WEBVTT

NOTE duration:”00:56:08.9170000”

NOTE language:en-us

NOTE Confidence: 0.8496607

00:00:00.000 --> 00:00:01.197 So welcome everyone,

NOTE Confidence: 0.8496607

00:00:01.197 --> 00:00:03.990 it is my great pleasure to introduce

NOTE Confidence: 0.8496607

00:00:04.067 --> 00:00:05.887 our seminar speaker today,

NOTE Confidence: 0.8496607

00:00:05.890 --> 00:00:07.279 Doctor Elizabeth Tipton.

NOTE Confidence: 0.8496607

00:00:07.279 --> 00:00:09.131 She's an associate professor

NOTE Confidence: 0.8496607

00:00:09.131 --> 00:00:11.846 statistics the Co director of the

NOTE Confidence: 0.8496607

00:00:11.846 --> 00:00:13.936 statistics or evidence based policy

NOTE Confidence: 0.8496607

00:00:13.936 --> 00:00:16.274 and practice Center and a faculty

NOTE Confidence: 0.8496607

00:00:16.274 --> 00:00:18.428 fellow in the Institute for Policy

NOTE Confidence: 0.8496607

00:00:18.430 --> 00:00:20.278 Research at Northwestern University.

NOTE Confidence: 0.8496607

00:00:20.278 --> 00:00:22.588 Unducted sentence research focuses on

NOTE Confidence: 0.8496607

00:00:22.588 --> 00:00:24.917 the design and analysis of randomized

NOTE Confidence: 0.8496607

00:00:24.917 --> 00:00:27.773 experiments with a focus on issues for

NOTE Confidence: 0.8496607

00:00:27.773 --> 00:00:29.745 external validity and generalizability,

NOTE Confidence: 0.8496607

00:00:29.750 --> 00:00:32.788 as well as meta analysis with the

NOTE Confidence: 0.8496607

00:00:32.788 --> 00:00:35.259 focus on dependent effect sizes.

NOTE Confidence: 0.8496607

00:00:35.260 --> 00:00:37.140 Um, today she's going to share with us

NOTE Confidence: 0.8496607

00:00:37.140 --> 00:00:39.014 how to design randomized experiments

NOTE Confidence: 0.8496607

00:00:39.014 --> 00:00:41.264 to better understand treatment effects.

NOTE Confidence: 0.8496607

00:00:41.270 --> 00:00:42.602 Head virginity welcome best.

NOTE Confidence: 0.8496607

00:00:42.602 --> 00:00:43.934 The floor is yours.

NOTE Confidence: 0.8496607

00:00:43.940 --> 00:00:44.610 Thank you.

NOTE Confidence: 0.823528

00:00:44.610 --> 00:00:47.616 Thank you. I'm very excited to be here today.

NOTE Confidence: 0.823528

00:00:47.620 --> 00:00:50.044 I really wish I hear I wasn't talking

NOTE Confidence: 0.823528

00:00:50.044 --> 00:00:52.463 about my office slash closet and was

NOTE Confidence: 0.823528

00:00:52.463 --> 00:00:54.950 actually with you guys in person and

NOTE Confidence: 0.823528

00:00:54.950 --> 00:00:57.638 this is my first time doing slides where

NOTE Confidence: 0.823528

00:00:57.640 --> 00:01:00.304 I'm on the slide so it's a little.

NOTE Confidence: 0.823528

00:01:00.310 --> 00:01:01.458 It's a little strange.

NOTE Confidence: 0.823528

00:01:01.458 --> 00:01:03.712 I don't know what is the protocol

NOTE Confidence: 0.823528

00:01:03.712 --> 00:01:05.934 for questions. How do you guys?

NOTE Confidence: 0.823528

00:01:05.934 --> 00:01:08.160 How do you usually set this up?

NOTE Confidence: 0.823528

00:01:08.160 --> 00:01:09.620 Do people what's the norm?

NOTE Confidence: 0.823528

00:01:09.620 --> 00:01:11.678 Do you guys usually up jump in

NOTE Confidence: 0.823528

00:01:11.678 --> 00:01:13.418 with questions or save them for

NOTE Confidence: 0.85336065

00:01:13.420 --> 00:01:15.756 the end? So I think as you prefer,

NOTE Confidence: 0.85336065

00:01:15.760 --> 00:01:17.800 we can do either way. OK, I'm

NOTE Confidence: 0.85336065

00:01:17.800 --> 00:01:19.844 just I won't be very good at

NOTE Confidence: 0.85336065

00:01:19.844 --> 00:01:21.593 checking the chat, so if there's

NOTE Confidence: 0.85336065

00:01:21.593 --> 00:01:23.343 a question if somebody can just

NOTE Confidence: 0.85336065

00:01:23.343 --> 00:01:25.680 speak up that would be will do that.

NOTE Confidence: 0.85336065

00:01:25.680 --> 00:01:28.136 I'll do that on the chat. OK, thank you.

NOTE Confidence: 0.85336065

00:01:28.136 --> 00:01:30.542 OK so I just want to set out background

NOTE Confidence: 0.85336065

00:01:30.542 --> 00:01:32.690 for what I'm talking about today,

NOTE Confidence: 0.85336065

00:01:32.690 --> 00:01:34.310 which is I'm talking about randomized

NOTE Confidence: 0.85336065

00:01:34.310 --> 00:01:36.417 trials an I realized that in a

NOTE Confidence: 0.85336065

00:01:36.417 --> 00:01:38.740 Biostatistics Department, you guys.

NOTE Confidence: 0.85336065

00:01:38.740 --> 00:01:40.900 The idea that randomized trials are

NOTE Confidence: 0.85336065

00:01:40.900 --> 00:01:43.046 common is probably almost absurdly basic

NOTE Confidence: 0.85336065

00:01:43.046 --> 00:01:45.293 for the world that you operate in,

NOTE Confidence: 0.85336065

00:01:45.300 --> 00:01:47.452 but I do a lot of my statistical

NOTE Confidence: 0.85336065

00:01:47.452 --> 00:01:49.816 work in the areas of education and

NOTE Confidence: 0.85336065

00:01:49.816 --> 00:01:52.490 psychology and kind of in the field

NOTE Confidence: 0.85336065

00:01:52.490 --> 00:01:54.610 experiments world and those areas.

NOTE Confidence: 0.85336065

00:01:54.610 --> 00:01:55.990 Randomized trials have only

NOTE Confidence: 0.85336065

00:01:55.990 --> 00:01:57.025 become common really,

NOTE Confidence: 0.85336065

00:01:57.030 --> 00:02:00.420 I'd say the last 20 years.

