybe you can't see it, I can. This one has $7.99 shipping. This one has $9.99 shipping. This one has $9.99 shipping. This one has $7.99 shipping.

So here it's clear these have Buy It Now prices. This is just $125 plus $9.99. This $125 plus $7.99. If anyone saw these two listings, obviously no one would buy this one. But they're just mixed in with that big list. People may only see one versus the other. It's unclear what's going to happen.

Anyway, so this is one seller's experiment. And their idea is that there are many, many, many of these experiments out there. But this is one of the 244,000 they're analyzing.

AUDIENCE: What's the logic from the seller's perspective? Like, where does it come from?

GLENN ELLISON: [INAUDIBLE] well, so I guess the seller probably doesn't know whether a $7.99 shipping or a $9.99 shipping produces better profits. So they're probably just auctioning both of them off, and they may have been doing this for a few weeks and then stopped and settled on whichever one was better. Certainly this Buy It Now versus the auction, they're probably some people who prefer one to the other. And so you probably sell both ways.

But, you know, something you see in their data is you see people conducting auctions where the starting price in the auction is above the Buy It Now price. So it does seem like there are people who just see an auction and decide to enter it and other people who just are looking for a fixed price and buy it and not really paying attention to the opposite ones.

OK, anyway, so how do we think about effects of reserve prices in IPV auctions? If you remember basic theory of reserve prices in an English auction is the reserve price in English auction is the monopoly price to one bidder. And the logic was just, you're selling this good, once you found out what the second highest value is, the willingness to pay of the high value bidder is just the distribution truncated at the point at which you know their value is at least that. And so you're just monopoly pricing on the single bidder distribution.

And so it's just the optimal reserve price is the monopoly price, right? You have this distribution of values. Once everyone else has dropped out here, all you know is that you're selling to someone whose draw is from this is the second highest bid. Nope.

OK. Anyway, so you have this distribution of the second highest bid. You're pricing to that. You set the monopoly price on that demand curve.

They also note that if we want to think about this empirically, how does a reserve price effect-- what's the optimal reserve price? Here's another good way to think about-- so this is density F of v. If you want to think about how does a reserve price affect things, what I can do is just imagine an experiment where the seller has used many, many different reserve prices. Like, they've started the bidding at $1, $5, $10, 20, 30, 40, 50.

For every reserve price, you can just draw a graph, which is, if I use a reserve price of 5% of the expected sale price, what's the eventual price at which I sell the good, and what's the-- if I use a very low reserve price, I sell the good with probability 0.9 or 0.95, but I get a fairly low average price. And then if I sell it with a $20 reserve, maybe I sell the good with probability 0.8 and I get a higher average price. And if I sell it with a reserve that's 120% of the average price for the good, maybe I sell it with just some very small probability, but I do get a very high price of selling.

So what you can basically do is just look at every reserve price in the data and figure out what's the probability that the good sells and what's the expected price conditional on its selling, and that's basically just your demand curve. You can view that as a demand curve and you're just trying to monopoly price on that curve. So that's a second way to think about optimal reserve pricing is just draw this demand curve like object and think about what would be the monopoly price on that. And that would be the way to figure out what the optimal reserve price is.

So that would be an experimental way to figure out my optimal reserve price without, you know-- theoretically you want to just say, what's the distribution from which the values are being drawn. But without thinking about what the distribution from which values are being drawn, you could just make this curve monopoly price relative to it.

OK. You know, several reasons why this simple theory may not work in practice. One is entry costs. If you imagine there's some cost to a bidder of starting to bid in an auction and examining the good in detail. If there's a low price, they may be willing to pay the entry cost. If they see that the reserve price is $110, they may not bother to read all the fine print and enter that auction.

There could be some kind of common value component. This theory was, once I know everybody else's bid, my bid is unchanged. There's common values, the theory changes. And then they also note that there could be-- speaking of this theory and why this theory may not work. You could also have behavioral preferences, like loss aversion.

If you imagine the loss aversion is always like, I feel bad losing something that I have. If you start the bidding at a very low price and people get used to the idea of owning this fancy golf club because they see it being bid for $20, once they've got used to the fact that they have this fancy golf club, then they may feel loss aversion if they don't win the golf club. You might, by getting people tied up in the bidding, they may get used to the thought of winning and then they may bid more in the end.

Anyway, to examine this, what they do is they find a data set where they found 19,000 experiments, where there are several features. One is there's variation in the start price, there's free shipping, which just simplifies things. There's no secret reserve price, no Buy It Now auction. So these are just straight, plain vanilla auctions of a good, many different times.

And so about 25 listings per experiment. So again, these are just, you know, sellers experimenting with, if I'm selling $125 golf club, should I start the bidding at $5, or $100, or 50, or something in between?

Figure 5 of the paper shows that, yes, sellers are doing this a lot. So let's look at a typical category like goods that cost between $30 and $100. And you can see goods where that cost, this is like the average price at which they're sold when they sell them through just a fixed price sale. And you can see goods that sell for $30 to $100, 20% of the sellers start them out at 0 in these experiments and then 37% start them out at between 85 to 100% of the average fixed price sale. 17% choose 60 to 80%. 13% are actually starting them off above the average Buy It Now average fixed price sales.

But we get this sort of bimodal distribution. Notably, there are very few sellers starting out in here. It seems like if you're going to go low, you just go all the way down to the bottom and start at almost 0.

And then there's with-- a lot of within experiment variation. So if you look at auctions in which their sellers are sometimes selling it at less than 5%, those same sellers were sometimes selling it at less than 5% are also sometimes selling 5 to 45%, 45 to 85%, 85 to 100%, 100 to 120%, greater than 120%. So there are just many experiments where sellers are trying out the whole range themselves. This is our experimental paper that we get to run without acting to actually pay for the experiments.

OK. So does the standard auction demand theory work? And the answer is no. There's one interesting feature, which is the demand curve turns up at very low prices. So remember, I drew this curve for you, saying, if I start the good at only $5, if I start the good at $1-- if I start the good at $1, I sell it almost entirely for sure, but I'm going to have a lower average price than if I start the good at a higher price.

What they find is that the curve doesn't look like this. The curve looks like this.