NOTE Confidence: 0.85336065

00:02:00.420 --> 00:02:02.690 So almost 20 years ago,

NOTE Confidence: 0.85336065

00:02:02.690 --> 00:02:05.060 the Institute for Education Sciences

NOTE Confidence: 0.85336065

00:02:05.060 --> 00:02:07.913 was founded in the Department of

NOTE Confidence: 0.85336065

00:02:07.913 --> 00:02:09.948 Education in the US government,

NOTE Confidence: 0.85336065

00:02:09.950 --> 00:02:12.314 and that has funded almost 500

NOTE Confidence: 0.85336065

00:02:12.314 --> 00:02:14.523 what are called efficacy and

NOTE Confidence: 0.85336065

00:02:14.523 --> 00:02:16.759 effectiveness trials and education.

NOTE Confidence: 0.85336065

00:02:16.760 --> 00:02:19.030 Previous to that there were

NOTE Confidence: 0.85336065

00:02:19.030 --> 00:02:20.846 very few of these.

NOTE Confidence: 0.85336065

00:02:20.850 --> 00:02:24.066 There's also an increasing number of

NOTE Confidence: 0.85336065

00:02:24.066 --> 00:02:26.875 nudge experiments in social psychology

NOTE Confidence: 0.85336065

00:02:26.875 --> 00:02:30.907 experiments that are occurring in the world.

NOTE Confidence: 0.85336065

00:02:30.910 --> 00:02:34.210 I know that there's a lot of rain in mice

NOTE Confidence: 0.85336065

00:02:34.290 --> 00:02:37.590 trials occurring in developing countries,

NOTE Confidence: 0.85336065

00:02:37.590 --> 00:02:40.260 so this is late in parallel,

NOTE Confidence: 0.85336065

00:02:40.260 --> 00:02:42.190 maybe 2.

NOTE Confidence: 0.85336065

00:02:42.190 --> 00:02:43.156 In public health,

NOTE Confidence: 0.85336065

00:02:43.156 --> 00:02:44.766 they're being randomized trials there,

NOTE Confidence: 0.85336065

00:02:44.770 --> 00:02:46.499 so I'm just sort of pointing out

NOTE Confidence: 0.85336065

00:02:46.499 --> 00:02:48.180 that these are becoming increasingly

NOTE Confidence: 0.85336065

00:02:48.180 --> 00:02:49.936 common for policy decisions,

NOTE Confidence: 0.85336065

00:02:49.940 --> 00:02:52.588 not just individual decisions.

NOTE Confidence: 0.85336065

00:02:52.590 --> 00:02:55.582 But the trials as there as they are

NOTE Confidence: 0.85336065

00:02:55.582 --> 00:02:57.521 designed currently are not necessarily

NOTE Confidence: 0.85336065

00:02:57.521 --> 00:03:00.479 ideal ideal in the sense that they are

NOTE Confidence: 0.85336065

00:03:00.479 --> 00:03:03.241 not as big as we would like them to be.

NOTE Confidence: 0.85336065

00:03:03.241 --> 00:03:05.889 In order to be able to really explore

NOTE Confidence: 0.85336065

00:03:05.889 --> 00:03:08.072 the data well, there often in,

NOTE Confidence: 0.85336065

00:03:08.072 --> 00:03:08.760 you know,

NOTE Confidence: 0.85336065

00:03:08.760 --> 00:03:10.565 sort of somewhat small samples

NOTE Confidence: 0.85336065

00:03:10.565 --> 00:03:13.400 of clusters in the in the kind of

NOTE Confidence: 0.85336065

00:03:13.400 --> 00:03:15.633 education world that I work in it.

NOTE Confidence: 0.85336065

00:03:15.640 --> 00:03:17.608 They're very often just simple to

NOTE Confidence: 0.85336065

00:03:17.608 --> 00:03:19.420 arm designs 5050 treatment control.

NOTE Confidence: 0.85336065

00:03:19.420 --> 00:03:21.695 I much less common to see things

NOTE Confidence: 0.85336065

00:03:21.695 --> 00:03:23.888 like step wedge or smart designs,

NOTE Confidence: 0.85336065

00:03:23.890 --> 00:03:27.590 so those are trickling in, I think.

NOTE Confidence: 0.85336065

00:03:27.590 --> 00:03:29.854 And the goal of these the is often

NOTE Confidence: 0.85336065

00:03:29.854 --> 00:03:32.361 to get into some things like clearing

NOTE Confidence: 0.85336065

00:03:32.361 --> 00:03:35.171 House of some places so that policy

NOTE Confidence: 0.85336065

00:03:35.171 --> 00:03:37.781 making decision makers can use the

NOTE Confidence: 0.85336065

00:03:37.781 --> 00:03:39.770 information from the trials to

NOTE Confidence: 0.85336065

00:03:39.770 --> 00:03:40.510 make decisions.

NOTE Confidence: 0.85336065

00:03:40.510 --> 00:03:42.722 But the problem which is the focus

NOTE Confidence: 0.85336065

00:03:42.722 --> 00:03:45.240 of my talk is that there very

NOTE Confidence: 0.85336065

00:03:45.240 --> 00:03:47.478 often been taking place in samples

NOTE Confidence: 0.85336065

00:03:47.548 --> 00:03:49.728 that are purely of convenience,

NOTE Confidence: 0.85336065

00:03:49.730 --> 00:03:51.570 which makes thinking about generalizability

NOTE Confidence: 0.85336065

00:03:51.570 --> 00:03:53.042 and heterogeneity rather difficult.

NOTE Confidence: 0.8614285

00:03:57.540 --> 00:03:59.820 If the treatment of X very if treatment

NOTE Confidence: 0.8614285

00:03:59.820 --> 00:04:01.435 effects vary across individuals or

NOTE Confidence: 0.8614285

00:04:01.435 --> 00:04:03.738 they vary across clusters in some way,

NOTE Confidence: 0.8614285

00:04:03.740 --> 00:04:05.210 then it's pretty straightforward to

NOTE Confidence: 0.8614285

00:04:05.210 --> 00:04:07.477 see as a group of statisticians here

NOTE Confidence: 0.8614285

00:04:07.477 --> 00:04:09.312 that the average treatment effect

NOTE Confidence: 0.8614285

00:04:09.312 --> 00:04:11.468 you would get in the population.

NOTE Confidence: 0.8614285

00:04:11.470 --> 00:04:13.678 Is probably not exactly the same thing as

NOTE Confidence: 0.8614285

00:04:13.678 --> 00:04:15.906 the average treatment effect in the sample,

NOTE Confidence: 0.8614285

00:04:15.910 --> 00:04:18.311 and that these could be quite different

NOTE Confidence: 0.8614285

00:04:18.311 --> 00:04:20.522 if treatment effects vary a lot aniff

NOTE Confidence: 0.8614285

00:04:20.522 --> 00:04:22.780 depending upon how the sample is selected.

NOTE Confidence: 0.8614285

00:04:22.780 --> 00:04:25.054 So there has been an increasing

NOTE Confidence: 0.8614285

00:04:25.054 --> 00:04:27.429 amount of work in this area.

NOTE Confidence: 0.8614285

00:04:27.430 --> 00:04:30.190 There's a couple of papers I think that

NOTE Confidence: 0.8614285

00:04:30.190 --> 00:04:32.077 are particularly helpful if there's

NOTE Confidence: 0.8614285

00:04:32.077 --> 00:04:34.333 a paper in education where they're

NOTE Confidence: 0.8614285

00:04:34.333 --> 00:04:36.690 looking at bias from non random treats.

NOTE Confidence: 0.8614285

00:04:36.690 --> 00:04:38.660 Non random treatment assignment or

NOTE Confidence: 0.8614285

00:04:38.660 --> 00:04:41.424 they show that the bias of external

NOTE Confidence: 0.8614285

00:04:41.424 --> 00:04:43.944 validity is on the same order as

NOTE Confidence: 0.8614285

00:04:43.944 --> 00:04:46.171 internal validity bias and so to do

NOTE Confidence: 0.8614285

00:04:46.171 --> 00:04:48.546 so they hear they sort of leverage

NOTE Confidence: 0.8614285

00:04:48.546 --> 00:04:50.506 and natural experiment with a

NOTE Confidence: 0.8614285

00:04:50.506 --> 00:04:52.350 randomized trial to look at this.

NOTE Confidence: 0.8614285

00:04:52.350 --> 00:04:54.130 And that's worked by Bell,

NOTE Confidence: 0.8614285

00:04:54.130 --> 00:04:55.474 Olson, Oregon, Stewart.

NOTE Confidence: 0.8614285

00:04:55.474 --> 00:04:57.714 There's also work showing that.

NOTE Confidence: 0.8614285

00:04:57.720 --> 00:05:00.429 In education and the kinds of schools

NOTE Confidence: 0.8614285

00:05:00.429 --> 00:05:03.205 and school districts that take part in

NOTE Confidence: 0.8614285

00:05:03.205 --> 00:05:05.140 randomized trials are different than

NOTE Confidence: 0.8614285

00:05:05.140 --> 00:05:07.930 the populations of various populations.

NOTE Confidence: 0.8614285

00:05:07.930 --> 00:05:10.186 At something like the Institute of

NOTE Confidence: 0.8614285

00:05:10.186 --> 00:05:12.599 Education Sciences might be interested in,

NOTE Confidence: 0.8614285

00:05:12.600 --> 00:05:16.092 so I have a paper out with Jessica Spy,

NOTE Confidence: 0.8614285

00:05:16.100 --> 00:05:17.975 Brooke Ann are students looking

NOTE Confidence: 0.8614285

00:05:17.975 --> 00:05:20.268 at 37 randomized trials and the

NOTE Confidence: 0.8614285

00:05:20.268 --> 00:05:22.368 samples of schools taking part in

NOTE Confidence: 0.8614285

00:05:22.368 --> 00:05:24.372 those studies and comparing them

NOTE Confidence: 0.8614285

00:05:24.372 --> 00:05:26.597 to various populations of schools.

NOTE Confidence: 0.8614285

00:05:26.600 --> 00:05:28.190 In the US.

NOTE Confidence: 0.8614285

00:05:28.190 --> 00:05:31.370 There's also work hidden behind me.

NOTE Confidence: 0.8614285

00:05:31.370 --> 00:05:32.963 By Liz Stewart.

NOTE Confidence: 0.8614285

00:05:32.963 --> 00:05:36.149 San colleagues looking at school districts

NOTE Confidence: 0.8614285

00:05:36.149 --> 00:05:39.599 and a couple of other papers as well,

NOTE Confidence: 0.8614285

00:05:39.600 --> 00:05:42.498 and these find fairly consistent things.

NOTE Confidence: 0.8614285

00:05:42.500 --> 00:05:43.714 For example,

NOTE Confidence: 0.8614285

00:05:43.714 --> 00:05:46.749 that large school districts are

NOTE Confidence: 0.8614285

00:05:46.749 --> 00:05:48.570 overrepresented in research.

NOTE Confidence: 0.8614285

00:05:48.570 --> 00:05:52.458 Relative to the size of districts in the US.

NOTE Confidence: 0.8614285

00:05:52.460 --> 00:05:54.791 There's been also a lot of work

NOTE Confidence: 0.8614285

00:05:54.791 --> 00:05:57.006 in this area of generalizability

NOTE Confidence: 0.8614285

00:05:57.006 --> 00:05:59.278 and post hoc corrections.

NOTE Confidence: 0.8614285

00:05:59.280 --> 00:06:02.325 I started into this work looking at

NOTE Confidence: 0.8614285

00:06:02.325 --> 00:06:04.995 using post stratification as a way

NOTE Confidence: 0.8614285

00:06:04.995 --> 00:06:07.135 of estimating a population average

NOTE Confidence: 0.8614285

00:06:07.135 --> 00:06:09.069 treatment effect from a sample.

NOTE Confidence: 0.8614285

00:06:09.070 --> 00:06:11.200 There's also been work using

NOTE Confidence: 0.8614285

00:06:11.200 --> 00:06:12.478 inverse probability weighting,

NOTE Confidence: 0.8614285

00:06:12.480 --> 00:06:13.758 maximum entropy weighting

NOTE Confidence: 0.8614285

00:06:13.758 --> 00:06:14.610 bounding approaches.

NOTE Confidence: 0.8614285

00:06:14.610 --> 00:06:16.740 There have been some approaches

NOTE Confidence: 0.8614285

00:06:16.740 --> 00:06:18.444 that focus on little,

NOTE Confidence: 0.8614285

00:06:18.450 --> 00:06:21.638 so I'm thinking like.

NOTE Confidence: 0.8614285

00:06:21.640 --> 00:06:23.796 Here's the paper and Stuart San Green.

NOTE Confidence: 0.8614285

00:06:23.800 --> 00:06:25.340 I think that does that,

NOTE Confidence: 0.8614285

00:06:25.340 --> 00:06:27.090 so there's been like a kind of

NOTE Confidence: 0.8614285

00:06:27.090 --> 00:06:29.029 a flurry of method development.