So as I start at a lower and lower price, not only do I sell with a higher probability, I sell at a higher price, conditional on selling. So I'm up here, I'm selling at 90% probability and I'm getting a higher price than I would get if I chose a reserve price that only sold with 70% probability.

So the demand curve turns up, suggesting that I put the $100 item out there, start the bidding at $1, lots of people see it at $1. 10 hours before the auction is over, they all start following the auction. They all start bidding. I get more entry than I would if I started at it at $20 instead. So the curve, instead of looking like this, looks like this instead.

Obviously, how do we do monopoly pricing? Like, the way to think about monopoly pricing is you do marginal revenue equals marginal cost. So they take this demand curve, they construct the marginal revenue curve, and the marginal revenue curve also turns up. Like, we're used to marginal revenue is downward sloping. You draw marginal cost across. Where marginal revenue hits marginal cost, that's the monopoly price. What do you do when marginal revenue turns up?

So if this is my marginal revenue curve, my marginal cost is flat. That's my normal monopoly pricing. What do I do if marginal revenue turns up? Then I can draw the line across where marginal revenue equals marginal cost. This is a locally optimal monopoly price. But the solution is either going to be this point or this point.

You know, it doesn't have to intersect again. It could be that you get a, here your revenue's increasing, so it may be that you want to price there. You may want to price here. That may be the global optimum or this could be the global optimum, but you're going to do one or the other.

Anyway. And so that's, I guess, in some sense the-- this is how they explain what's going on in the data. If you remember in the data, there was a lot of firms-- figure 5. Like, there were a lot of firms charging fairly high reserve prices and then some firms charging very low reserve prices. And so their thought is that what's going on in the data is probably something like this, where the marginal cost is pretty high. So you either set a quite high reserve price and this is the optimum or you set this right end point. And for different sellers, different ones are optimal.

And in general, if your marginal cost is high, you're going to just set the high reserve price. Your marginal cost is low, you're going to be setting the right end point.

You know, and then the paper tries to do other predictions. It tries to have a test of upper tail hypothesis of, does the upper tail of the price distribution look the same regardless of what reserve price you use? And that test only works so well. So it does seem like the classic theory, this sort of upper tail distribution is just what you get by truncating isn't quite right here. And it does seem to enter-- it's entry costs are important for online auctions, I think is the main message there. OK, questions?

Yeah. And obviously, I think that's something, you know-- that work was done inside eBay. For a long time, eBay actually had this sort of very, very open policy to academics. And we actually had several students, like, after they finished their PhDs here, go to a postdoc for a year and stay at eBay because eBay would let you sit there and operate inside the company and do all kinds of stuff with their full data set. That seems to have faded away three or four years ago. But obviously one could still scrape eBay and do the same kinds of things.

OK, so then the second topic that I guess is common to the rest of the papers I'm going to talk about is just understanding how online retail works and why it's successful. You know, if you read people writing about E-commerce early in the internet era, there was this question of, is online retail ever going to really take off and displace traditional retail? And the biggest argument against online retail taking off was that it's just much more expensive.

You know, like, if you look at the long history of retail in America, you know, in the 1800s, people shopped in department stores. And then there was this golden era of catalog retail where, like the Sears catalog in some sense brought products to-- you know, they were only formally available in cities to all of America. And they had like this phone book, you know, 1,000-page thick catalog that they would mail you. You could buy anything from the Sears catalog.

You know, the Sears catalog just took off in the, I guess it's late 1800s after there were trains, or whatever, that would bring packages to people. When the car was invented, the Sears catalog just went into a prolonged decline. And Sears opened, you know, massive, you know, not the size of a Walmart, but, you know, massive department stores in many US cities. And again, the catalog business declined. The physical retail business took off.

Thoughts on why that was. Catalog retail was just always much more expensive than traditional retail. Like, if you look at Walmart, Walmart's cost of selling goods is roughly 20% of the merchandise that they sell. Catalog retail had costs more like 40 or 50%. And why does it cost Walmart only 20% when it costs catalog retail 50%?

You know, it's because catalog retail, you need to have stuff in a warehouse. Every single item that gets sold, someone needs to put it in a box, put the box in the mail, pay for the mail. Whereas Walmart's cost of putting-- like, the simplest example of things that should never be online.

Like, a kiddie swimming pool, one of those plastic four-foot diameter swimming pools. You can put a whole bunch of four-foot diameter plastic swimming pools, stack them together, put them in a truck, take them to the Walmart, and the cost of getting them there is, you know, $0.50 or something like that. You try to put a kiddie swimming pool in the mail and send it to somebody, you've spent, you know, $30 mailing something that's only worth $5 or something.

And, you know, putting things in a truck is just much, much cheaper than handling every package individually. And you're getting some sense the free labor of people to just walk through the store and take it off the shelf and carry it home.

Anyway, early on, Amazon was undercutting Barnes & Noble and selling books for lower than you would buy the books in a physical bookstore. And this question of, could Amazon ever actually make a profit selling things not at a loss? And if Amazon were pricing things at cost or above cost, could it be competitive with traditional retail?

One of the big arguments that was made early on was that, yes, Amazon doesn't need to have a cost advantage to sell things online. It could be that you have a cost disadvantage that people buy online because of the product variety. The product variety advantages are much bigger than the cost disadvantage. So therefore people buy online to get the book they want or whatever product they want and have a much bigger selection.

And they note that when Amazon launched, Amazon had one million titles. Obviously, it later expanded to several million titles. And a very large physical bookstore at the time would have 100,000 titles. You're typical bookstore might have 50,000 titles in it. Perhaps there's some superstores that would have 200,000. But Amazon had many, many more books, even from the very beginning, than physical bookstores did.

And so what Brynjolfsson, Hu, and Smith tried to do is just say, could we estimate how big the product variety benefit is by just, in some sense, looking at this demand curve? Like, finding out what fraction of the sales of Amazon are to books that are beyond the 100,000th book? So let's say there are 100,000 books that are available in your physical bookstore. There are millions of other books that are not available in a physical bookstore. How much utility do people get from buying all those books from 100,000 out to 2 million?

And the way they do that is to just basically convert Amazon sales ranks to quantities, figure out how many copies of every book are being sold, and then just count up how many books-- what fraction of Amazon sales are books outside the top 100,000 and then multiply them by just a consumer surplus estimate per title. If you have a general sense of the demand elasticity, you know the average price, you know the consumer, you can just guess what the consumer surplus is and estimate that.

And what they show in this figure is that if you look at books that are outside the top 100,000, there seems to be this very, very thick or very, very slow decline in sales. So book number 100,000 is selling 1 and 1/2 copies a week. Book number 200,000 is selling 0.8. Book number 400 is selling 0.6 copies a week. Book number 6 is selling like 0.4.

But then when you get to book 2.2 million, book 2.2 million is not selling that many fewer copies than book 600,000. And so because this tail doesn't just go to 0 very quickly, there are lots of sales on Amazon that are these extra books. And the thought is that all these extra books are contributing a lot to consumer surplus.

What they estimate in their paper is that the product variety back in 2000, when Amazon was still young, was adding $700 million to a billion to consumer surplus, and that was about-- Amazon was lowering prices, I think it was by-- Amazon was 5 to 10% cheaper than physical bookstores at the time. And what they find is that look at the gain from the lower prices, the 5 to 10% lower prices. Look at the gains from the consumer surplus.

And the gains from the lower prices were only about $100 million a year. The gains from the consumer surplus is like $1 billion a year. So consumer surplus is 10 times as important as the price cut. And so if you believe that, if consumer surplus is 10 times as important as a price cut, you could take away the $100 million or even have a negative $100 million or negative $200 million, and they say that Amazon really could raise prices and still be a positive value for consumers because the product variety effect.

It's a very nice paper. The calculation itself is a little speculative in that it is a-- we don't have a lot of evidence. Like, if you think about it, like, any of these books that are sold very rarely, it's very hard to get much evidence of what the demand curves look like for them because they're hardly selling any copies. You are extrapolating on some things there.

So then the paper I want to talk about more is Quan and Williams'*Product Variety, Market Heterogeneity.* And this is an argument that the Brynjolfsson, Hu, and Smith paper may really be overestimating product variety advantages. And the reason they would argue this could be really overestimating profit variety advantages is because different physical bookstores carry different books.

So, for instance, you know, people at the time needed to get textbooks for all of their graduate economics classes. Where would you get textbooks for graduate economics classes? None of them are in the top 100,000 most popular books in the United States. However, you go to the MIT Coop in the week before classes start, and they would have all of the graduate economics textbooks and all of the other textbooks for all of your other classes would be in the MIT Coop.

And, you know, not just the graduate economics textbooks from MIT Coop, but the MIT Coop also has a big selection of, you know, technical books. And if you just want books about computer programming or whatever, not for your classes, they have lots of those books in that bookstores. And then there are other bookstores in Boston where if you're interested in art that have lots of books about art and lots of big coffee table books or whatever.

And so their thought was that-- the Quan and Williams argument is that you really could be overestimating what Amazon is doing because there are many different bookstores and many different bookstores have many different books. So the gain from the online over going to the bookstore that you know would have the books you want could be much smaller.

They study this not in the case of books. They pick a different industry. Their industry is shoes. They don't say it, one suspects that they have-- they had a friend who worked at Zappos, because they are working with what they describe as a very large online shoe store in the year 2012, 2013.

And just here's one example of why we would think that physical shoe stores would carry different things. This graph is entirely done with sales on the online site. But what they graph is, this is a state level analysis for, I think, 49 states. And what they put on this is just a dot for every state of what fraction of the shoes that are sold to that state are boots and what fraction are sandals. And what do you notice on the right side of this graph?

We take the two warmest US states, which would be Hawaii and Florida. And in Hawaii and Florida, people don't buy a whole lot of boots and they buy a lot of sandals. I guess especially in Florida, people buy a lot of sandals.

And then you go down here and you look at what people are buying in the coldest states. So it's, I don't know if that's North Dakota, Wyoming, Minnesota, places like that, people buy a lot of boots in North Dakota. People buy fewer sandals in North Dakota.

Although I guess sort of surprising things you learned from this. It turns out people in North Dakota and Wyoming actually buy more sandals than people in middle temperature states. I guess it may be-- I don't know. Maybe if you live in a cold place, you go on vacation to Florida and you need to buy sandals. Or if you live in a medium temperature state, you just don't go on as many vacations to warm weather states. I don't know.

But anyway, but their thought from this graph is if you look at if this is what happens for online retailers. You imagine that you walk into-- even take a national chain store, you imagine that the national chains in their North Dakota stores, they have lots of boots on the shelves, and in their Florida stores, they have very few winter boots on the shelves. And so this would just be an example.

You know, this doesn't prove that local stores tailor their selections to what customers want. But this at least suggests that local stores ought to tailor their selections to their customers, and that ought to give you back some of the gains that we think is Amazon's gains from product variety should be gains from product variety at these stores.

You know, the paper has a second data set where they also got access to the inventories of local stores at Macy's and Payless Shoe stores, which are just two chain retail stores. And they find that there is substantial differences in the SKUs carried at the different stores. They don't really have enough data to say whether it's just completely random or how much of it is this store carrying boots and this one carrying sandals, or this one having just different brands from the others.

But what they find is that especially at the Macy's, there are almost no shoes that all the Macy's have. Different Macy's just have different shoes. There are a lot of them, shoes, that are only found in one or two stores. Very few that are found in 100% of stores.

At the Payless, there are a lot of store shoes that are found in every Payless and a lot of shoes that are found in like one or two Paylesses.

OK, anyway, so how do they estimate? So what they want to do is estimate how important is this variation in consumer preferences and to estimate how important if stores just match their local preferences, would stores be able to get most of the benefits that you get from Amazon?

So what they do is demand estimation, it's going to be a standard kind of nested logit demand. So this is utility, customer I, location L, buying product J. It's going to be a mean utility of product J in location L, plus sort of nested logit models, there's a shock. You have this idiosyncratic preference of a customer for all products in category J. So if I'm looking for snow boots, I have a snow boot specific shock, and then a weight.