NOTE Confidence: 0.8614285

00:06:29.030 --> 00:06:31.186 I think here in this area of

NOTE Confidence: 0.8614285

00:06:31.186 --> 00:06:32.110 thinking you know,

NOTE Confidence: 0.8614285

00:06:32.110 --> 00:06:33.958 how do I actually estimate this?

NOTE Confidence: 0.8614285

00:06:33.960 --> 00:06:36.224 If I have population data of different forms

NOTE Confidence: 0.8614285

00:06:36.224 --> 00:06:38.890 and I have sample data of different forms,

NOTE Confidence: 0.8614285

00:06:38.890 --> 00:06:41.302 how can I actually estimate a

NOTE Confidence: 0.8614285

00:06:41.302 --> 00:06:42.910 population average treatment effect?

NOTE Confidence: 0.8614285

00:06:42.910 --> 00:06:45.550 But when I first started doing this work,

NOTE Confidence: 0.8614285

00:06:45.550 --> 00:06:47.384 I realized in a series of examples

NOTE Confidence: 0.8614285

00:06:47.384 --> 00:06:49.601 that I was working on that the

NOTE Confidence: 0.8614285

00:06:49.601 --> 00:06:51.316 effectiveness of these methods is

NOTE Confidence: 0.8614285

00:06:51.316 --> 00:06:53.211 often severely limited in practice

NOTE Confidence: 0.8614285

00:06:53.211 --> 00:06:54.699 because of undercoverage and

NOTE Confidence: 0.8614285

00:06:54.699 --> 00:06:57.430 what I mean is that it you can't.

NOTE Confidence: 0.8614285

00:06:57.430 --> 00:06:59.537 If it turns out that your population

NOTE Confidence: 0.8614285

00:06:59.537 --> 00:07:01.875 has there's a part of the population

NOTE Confidence: 0.8614285

00:07:01.875 --> 00:07:04.360 that's just not represented in the trial,

NOTE Confidence: 0.8614285

00:07:04.360 --> 00:07:06.010 there's really not much statistical

NOTE Confidence: 0.8614285

00:07:06.010 --> 00:07:07.330 magic you can do.

NOTE Confidence: 0.87749577

00:07:07.330 --> 00:07:08.980 You can make some assumptions,

NOTE Confidence: 0.87749577

00:07:08.980 --> 00:07:10.960 but you can't really re wait

NOTE Confidence: 0.87749577

00:07:10.960 --> 00:07:12.940 something that doesn't exist, and the.

NOTE Confidence: 0.8677029

00:07:13.970 --> 00:07:16.175 It's it's really a reflection of lack

NOTE Confidence: 0.8677029

00:07:16.175 --> 00:07:18.270 of positive ITI in the study. Yes,

NOTE Confidence: 0.8677029

00:07:18.270 --> 00:07:19.494 exactly thanks. Yeah exactly.

NOTE Confidence: 0.8677029

00:07:19.494 --> 00:07:21.338 And yeah, I'm just using survey

NOTE Confidence: 0.8677029

00:07:21.338 --> 00:07:23.180 sampling language for the same thing.

NOTE Confidence: 0.8677029

00:07:23.180 --> 00:07:24.720 That's right, yeah, and so.

NOTE Confidence: 0.8677029

00:07:24.720 --> 00:07:26.869 And that's the lack of positive ITI

NOTE Confidence: 0.8677029

00:07:26.869 --> 00:07:28.535 often arises because people aren't

NOTE Confidence: 0.8677029

00:07:28.535 --> 00:07:30.605 thinking about what the population is

NOTE Confidence: 0.8677029

00:07:30.605 --> 00:07:32.638 in advanced and so it's very tricky

NOTE Confidence: 0.8677029

00:07:32.638 --> 00:07:34.550 for them after the fact to generalize,

NOTE Confidence: 0.8677029

00:07:34.550 --> 00:07:36.410 because it turns out that maybe

NOTE Confidence: 0.8677029

00:07:36.410 --> 00:07:38.249 this what I as analyst him now

NOTE Confidence: 0.8677029

00:07:38.249 --> 00:07:40.284 trying to think of as the population

NOTE Confidence: 0.8677029

00:07:40.284 --> 00:07:41.908 isn't exactly the population,

NOTE Confidence: 0.8677029

00:07:41.910 --> 00:07:43.905 but it's very hard for people to

NOTE Confidence: 0.8677029

00:07:43.905 --> 00:07:45.589 articulate what the population is,

NOTE Confidence: 0.8677029

00:07:45.590 --> 00:07:48.810 and so spent a lot of time just trying to.

NOTE Confidence: 0.8677029

00:07:48.810 --> 00:07:50.826 You're out what the population actually is,

NOTE Confidence: 0.8677029

00:07:50.830 --> 00:07:52.546 and if that's population is meaningful,

NOTE Confidence: 0.8677029

00:07:52.550 --> 00:07:54.854 as if that's a population that even matters.

NOTE Confidence: 0.8677029

00:07:54.860 --> 00:07:56.869 So I realized I pivoted a bit.

NOTE Confidence: 0.8677029

00:07:56.870 --> 00:07:58.346 I realized that you could do

NOTE Confidence: 0.8677029

00:07:58.346 --> 00:08:00.330 a lot of this with statistics,

NOTE Confidence: 0.8677029

00:08:00.330 --> 00:08:02.101 but you were going to be limited

NOTE Confidence: 0.8677029

00:08:02.101 --> 00:08:04.070 if you didn't design better trials.

NOTE Confidence: 0.881084

00:08:08.490 --> 00:08:10.394 And so that's allowed me to think.

NOTE Confidence: 0.881084

00:08:10.400 --> 00:08:12.830 Well, why don't we just start?

NOTE Confidence: 0.881084

00:08:12.830 --> 00:08:14.490 The beginning and do this

NOTE Confidence: 0.881084

00:08:14.490 --> 00:08:16.730 do a better job this so why?

NOTE Confidence: 0.881084

00:08:16.730 --> 00:08:18.356 What have we started at the

NOTE Confidence: 0.881084

00:08:18.356 --> 00:08:20.119 beginning of our studies by asking

NOTE Confidence: 0.881084

00:08:20.119 --> 00:08:21.991 what the target population of the

NOTE Confidence: 0.881084

00:08:21.991 --> 00:08:23.510 intervention was thinking about

NOTE Confidence: 0.881084

00:08:23.510 --> 00:08:25.178 inclusion and exclusion criteria,

NOTE Confidence: 0.881084

00:08:25.180 --> 00:08:26.592 I think it helped.

NOTE Confidence: 0.881084

00:08:26.592 --> 00:08:28.710 This probably matters even more with

NOTE Confidence: 0.881084

00:08:28.782 --> 00:08:30.837 like comorbidities and you know.