I have two shocks. One is the category-specific shock, and then one is the brand-specific shock. And they're going to estimate this lambda parameter so they can tell whether people's shocks are mostly category-specific or mostly product-specific.

And then what they're going to assume is that the mean utility provided by product J in location L is a function of the product characteristics, a function of the product price. And then this other xi sub LJ term, which that's the idiosyncratic preference for this product in location L.

They don't have the temperature here. So the people in Florida don't like boots would be here or people in Hawaii like sandals would be here. But then you could imagine there's also just like these mayonnaise-like preferences. Like, there's just UGGs boots are popular in New England. They're unfashionable in California. Or this brand of shoes is popular here, it's not popular here. All of that would go into the xi LJs.

A real challenge to estimating demand is that most products, you know, like there's a very, very large number of shoes in the world. And it turns out that most shoes in most cities in most months have zero sales. And so dealing with quantity data when you have zeros is very difficult.

If you remember, one of the things we-- standard approaches people do in the load to demand estimation is you have, you know, log of the share of product J minus log of the share of the outside good was X, J, beta minus Alpha, P, J, something like that. You convert to the market shares to get back to the product characteristics, and then you can run IV regressions to estimate demand. You know, you can't do these log transformations when you have a whole bunch of zeros in your data.

And so what do you do when you have a whole bunch of zeros in your data? Anyway, what they do in this paper is discuss a GMM approach to try to deal with all the zeros You also don't want to throw out-- another thing people do is sometimes just throw out all the zeros. You clearly don't want to throw out all the zeros because the goods that have zero sales in one month will have positive sales in the next month. They all have positive mean utilities. You don't want to remove them and say, no one wants to buy from those goods.

In particular, what they note is that the number of shoes that are sold by this online website in a particular city, in a particular month is smaller than the inventory of a large shoe store. So if you actually were to take the realized shoe sales and say, those people's deltas are only based on the shoes that actually got bought, what you'd conclude is that a single physical store could actually supply the entire market. And so it's just, it's very important to correctly deal with all of these shoes that are in the store and have zero sales in a month.

But, you know, people are buying. You just have these many products all with a fraction of expected sales. But those are what people buy among. And so they have to deal with the zeros.

Anyway, one thing you'll notice in the paper, like, they spend a while talking about how to deal with demand estimation when you have lots of zeros. If you ever have a paper where you're trying to do demand estimation with lots of zeros, they discuss their method and other methods that people have used to deal with zeros. I think theirs is thought to be a fairly good method for doing demand estimation when you have a whole bunch of zeros.

But so anyway, the basic idea is they aggregate sales to the location month level. They have 213 locations. Most of the 213 locations are metropolitan areas. But then in addition, they have each metropolitan area as a location observation, and then they have the rest of the state not in metropolitan areas as an observation.

Anyway, so they have 213 different locations. They do something that's kind of like a BLP estimation except dealing with the zeros. So they aggregate sales. They do something very much like, I don't know if you remember back to Tobias's lecture, you're BLP instruments for price are some function of the characteristics of all the competing brands. So they have not a very convincing instrumentation strategy, but that's what's available.

So they have a BLP style instrumentation strategy dealing with the zeros. And it's not a pure cross section because there's price variation, because the prices vary across months. And you can see when this shoe is on sale in some month, do sales go up. Product reviews are also time varying. Shoes come out and they have no reviews over time. They build up reviews. Obviously, higher reviews are good for sales.

It turns out, though, also zero reviews are good for sales or positively associated with sales as well. And this may be because zero reviews means it's a brand new product that's just being introduced with the latest fashion or that's being advertised or something like that.

Anyway, and then the products, that is also time varying. So anyway, with all of that, they're able to estimate all these coefficients. And they can estimate a distribution of the delta LJ's. And the delta LJ's indicate that there is a lot of cross-market heterogeneity in preferences.

And, you know, you have to have faith that they're dealing with the zeros appropriately to deal with it. It's not just in some markets these shoes actually have sales, in some markets these have actually have sales. But I think they've done a fairly good job of trying to argue that, no, there's underlying cross sectional variation in the propensities to purchase different shoes in different places.

And again, I think some of that is like the boots versus sandals approach. And some of that is just, you know, shoes are a fashion good. And there are brands of sneakers, brands of shoes, brands of whatever that are more or less fashionable in different cities. For whatever reason, this has become something people wear and people are willing to pay more for those.

OK, so how do they estimate the gains from product variety? Estimated gains from product variety, this is done with a counterfactual. So they know all the sales that Zappos, or whatever the retailer is, made. And then they think about a counterfactual world where instead of choosing from all the shoes at Zappos, people could only choose from a limited supply. What do they do for the limited supply?

Well, the first thing they do is they use the data on the Macy's and Payless Shoe Stores to estimate the total number of products available in each market as a function of the population. So they would be doing, how many different brands of shoes are available in Boston? Which you get by looking at the overlap of all the Macy's in Boston and the overlap of the Macy's in Boston with the Payless Shoe Stores in Boston.

And then they have census data knowing how many shoe stores there are in Boston and how large they are in Boston. They assume that they have the same overlap as they're seeing in the Macy's and Paylesses. And they're trying to estimate how many total shoes are there in the Boston area that you can buy at physical retail stores. Obviously, that's a bit of an aggressive assumption, and that is assuming that people could actually know which local Boston shoe store to go to buy whatever particular brand of boots they want.

And then what they assume is that offline consumers have the choice of the NL products that have the highest delta LJ's to choose from so that the physical stores in each state or each city carry the most popular brands. So again, it's kind of an aggressive assumption in that, you know, local stores are tailoring their demand to what people want and they're tailoring it perfectly. They're picking the most popular brands.

And so, you know, it's both we observe. It's both we know where to buy shoes physically, depending on what we want, which store to go to, and the shoe stores are doing perfect tailoring and carrying the most popular shoes, not making mistakes.

But anyway, so subject to that caveat, as the paper suggests, the online product variety benefits are much smaller than what Brynjolfsson, Hu, and Smith said. And they're much smaller in two different senses. The first is that if you look at the gain from product variety, they argue that about half of the gain from product variety would be provided anyway by local stores by just having heterogeneity in local stores. So if you think about this sort of continuum.

Let me do the Brynjolfsson, Hu, and Smith numbers. So Brynjolfsson, Hu, and Smith are all looking at sort of, what's the consumer surplus from there being one bookstore that has 100,000 titles? And then we have this sort of Amazon with 2 million titles and they're arguing that this is 10. This is like, if this is 1, this is 10. You get 10 times-- I'm putting the consumer surplus in this axis.

And then what Quan and Williams are saying is that we have-- so this is-- Quan and Williams saying, well, our benchmark is going to be not one bookstore, but all Boston bookstores. They don't know what all Boston bookstores have, but they just assume that the Boston bookstores have these non-overlapping sets of books and the non-overlapping sets of books tailored to what the MIT students want, tailored to what everybody else in Boston wants.

And what they find is that, you know-- well, they do all shoe stores, but all shoe stores thing compared to the one shoe store, they're arguing that all shoe stores get you about half of the way to the Amazon rankings. So in some sense that Quan and Williams may be overstating the gains by 100% or something.

And so that's one point. And then the second point of the paper that is not emphasized as much in just in the title and the introduction is that they also find that the gains from product variety are just much smaller than are reported elsewhere.

So that whereas Brynjolfsson, Hu, and Smith are saying that these gains are 10 times the gain from a 5% price cut. What they find is that the gains in the shoe store data look like they're, you know, 2 times the benefit of a-- so anyway, sorry. If I make this 1, this is 2. So the gain is just equivalent to a 5. The gain is like equivalent to a 5% price cut.

So it's like this number is here, this number is here. So what they're arguing is that the-- actually, no, I take that back. It's this one compared to this one I think is roughly equivalent to a 5% price cut. So what they're finding is that the gains from product variety are just much, much smaller in shoes than Brynjolfsson, Hu, and Smith said they were in books. And because the gains are so much smaller, and you take out this thing is halfway there already, what you're getting is the product variety gains are really only like equivalent to a 5% price cut.

So it's saying, at least in the shoe sense, product variety gains are, you know, 10 times smaller than you would have might have thought from the Brynjolfsson, Hu, and Smith paper. And obviously, you know, is that because shoes are very different from books? That's one possibility.

You know, like, every book is very, very different from every other book. You could imagine, once you've got a million shoes, the ability to make a million first shoe that's very different from all existing shoes is very hard. Whereas every new book is just completely different text from every other book. There may be more inherent variety in books than there are in shoes, but I think it may also be that they may have done a better job of dealing with estimating what's true product variety versus spurious product variety. You know, like the-- I don't if I remember.

Like, I've discussed, like, when you're trying to estimate consumer variety in demand, like, consumer welfare from demand and you have heterogeneous products, you know, everyone's always doing this X, J, Beta minus alpha, P, J plus Epsilon, IJ. And they're always assuming these Epsilon IJ's are logit distributed, which has a fairly thick tail, and that every Epsilon IJ is independent of every other Epsilon IJ.

Again, once you get to the millionth different shoe, is it really an Epsilon IJ that's completely independent of your taste for every other shoe out there? It's highly likely those Epsilons are highly correlated across shoes. At least by having the-- you know, they've got those category-specific shocks by putting in the category-specific shocks and assuming the magnitude and doing a careful job of thinking about the zeros, they may be avoiding overstating how big all these Epsilon IJ's are and all this surplus we're getting from all these new different products.

And I guess one other paper I thought I'd mention-- one other thing I thought I'd mentioned in the paper is that they know that Macy's and Payless carry different shoes in different stores. They ask, how important is it for stores to tailor their selections? And what they find is that there's a gain that firms should do it. It's not that large.

The gain from customizing stores is estimated to be about 6% in revenue. So you could just carry all the standard shoes or you could tailor the demand in your city. You tailor the demand in your city, you get 6% more profit. So it seems like it's worth doing. Firms should want to do this, especially if they have thin profit margins. But it's boundedly-- tailoring is boundedly profitable? Yeah?

AUDIENCE: Was this paper actually response to names on paper? Or was it just kind of disjointed? Because couldn't you just do the same exercise they did in the Amazon paper but with shoes [INAUDIBLE] exchange?

GLENN ELLISON: Yeah, so certainly they could have-- you know, so they actually have-- I don't know. They do several different calculations in the paper. Like, suppose we estimate demand this way. Suppose we estimate demand this way. Suppose we estimate demand this way.

You know, in some sense doing demand worse and worse and worse, doing worse and worse demand estimation. And they do present several things. If you've done this estimate, instead you'd get numbers that are two times as big. If you do these estimates instead, you get things that are five times as big. So they do get-- you know, it's not exactly the same calculation as the Amazon data, but they do something that's more like the naive calculations where you treat everything as national market shares and use logit models and show that you can get things many times bigger than the estimates they've gotten.

Yeah, I think these are the things that obviously you couldn't have done with the Amazon data. Like, the Brynjolfsson, Hu, and Smith didn't have anything about what is in local bookstores and how much overlap there is across local bookstores and stuff like that.

But certainly, the paper is very much written in response to literature saying there are these really big private-- really big product variety benefits. And I think the paper is more about the benefits are smaller because of the variety in retail than the benefits are smaller if you just do the econometrics correctly. But I think it is about both.

OK. So anyway, final paper I want to cover. This is a working paper that Sara and I have, which I guess has two different motivations. You know, one is also this is a paper about trying to estimate how big product variety benefits are and how big product variety benefits are in an area where we think the product variety benefits would be bigger, which is in used books.

And then the other is also sometimes you can say it is a, you know, Chicago school critique of our search and obfuscation paper. Where in the search and obfuscation paper we argued that you see these not super low and dispersed prices for computer components because the firms have an incentive to obfuscate what they're selling the price is at and invent these add-ons or invent complicated things that raise search costs and keep prices high and disperses as they are in a search cost model rather than letting perfect competition come out here.

This paper is going to provide an alternate explanation for why prices can be high in dispersed. It's going to be that in some sense, prices are high and dispersed for two different reasons. One is prices are high on the internet because the internet is so good at matching products to people. And if you think about a used book that you had as a child.