NOTE Confidence: 0.881084

00:08:30.840 --> 00:08:32.280 I like rolling out people.

NOTE Confidence: 0.881084

00:08:32.280 --> 00:08:33.996 You're doing a study on depression,

NOTE Confidence: 0.881084

00:08:34.000 --> 00:08:36.009 but you rule out people with anxiety,

NOTE Confidence: 0.881084

00:08:36.010 --> 00:08:37.738 and that's like a big problem

NOTE Confidence: 0.881084

00:08:37.738 --> 00:08:39.225 for the interpretation since they

NOTE Confidence: 0.881084

00:08:39.225 --> 00:08:40.887 were highly related to each other.

NOTE Confidence: 0.881084

00:08:40.890 --> 00:08:42.885 This is true in education as well.

NOTE Confidence: 0.881084

00:08:42.890 --> 00:08:43.558 So like,

NOTE Confidence: 0.881084

00:08:43.558 --> 00:08:45.228 what are the inclusion exclusion

NOTE Confidence: 0.881084

00:08:45.228 --> 00:08:47.587 criteria for your trial and how might

NOTE Confidence: 0.881084

00:08:47.587 --> 00:08:49.525 that affect where you can generalize?

NOTE Confidence: 0.881084

00:08:49.530 --> 00:08:51.275 And then also thinking about

NOTE Confidence: 0.881084

00:08:51.275 --> 00:08:52.322 background characteristics and

NOTE Confidence: 0.881084

00:08:52.322 --> 00:08:53.641 contextual variables that might

NOTE Confidence: 0.881084

00:08:53.641 --> 00:08:54.777 moderate the intervention's effect

NOTE Confidence: 0.881084

00:08:54.777 --> 00:08:56.310 and the tricky part here is,

NOTE Confidence: 0.881084

00:08:56.310 --> 00:08:58.529 there's a little bit of a circularity

NOTE Confidence: 0.881084

00:08:58.529 --> 00:09:00.419 which I'm going to keep coming

NOTE Confidence: 0.881084

00:09:00.419 --> 00:09:02.470 back to in what I'm talking about,

NOTE Confidence: 0.881084

00:09:02.470 --> 00:09:04.619 which is in order to know these,

NOTE Confidence: 0.881084

00:09:04.620 --> 00:09:05.236 you know,

NOTE Confidence: 0.881084

00:09:05.236 --> 00:09:07.700 we don't know what these are in advance,

NOTE Confidence: 0.881084

00:09:07.700 --> 00:09:09.660 and we don't have a lot of

NOTE Confidence: 0.881084

00:09:09.660 --> 00:09:11.254 knowledge generated to date about

NOTE Confidence: 0.881084

00:09:11.254 --> 00:09:12.630 what these variables are,

NOTE Confidence: 0.881084

00:09:12.630 --> 00:09:14.492 because studies have not been designed to

NOTE Confidence: 0.881084

00:09:14.492 --> 00:09:16.630 estimate or test hypothesis about moderation,

NOTE Confidence: 0.881084

00:09:16.630 --> 00:09:19.294 and so instead we have to sort of think

NOTE Confidence: 0.881084

00:09:19.294 --> 00:09:21.357 through what we think might matter.

NOTE Confidence: 0.881084

00:09:21.360 --> 00:09:21.707 Using,

NOTE Confidence: 0.881084

00:09:21.707 --> 00:09:22.401 you know,

NOTE Confidence: 0.881084

00:09:22.401 --> 00:09:24.830 not a great source of knowledge here,

NOTE Confidence: 0.881084

00:09:24.830 --> 00:09:26.910 but the idea is that you sort of

NOTE Confidence: 0.881084

00:09:26.910 --> 00:09:28.788 take all of this information and

NOTE Confidence: 0.881084

00:09:28.788 --> 00:09:31.241 then you use this to create to

NOTE Confidence: 0.881084

00:09:31.241 --> 00:09:33.461 use sampling methods to actually

NOTE Confidence: 0.881084

00:09:33.461 --> 00:09:34.793 design recruitment procedures

NOTE Confidence: 0.881084

00:09:34.793 --> 00:09:36.245 like using stratified sampling.

NOTE Confidence: 0.881084

00:09:36.245 --> 00:09:37.920 Figuring out if you should

NOTE Confidence: 0.881084

00:09:37.920 --> 00:09:39.400 you know within Strata,

NOTE Confidence: 0.881084

00:09:39.400 --> 00:09:41.360 using balanced sampling or random

NOTE Confidence: 0.881084

00:09:41.360 --> 00:09:43.677 sampling and thinking about sort of

NOTE Confidence: 0.881084

00:09:43.677 --> 00:09:45.714 ways in which you can increase the

NOTE Confidence: 0.881084

00:09:45.714 --> 00:09:47.740 coverage so you can have positive

NOTE Confidence: 0.881084

00:09:47.740 --> 00:09:49.810 ITI for the whole target population,

NOTE Confidence: 0.881084

00:09:49.810 --> 00:09:52.344 so that when you do you know.

NOTE Confidence: 0.881084

00:09:52.350 --> 00:09:53.981 When you do need to make adjustments

NOTE Confidence: 0.881084

00:09:53.981 --> 00:09:56.030 at the end of your trial using

NOTE Confidence: 0.881084

00:09:56.030 --> 00:09:57.005 these statistical methods,

NOTE Confidence: 0.881084

00:09:57.010 --> 00:09:58.865 they are in a realm in which

NOTE Confidence: 0.881084

00:09:58.865 --> 00:10:00.020 they can perform well.

NOTE Confidence: 0.8562651

00:10:03.030 --> 00:10:05.370 This this sort of lad me to thinking about

NOTE Confidence: 0.8562651

00:10:05.370 --> 00:10:07.368 tools for general for generalization,

NOTE Confidence: 0.8562651

00:10:07.370 --> 00:10:10.090 and so I just want to highlight this

NOTE Confidence: 0.8562651

00:10:10.090 --> 00:10:13.113 because I think this is a good strategy

NOTE Confidence: 0.8562651

00:10:13.113 --> 00:10:15.589 for methods people to think about.

NOTE Confidence: 0.8562651

00:10:15.590 --> 00:10:18.520 So I I thought will nobody is going to do

NOTE Confidence: 0.8562651

00:10:18.592 --> 00:10:21.418 what I'm telling them to do if I don't

NOTE Confidence: 0.8562651

00:10:21.418 --> 00:10:24.238 build a tool because the kind of people.