You know, and then you have a used book as a child, you now have a niece or nephew. You want to buy them the same picture book that you remember from when you were a kid. You're willing to pay many dollars for that book.

There are a million other people out there who didn't remember that book when they were kids. There are probably hundreds of copies of it sitting in physical used bookstores with $0.10 marked inside the cover that no one is ever buying. If we get those books in the hands of the right people, the willingness to pay is way, way higher.

In fact, actually, we say the paper started when Sara was writing a paper on pharmaceuticals and needed some obscure pharmaceutical book that someone had written in the 1970s that the MIT libraries didn't carry. You know, websites selling used books online were new.

She actually went on to websites selling used books, found the book, and it's like, wow, I can find this book. I've been trying to find this. It's only $15. I can pay it out of my research account. I'm thrilled to get it for $15 instead of $100.

Gets it and then got the book in like $0.25 was penciled inside the front cover of it. And she realized it had been sitting in a bookstore in Iowa City since some economics professor retired for the last decade, and no one wanted to pay $0.25 for this. And she was thrilled to get it for $15 and realizing that this is potential huge increase in social welfare of match quality.

Anyway, so part of what we're doing is underestimate that match quality. And then the second thought in the paper is that, it's nice to explain two things with one story. Like, prices are high and dispersed because of search costs. But we think there are two stories here. Prices are high because of match quality. And prices are dispersed because some websites are better at selling things than others, and the websites that are better at selling things are able to command much higher prices than the ones that have limited attention.

Anyway, so what does the paper do? It's a three-part paper. First, it's developing simple models to basically illustrate the ideas. You know, one and two are really, again, just me always aspiring to be like Hendrix and Porter and having this theory, having predictions of the theory, testing the predictions of the theory to convince you that it's right. And then the back part of the paper is structural estimates to try to say how big are the welfare gains from improved accessibility of used books sold online versus in physical stores.

OK. So first, let me start with the basic theory. And this is a simple model of dynamic monopoly pricing of a variety that we don't-- I guess I never covered in the IO class. But so imagine that you have a seller and the seller just has one unit to sell. There's no production here. So this is a unique item. You're just trying to sell it.

And then you have customers who have values VJ drawn from some distribution F and they just arrive at some Poisson rate gamma. They walk into your store. When they walk into your store and are interested in this pharmaceutical textbook, you have to either sell it to them immediately or not at all. Because if they walk into your store, they see the book, they look at the price, they don't want to buy it, they walk out. You'll never see them again.

How should you set your prices? Well, you know, the demand from these are one-time consumers who are myopic. So the demand at price P is just 1 minus capital F of P, the probability that the consumer's value is above P. So the expected profit is the expectation of setting a price of P and leaving it there is the expectation of pe to the minus rt, where t is the random time at which you will sell the book.

In this model, as long as you set any p that's less than the upper bound of F, you'll always sell the book eventually, you just may sell it a million years in the future and the [INAUDIBLE] rt will eat up all your profits.

OK, so what we're going to do is integrate pe to the minus rt times the density of the probability of selling the book in the interval tt plus dt. Gamma is the Poisson arrival rate. Gamma D of P is the Poisson arrival rate of people willing to pay at least P. And so the density of the time you sell it is gamma D of pe to the minus gamma D of pt.

You may have seen like if you have a Poisson arrival rate, gamma, you know, the density of arrivals is gamma e to the minus gamma t. And so the arrival rate of people willing to pay p is this. So this is the density of the arrival distribution. And it turns out this is a nice, well-behaved integral. And so the price is a function of p as this expression.

It's p times gamma D of p over r plus gamma D of p. And so as gamma D of p goes to 1, you're selling it almost instantly and your profit is p. As gamma D of p goes to zero, the r over outweighs the gamma D of p, and you basically are getting something that's going towards zero because this is the discounting and this is the price at which you sell it.

And so we can think of monopoly pricing in two different ways. One way to think about it is just the monopoly price is the maximizer of this expression. Second way to think about it is like a dynamic programming approach. This good doesn't have any production costs, but it has an opportunity cost because you can only sell it once. And so if pi star is your equilibrium expected profit from holding one unit of this good, you're thinking about someone's in my store now, what price should I quote them?

What I want to do is maximize p minus pi star D of p. I'm maximizing price minus opportunity cost times the demand. And so another way to think about it is that pm has to solve this where pi star is pi of pm. So it's two different ways to think about monopoly pricing for a unique item when you have a customer arrivals.

Observations from this monopoly pricing model. Monopoly prices increasing in the customer arrival rate. Why is the increase in customer arrival rate? Because the higher is the customer arrival rate, the higher is your opportunity cost. This is, again, a 14.121 standard exam question. Your monopoly price is increasing in your cost, full stop because of this function as increasing differences.

So monopoly price is increasing opportunity cost, therefore dynamic monopoly price increases in arrival rate because arrival rate increases your opportunity set. Opportunity cost goes up. So anyway, if consumers start to arrive at higher rates online than they did in physical stores, you're going to get higher prices. And if there's a rival rate heterogeneity across online stores, that's going to lead to price dispersion. And stores that are frequently visited by online consumers are going to be able to set higher prices because they're more patient. Those that have fewer visitors are going to set lower prices.

How sensitive prices are to arrival rates depends on the thickness of the upper tail. If you have an extremely thick tailed distribution, like constant times p to the minus 1, then the monopoly price increases almost linearly in arrival rates. If you have a very thin tailed distribution with a greater than 1, you know, things increase more slowly with arrival rates.

And then with if demand has this constant elasticity form, like this, then you get this very unusual theorem. So the theorem is, suppose the distribution of consumer valuations is such that demand has a truncated, constant elasticity form. Suppose the monopoly price is not at the kink of this demand curve at 1, but it's on the downward sloping part. Then the monopoly price maximizes social welfare, and social welfare is the price numerically, just numerically.

Like, I want to say, what's the expected producer surplus plus consumer surplus from selling this book? The answer is, if that book is being sold for $12, the answer expected welfare is $12. So in some sense, like, it's the most powerful structural estimation theorem of all time, saying, if you want to estimate the social welfare from dynamic sales, you just write down the price.