NOTE Confidence: 0.8562651

00:10:24.240 --> 00:10:25.134 Clan randomized trials,

NOTE Confidence: 0.8562651

00:10:25.134 --> 00:10:26.624 at least in my domain,

NOTE Confidence: 0.8562651

00:10:26.630 --> 00:10:28.466 don't often have statisticians ready at

NOTE Confidence: 0.8562651

00:10:28.466 --> 00:10:30.818 the ready to work with them on things,

NOTE Confidence: 0.8562651

00:10:30.820 --> 00:10:32.536 and they are often writing grant

NOTE Confidence: 0.8562651

00:10:32.536 --> 00:10:34.110 proposals before they've got funding,

NOTE Confidence: 0.8562651

00:10:34.110 --> 00:10:36.670 and so it's very possible that they're not

NOTE Confidence: 0.8562651

00:10:36.670 --> 00:10:39.223 going to think about generalization or or

NOTE Confidence: 0.8562651

00:10:39.223 --> 00:10:41.889 have the training or tools to do it so.

NOTE Confidence: 0.8562651

00:10:41.890 --> 00:10:44.074 I got a grant from the Spencer

NOTE Confidence: 0.8562651

00:10:44.074 --> 00:10:46.479 Foundation and then I've had follow up

NOTE Confidence: 0.8562651

00:10:46.479 --> 00:10:48.579 money from the Institute of Education

NOTE Confidence: 0.8562651

00:10:48.643 --> 00:10:50.815 Sciences to build this tool called

NOTE Confidence: 0.8562651

00:10:50.815 --> 00:10:52.962 the Generalize are that uses some

NOTE Confidence: 0.8562651

00:10:52.962 --> 00:10:54.346 basic design principles standalone.

NOTE Confidence: 0.8562651

00:10:54.350 --> 00:10:55.246 It's got.

NOTE Confidence: 0.8562651

00:10:55.246 --> 00:10:57.934 It's very focused on the user

NOTE Confidence: 0.8562651

00:10:57.934 --> 00:11:00.512 experience and it in the background

NOTE Confidence: 0.8562651

00:11:00.512 --> 00:11:03.740 has the Common Core of data which is.

NOTE Confidence: 0.8562651

00:11:03.740 --> 00:11:05.738 An annual census of the public

NOTE Confidence: 0.8562651

00:11:05.738 --> 00:11:08.473 schools in the US so that the data

NOTE Confidence: 0.8562651

00:11:08.473 --> 00:11:10.750 is already been cleaned and set up.

NOTE Confidence: 0.8562651

00:11:10.750 --> 00:11:13.430 We're adding in right now the iPads data,

NOTE Confidence: 0.8562651

00:11:13.430 --> 00:11:15.563 which is higher Ed data in the US and

NOTE Confidence: 0.8562651

00:11:15.563 --> 00:11:17.905 so the idea is somebody could go in

NOTE Confidence: 0.8562651

00:11:17.905 --> 00:11:19.930 and walk through inclusion exclusion

NOTE Confidence: 0.8562651

00:11:19.930 --> 00:11:22.110 criteria identified moderate orsan.

NOTE Confidence: 0.8562651

00:11:22.110 --> 00:11:23.845 It would build you stratified

NOTE Confidence: 0.8562651

00:11:23.845 --> 00:11:26.120 recruitment plan in less than an hour.

NOTE Confidence: 0.8562651

00:11:26.120 --> 00:11:28.632 You could leave with a list of all

NOTE Confidence: 0.8562651

00:11:28.632 --> 00:11:30.730 the schools and start being able

NOTE Confidence: 0.8562651

00:11:30.730 --> 00:11:32.130 to recruit with the.

NOTE Confidence: 0.83171284

00:11:34.550 --> 00:11:36.657 Great, I've had this going since 2015

NOTE Confidence: 0.83171284

00:11:36.657 --> 00:11:39.320 and it was very slow going for awhile.

NOTE Confidence: 0.83171284

00:11:39.320 --> 00:11:41.896 This is sort of. I just realized this

NOTE Confidence: 0.83171284

00:11:41.896 --> 00:11:43.902 year that I could actually extract

NOTE Confidence: 0.83171284

00:11:43.902 --> 00:11:46.794 a lot of user data and So what you

NOTE Confidence: 0.83171284

00:11:46.794 --> 00:11:49.177 can see here is actually it was slow

NOTE Confidence: 0.83171284

00:11:49.177 --> 00:11:51.396 going and I had some early adopters.

NOTE Confidence: 0.83171284

00:11:51.400 --> 00:11:53.944 These are people that would be star users,

NOTE Confidence: 0.83171284

00:11:53.950 --> 00:11:55.852 so many of them are planning

NOTE Confidence: 0.83171284

00:11:55.852 --> 00:11:56.486 randomized trials,

NOTE Confidence: 0.83171284

00:11:56.490 --> 00:11:58.625 but there was actually a very big

NOTE Confidence: 0.83171284

00:11:58.625 --> 00:12:00.585 jump that occur this summer and

NOTE Confidence: 0.83171284

00:12:00.585 --> 00:12:02.220 that's based on this jump.

NOTE Confidence: 0.83171284

00:12:02.220 --> 00:12:03.880 I actually started digging through

NOTE Confidence: 0.83171284

00:12:03.880 --> 00:12:05.208 things and realized that.

NOTE Confidence: 0.83171284

00:12:05.210 --> 00:12:07.442 Institute of Education Sciences that actually

NOTE Confidence: 0.83171284

00:12:07.442 --> 00:12:08.930 enacted requirements for generalizability.

NOTE Confidence: 0.83171284

00:12:08.930 --> 00:12:10.790 In their request for proposals,

NOTE Confidence: 0.83171284

00:12:10.790 --> 00:12:13.044 and so you can see that what

NOTE Confidence: 0.83171284

00:12:13.044 --> 00:12:14.880 I already always speculated,

NOTE Confidence: 0.83171284

00:12:14.880 --> 00:12:17.484 which is that funders really drive change.

NOTE Confidence: 0.83171284

00:12:17.490 --> 00:12:19.583 So once funders said you need to

NOTE Confidence: 0.83171284

00:12:19.583 --> 00:12:21.580 pay attention to generalizability,

NOTE Confidence: 0.83171284

00:12:21.580 --> 00:12:24.090 people actually started paying attention

NOTE Confidence: 0.83171284

00:12:24.090 --> 00:12:26.600 to generalizability in their proposals.

NOTE Confidence: 0.83171284

00:12:26.600 --> 00:12:28.432 OK, so this I just wanted to give

NOTE Confidence: 0.83171284

00:12:28.432 --> 00:12:30.322 you all of this background as a

NOTE Confidence: 0.83171284

00:12:30.322 --> 00:12:32.232 way of explaining sort of my like

NOTE Confidence: 0.83171284

00:12:32.232 --> 00:12:33.912 where I'm coming from in the in

NOTE Confidence: 0.83171284

00:12:33.912 --> 00:12:35.925 the in heterogeneity and how I'm

NOTE Confidence: 0.83171284

00:12:35.925 --> 00:12:36.984 thinking about this.