You don't need to compute the demand elasticities. You don't need to figure out consumer surplus, nothing. You just write down the price and that is the welfare. And then there's also this optimality theorem that there's no monopoly distortion.

And so, like, why does this work out? You know, obviously the monopoly price maximizes monopolist profits. But what is consumer surplus? So if I have this distribution of consumer valuations. So I have this distribution of consumer valuations, F of V, and imagine it looks something like that. And suppose I set some price, pm. OK.

When I sell the good, there's some consumer surplus. What's the consumer surplus? That's the expectation of V minus pm conditional on V greater than pm. How does that vary with the monopoly price? Like, in a linear demand curve, as I set the price really high, you know, the expected value conditional on being bigger than this is almost nothing. And so consumer surplus drops substantially with the monopoly price.

But in a constant elasticity model, the expected value of V minus pm conditional on V being bigger than pm is proportional to pm. So in a constant elasticity demand curve, consumer surplus is proportional to the price. So if you set a really, really high price, you don't sell for a very long time, and so the pe to the minus rt part gets really small. But when you do sell it, the person because they're willing to pay $1 million for this book, their expected surplus is also on the order of $1 million.

And so with constant elasticity demand, the consumer surplus is just always proportional to the price being paid, expected proportional to price. Therefore, by maximizing expected price paid discounted expected price, you're also maximizing discounted expected social welfare. So I view this as more a-- I mean, it's a neat estimation trick or something, but I view it more as a cautionary tale.

The cautionary tale being, if in this demand estimation, you can't tell the demand apart from-- if you can't tell the valuation curve apart from a constant elasticity demand curve, then you're going to end up finding that the expected welfare is the price, regardless of what the demand elasticity is and everything else are. So it can be a first approximation. And it's also, if you want a better approximation, you need to be able to tell the shape of this curve apart from a constant elasticity one, which is going to be hard to do when the upper tail because you don't see things in the upper tail very often.

OK. so now let me think about that's a monopoly model. Again, think about an oligopoly model. So you have two types of consumers. You have some people who just go straight. Like, they know some online bookstore, like Better World Books or Powell's Books. And they just go directly to PowellsBooks.com, look for a book there.

We're going to have some consumers who have this, or you do a Google search. Google search reveals the name of one bookstore for the title you want. You just go to that bookstore and you buy from the first, the one that's highest ranked on Google. And then we have other people who are going to be shoppers who are going to arrive at a Poisson rate and they observe the prices at all the stores.

Do I have a picture of AbeBooks? I don't have a picture of AbeBooks. OK. Anyway, we use data in the paper from a website called AbeBooks.com.

AbeBooks, it's a used book aggregator site. You can go on there. You could type in any book you want that's out of print. It will find 100 online bookstores that sell that particular used book. OK.

Actually, you know, AbeBooks used to be an independent company. It was purchased by Amazon. So if you go to the Amazon Buy Used button, Amazon Buy Used button is the AbeBooks listings. In some sense, Amazon invented the Buy Used button first. When they invented the Buy Used button, there was nothing there under Buy Used.

You'd click Buy Used and there were no copies available. And then they realized, well, why don't we just buy AbeBooks that has 10 million used books listed and put all their listings under the Buy Used button? So now if you go to the Buy Used button, you're seeing the same AbeBooks listings.

OK. Anyway, so our model is that there are sort of these two types of customers, the ones going to Amazon and clicking Buy Used and getting this curated list of all the things sorted by price, and then the people who are searching for a book by just typing it into Google or going directly to a used bookstore that they know online and finding it that way.

You know, we know that this model can have pure strategy dispersed price equilibrium, not so different from the dispersed price equilibrium I did when I did differentiated product search models. You know, the firms here, the dispersion is going to be the firms that have the higher arrival rates are going to set high prices and mostly serve the non-shoppers. If you don't have a lot of non-shoppers, you're going to set a low price and compete for the shoppers.

And when I'm thinking about, like, how does the online prices differ from the offline prices? It helps think that there are two reasons. One is the match quality reason. The higher arrival rates shift the distribution of monopoly prices to the right. The second is the competition effect, which is there's a desire to attract the shoppers. So the firms that have low arrival rates of non-shoppers are going to be competing against each other to try to sell in a more Bertrand-like environment to the people who are using the search engines.

So I think on the next page I have this illustration. So here's a numerical simulation of the model. So this dashed gray line would be the distribution of monopoly prices if there are exponentially distributed arrival rates. And so the firms with high arrival rates at high prices, those with low arrival rates at low prices.

If I increase the value of the good because people are getting a better matched good, the monopoly price distribution would shift to the right from the gray line to the blue line. But then if I add in this atom of shoppers who are just going to buy for the lowest price book and are fairly price insensitive, what you get is the distribution of prices. All the firms that are going to be setting the lowest prices are now Bertrand competing against the other firms who are also going to set very low prices. And so the low end of the price distribution gets pulled down.

The upper tail really doesn't get affected at all. The upper tail doesn't get affected at all because these firms were just monopoly pricing and serving the high value customers who happen on their website occasionally. They don't bother to compete for the shoppers. They just leave the high price and just exploit the non-shoppers.

And so what you get is this increase in dispersion, which comes from the upper tail really having moved up because people have higher values, and then the lower tail getting pulled down by the competition. And so we think those are the two things you should expect is that the online price distribution has this thicker upper tail relative to the offline price distribution, but then it also has more mass in the lower tail, perhaps.

OK, so how do we try to test this in the data? So we gathered a data set that had 335 books representative of traditional used bookstores' inventories. The way this worked is we just-- actually, Sara did some of this herself. We also just sent an RA to some physical bookstores with the instructions, I want you to go to the physical bookstores and just go to the first shelf, take every 17th book off the shelf, take that book, write down the title, what it is, write down the price, and then just go shelf to shelf to shelf picking out books.

So we have hundreds, what we call standard books in the data set. We also oversampled what we call local interest books. Like, if you go into a Boston bookstore, you'll notice there's a shelf of books about Boston and Boston stuff. And so what we said is, if there's a local interest section, so there's a Boston shelf in a Boston bookstore, one of the bookstores was in Atlanta, there if there's a book about Christmas in Georgia, you know, like, find the shelf that has the Georgia books and oversample those.