NOTE Confidence: 0.83171284

00:12:36.990 --> 00:12:37.391 Um?

NOTE Confidence: 0.83171284

00:12:37.391 --> 00:12:39.396 So everything I've talked about

NOTE Confidence: 0.83171284

00:12:39.396 --> 00:12:41.791 is sort of averaging over hedge

NOTE Confidence: 0.83171284

00:12:41.791 --> 00:12:44.297 and 80 when we talk about analyze.

NOTE Confidence: 0.83171284

00:12:44.300 --> 00:12:45.648 Estimate an average treatment

NOTE Confidence: 0.83171284

00:12:45.648 --> 00:12:46.996 effect for a population,

NOTE Confidence: 0.83171284

00:12:47.000 --> 00:12:49.060 assuming that there's variation of

NOTE Confidence: 0.83171284

00:12:49.060 --> 00:12:51.560 effects and we're averaging over those.

NOTE Confidence: 0.83171284

00:12:51.560 --> 00:12:53.576 But to average over those requires

NOTE Confidence: 0.83171284

00:12:53.576 --> 00:12:55.686 that we know something about how

NOTE Confidence: 0.83171284

00:12:55.686 --> 00:12:57.804 treatment affects very and very often,

NOTE Confidence: 0.83171284

00:12:57.810 --> 00:13:00.085 and I would say this is the

NOTE Confidence: 0.83171284

00:13:00.085 --> 00:13:01.620 in general we don't,

NOTE Confidence: 0.83171284

00:13:01.620 --> 00:13:04.100 and the reason that we don't have a

NOTE Confidence: 0.83171284

00:13:04.100 --> 00:13:06.408 great handle on this is because sample

NOTE Confidence: 0.83171284

00:13:06.408 --> 00:13:09.259 size and sample sizes in randomized trials.

NOTE Confidence: 0.83171284

00:13:09.260 --> 00:13:10.990 I've been very focused on

NOTE Confidence: 0.83171284

00:13:10.990 --> 00:13:12.374 the after treatment effect.

NOTE Confidence: 0.83171284

00:13:12.380 --> 00:13:15.156 Moderators have only become more of a focus,

NOTE Confidence: 0.83171284

00:13:15.160 --> 00:13:16.544 at least in education.

NOTE Confidence: 0.83171284

00:13:16.544 --> 00:13:17.236 More recently,

NOTE Confidence: 0.83171284

00:13:17.240 --> 00:13:20.066 and I think that's true in

NOTE Confidence: 0.83171284

00:13:20.066 --> 00:13:22.730 psychology and related areas as well.

NOTE Confidence: 0.83171284

00:13:22.730 --> 00:13:24.434 And they are often more like

NOTE Confidence: 0.83171284

00:13:24.434 --> 00:13:25.830 exploratory analysis at the end,

NOTE Confidence: 0.83171284

00:13:25.830 --> 00:13:27.685 so you end up with these problems

NOTE Confidence: 0.83171284

00:13:27.685 --> 00:13:29.130 where moderador effects don't get

NOTE Confidence: 0.83171284

00:13:29.130 --> 00:13:30.620 replicated and they don't get

NOTE Confidence: 0.83171284

00:13:30.620 --> 00:13:31.750 replicated because there was.

NOTE Confidence: 0.83171284

00:13:31.750 --> 00:13:32.500 You know,

NOTE Confidence: 0.83171284

00:13:32.500 --> 00:13:34.750 who knows how many statistical tests

NOTE Confidence: 0.83171284

00:13:34.750 --> 00:13:37.538 conducted in order to find those moderators.

NOTE Confidence: 0.83171284

00:13:37.540 --> 00:13:39.110 So they're not very stable,

NOTE Confidence: 0.83171284

00:13:39.110 --> 00:13:40.730 and we don't really necessarily

NOTE Confidence: 0.83171284

00:13:40.730 --> 00:13:42.026 understand or their underpowered

NOTE Confidence: 0.83171284

00:13:42.026 --> 00:13:43.576 deeply underpowered like you just

NOTE Confidence: 0.83171284

00:13:43.576 --> 00:13:45.046 have a very homogeneous sample.

NOTE Confidence: 0.83171284

00:13:45.050 --> 00:13:47.138 And so how are you going to find

NOTE Confidence: 0.83171284

00:13:47.138 --> 00:13:48.814 a treatment effect variation if

NOTE Confidence: 0.83171284

00:13:48.814 --> 00:13:50.986 there's not much variation in your

NOTE Confidence: 0.83171284

00:13:50.986 --> 00:13:52.248 sample to start with,

NOTE Confidence: 0.83171284

00:13:52.250 --> 00:13:53.820 so they're often an afterthought,

NOTE Confidence: 0.83171284

00:13:53.820 --> 00:13:55.857 but I what I noticed overtime is

NOTE Confidence: 0.83171284

00:13:55.857 --> 00:13:57.453 that as generalizability has become

NOTE Confidence: 0.83171284

00:13:57.453 --> 00:13:59.445 something people are paying attention to,

NOTE Confidence: 0.83171284

00:13:59.450 --> 00:14:01.124 people are also starting to pay

NOTE Confidence: 0.83171284

00:14:01.124 --> 00:14:03.018 attention to the idea that you

NOTE Confidence: 0.83171284

00:14:03.018 --> 00:14:04.458 could predict treatment effects,

NOTE Confidence: 0.83171284

00:14:04.460 --> 00:14:06.410 or that you could identify subgroup

NOTE Confidence: 0.83171284

00:14:06.410 --> 00:14:08.151 effects and that this might

NOTE Confidence: 0.83171284

00:14:08.151 --> 00:14:09.659 be very useful information.

NOTE Confidence: 0.83171284

00:14:09.660 --> 00:14:12.145 Which led me to start thinking about

NOTE Confidence: 0.83171284

00:14:12.145 --> 00:14:14.917 how you would design trials for this.

NOTE Confidence: 0.83171284

00:14:14.920 --> 00:14:16.330 So what I'm going to,

NOTE Confidence: 0.83171284

00:14:16.330 --> 00:14:18.129 what I'm leading up to is talking

NOTE Confidence: 0.83171284

00:14:18.129 --> 00:14:19.614 about designing trials to think

NOTE Confidence: 0.83171284

00:14:19.614 --> 00:14:20.280 about heterogeneity.