And then we also separate out popular books. Popular books are the things like-- I mean, not necessarily Harry Potter, but things that are popular enough that every used bookstore is probably going to have a copy. There are some things that these are former mega bestsellers. So it's not so rare to find these, whereas the standard books are all going to be completely non-overlapping across stores because just the average used book-- there are millions of books. The average used book is just no one's going to have another copy of that. You could search the bookstores all your life and probably not find another copy of a standard book.

OK. Anyway, so we have four data collections. What we did is we first just randomly selected these books from physical bookstores, and then we went online and searched for those books at AbeBooks.com. And then we noticed that in between the time we collected this and we wrote the paper, Amazon bought AbeBooks. And we thought, wow, that's a great second accidental opportunity here is that Amazon buying AbeBooks and putting the listings under Buy Used probably just dramatically increased the number of people who are looking at the curated search engine listings for them.

And so then after Amazon bought and incorporated the AbeBooks listings, we went out and collected the data again a second time. And we collected the data again a second time so that we would have a before and after on prices, before and after the AbeBooks listings. And then we went and collected the data fourth time a few months later. We realized it would be useful to have quantity data, and a way to get quantity data is to search a bookstore for the book, search that same bookstore two months later, see if they're still listing the book and infer if they're not still listing the book two months later, they may have sold the book.

Now they may have just also just taken it down, but it's a pain to take things down from your website. You probably take them down when you sell them. So anyway, that's so we have these four data collections. And last two is how we got quantity data.

So what are the predictions we're trying to test? Prediction 1 is that online prices should be more dispersed than offline prices and have a thicker upper tail. Prediction 2 is, local interest books may already be well matched by the offline stores, and so the online-offline gap should be smaller for the local interest books. In particular, I put this quote up here.

I found this online review of one of the books in our data set, the*Mount Vernon Street Warrens-- A Boston Story.* I don't know if you guys have been to Beacon Hill. Mount Vernon is one of the streets on Beacon Hill. It's a cute street or whatever, with the gas lamps and the brick sidewalks and stuff. Anyway, if you go into a local Boston bookstore, one of them, some RA found a copy of this *Mount Vernon Street Warrens.*

Anyway, if you read that book, like, what's your overwhelming thing? So this is, you know, "telling the stories of these children who became notable for eccentricity and philanthropy. Green focuses on Ned Warren, a homosexual and mover in the international movement of aestheticism, who was determined to lead a 'grand but blighted life.' This somewhat jumbled tale of a family sundering through greed and suicide."

You know, basically I read this and I'm like, I don't want to read that book. I mean, I don't know so much about what aestheticism was, but how much do I want to really read about greed and suicide in Boston in the 19th century? And the whole point is, in fact, nobody wants to read that book. So who is the high value purchaser of that book?

A tourist who's just been to Beacon Hill and seen Mount Vernon Street and finds a copy of a book about a family on Mount Vernon Street in the 19th century and thinks, wow, this will be great to read in my hotel room. And then you get partway through and get the suicide side, and you're like, this is not my vacation reading. And then you leave it there and then the hotel turns it back in and then sells it to the used bookstore, and then they sell it again to the next tourist or something.

But anyway, but so for these books, like, the online price is not really much higher because no one is out there searching for that book. But it is a good match to the local physical bookstore.

Anyway, other predictions, the online-offline gap should also be smaller for popular titles because there's a thinner upper tail for popular titles. There's not the match quality issue. And then Amazon's Buy Used feature should lead to more low prices in 2012 because it increases that competition for the lower tail.

Anyway, so those are the predictions we try to test. And let me just say basically the story is, almost all of these things seem to be true in the data. In each of these figures, the left is the distribution of offline prices. The right side is an estimate of the same object, an estimate of the average across titles of the distribution of prices for that title. So these should be estimates of the distribution of offline prices for these books and the distribution of online prices for these books. And if you see for the standard books, there's just this massively higher upper tail of prices online than there was offline.

If you go to the local interest books, the local interest books already had something of a tail online because the ones that are well-matched. That book that finds its way to the Beacon Hill bookstore is already well-matched, can charge a very high price offline. And so the offline-- you know, more of an upper tail offline. The online-offline gap is smaller for those local interest books.

Popular ones also works, but we skip that. This is a 2009-2012 comparison. 2009-2012 comparison, the gray bars are the 2009 price distribution. The white outlined bars for the 2012 distribution. And you see these distributions almost lie exactly on top of each other, with the exception that here on the standard titles, the number of standard titles sold at a very low price just shoots up in 2012 after the Amazon acquisition.

And so this is if there are many more people searching. And so that left tail just gets pushed down, but it doesn't affect the upper tail at all. And you see that in all three categories of books that the lower tail gets much thicker. And it just seems to be drawing from these slightly higher prices, as if it's competition is pulling down the bottom end.

And then, you know, the paper also then develops a structural model to estimate what are the welfare gains from selling books online. And our answer is, there seems to be a substantial welfare gain, but even in this used books case, it's not enormous. So we're more with Quan and Williams than we are with Brynjolfsson, Hu, and Smith in terms of what are the welfare gains.

What we find is that the estimated social welfare from selling one of these 2009 books in the offline world was $7.76. Estimated social welfare from selling it online was $16.87. So it's a more than doubling of welfare that comes from selling online, but it's only a more than doubling of welfare. It's not a massive, not a 10 times increase in welfare from selling these things online.

And if you look at the firm profits, the profit is coming from two different dimensions. It's like the average price is going from $12 to $16 when they're sold online and the books are being sold sooner. So instead of the e to the minus rt being 0.3, the e to the minus rt is 0.48. So the books are getting sold for higher prices and they're getting sold sooner. Those two things both make profit more than double.

We're also estimating that the consumer surplus, as best we can, is also more than doubling. And the consumer surplus is more than doubling because the people who are willing to buy $16 books get more utility than people who are willing to buy $12 books, and they're getting them sooner. So the consumer surplus is discounted by this instead of that. OK.

Sorry, I'm out of time. So [INAUDIBLE] you say, that's what I've got on retail for today. And then--