NOTE Confidence: 0.83171284

00:14:20.280 --> 00:14:22.156 So I'm just going to start with

NOTE Confidence: 0.83171284

00:14:22.156 --> 00:14:22.960 like a little

NOTE Confidence: 0.84072816

00:14:23.021 --> 00:14:24.506 bit of a background here.

NOTE Confidence: 0.84072816

00:14:24.510 --> 00:14:26.758 So we're going to assume that you've got.

NOTE Confidence: 0.84072816

00:14:26.760 --> 00:14:28.170 I'm assuming we've got units

NOTE Confidence: 0.84072816

00:14:28.170 --> 00:14:29.298 which are usually here.

NOTE Confidence: 0.84072816

00:14:29.300 --> 00:14:30.428 Let's say students insights

NOTE Confidence: 0.84072816

00:14:30.428 --> 00:14:31.556 which might be schools,

NOTE Confidence: 0.84072816

00:14:31.560 --> 00:14:33.246 and I'm doing a randomized trial,

NOTE Confidence: 0.84072816

00:14:33.250 --> 00:14:36.460 and I've got these potential outcomes.

NOTE Confidence: 0.84072816

00:14:36.460 --> 00:14:38.674 And so we've got both an

NOTE Confidence: 0.84072816

00:14:38.674 --> 00:14:40.680 average and intercept in these,

NOTE Confidence: 0.84072816

00:14:40.680 --> 00:14:43.354 and we've also got some sort of

NOTE Confidence: 0.84072816

00:14:43.354 --> 00:14:45.678 fixed variation that we can explain.

NOTE Confidence: 0.84072816

00:14:45.680 --> 00:14:48.384 And then we have this other parts that

NOTE Confidence: 0.84072816

00:14:48.384 --> 00:14:51.049 are not affected by the treatment.

NOTE Confidence: 0.84072816

00:14:51.050 --> 00:14:53.696 We've got some site level and individual

NOTE Confidence: 0.84072816

00:14:53.696 --> 00:14:56.250 residuals and some idiosyncratic errors.

NOTE Confidence: 0.84072816

00:14:56.250 --> 00:14:57.965 But what we're interested in

NOTE Confidence: 0.84072816

00:14:57.965 --> 00:15:00.188 really is in these these moderate

NOTE Confidence: 0.84072816

00:15:00.188 --> 00:15:01.948 yrs of treatment effects,

NOTE Confidence: 0.84072816

00:15:01.950 --> 00:15:04.106 and so you could say that Delta

NOTE Confidence: 0.84072816

00:15:04.106 --> 00:15:06.888 0 is the difference in averages.

NOTE Confidence: 0.84072816

00:15:06.890 --> 00:15:09.170 I'm assuming these are centered variables,

NOTE Confidence: 0.84072816

00:15:09.170 --> 00:15:11.557 so this is nicely the difference in

NOTE Confidence: 0.84072816

00:15:11.557 --> 00:15:13.790 averages and that the vector Delta

NOTE Confidence: 0.84072816

00:15:13.790 --> 00:15:15.690 is the difference between these

NOTE Confidence: 0.84072816

00:15:15.690 --> 00:15:18.149 effects of the treatment and then

NOTE Confidence: 0.84072816

00:15:18.149 --> 00:15:20.189 under treatment and under control.

NOTE Confidence: 0.82586044

00:15:25.520 --> 00:15:27.432 A lot of so as you have to

NOTE Confidence: 0.82586044

00:15:27.432 --> 00:15:28.341 think about interpretability

NOTE Confidence: 0.82586044

00:15:28.341 --> 00:15:30.875 here of what I mean by Delta.

NOTE Confidence: 0.82586044

00:15:30.880 --> 00:15:32.698 By this by these deltas and

NOTE Confidence: 0.82586044

00:15:32.698 --> 00:15:34.246 how to standardize because we

NOTE Confidence: 0.82586044

00:15:34.246 --> 00:15:35.646 wanted to talk about these,

NOTE Confidence: 0.82586044

00:15:35.650 --> 00:15:38.034 they need to have a mean of 0,

NOTE Confidence: 0.82586044

00:15:38.040 --> 00:15:39.822 but also in order to talk

NOTE Confidence: 0.82586044

00:15:39.822 --> 00:15:40.713 about treatment effects.

NOTE Confidence: 0.82586044

00:15:40.720 --> 00:15:42.508 Sort of done in general for

NOTE Confidence: 0.82586044

00:15:42.508 --> 00:15:43.700 developing things like power,

NOTE Confidence: 0.82586044

00:15:43.700 --> 00:15:45.080 we often standardize them so

NOTE Confidence: 0.82586044

00:15:45.080 --> 00:15:46.844 often we have effect sizes for

NOTE Confidence: 0.82586044

00:15:46.844 --> 00:15:48.168 the average treatment effect,

NOTE Confidence: 0.82586044

00:15:48.170 --> 00:15:49.422 their standardized in relation

NOTE Confidence: 0.82586044

00:15:49.422 --> 00:15:51.690 to the variation in the in the

NOTE Confidence: 0.82586044

00:15:51.690 --> 00:15:53.010 sample and the population.

NOTE Confidence: 0.82586044

00:15:53.010 --> 00:15:55.034 And so here I'm going to sort of

NOTE Confidence: 0.82586044

00:15:55.034 --> 00:15:57.612 say we what we need to do is we

NOTE Confidence: 0.82586044

00:15:57.612 --> 00:15:59.301 need to standardize the covariates

NOTE Confidence: 0.82586044

00:15:59.301 --> 00:16:01.599 and we need to standardize the

NOTE Confidence: 0.82586044

00:16:01.599 --> 00:16:03.303 covariates in relation to the

NOTE Confidence: 0.82586044

00:16:03.303 --> 00:16:04.236 population standard deviation.

NOTE Confidence: 0.82586044

00:16:04.240 --> 00:16:05.902 This might not seem like this

NOTE Confidence: 0.82586044

00:16:05.902 --> 00:16:07.670 is like a radical statement,

NOTE Confidence: 0.82586044

00:16:07.670 --> 00:16:10.014 but if you look into the power analysis

NOTE Confidence: 0.82586044

00:16:10.014 --> 00:16:11.967 literature on how to conduct power

NOTE Confidence: 0.82586044

00:16:11.967 --> 00:16:13.597 analysis for moderate are tests,

NOTE Confidence: 0.82586044

00:16:13.600 --> 00:16:15.070 they are typically standardizing in

NOTE Confidence: 0.82586044

00:16:15.070 --> 00:16:17.030 relation to the sample standard deviation,

NOTE Confidence: 0.82586044

00:16:17.030 --> 00:16:18.166 and in doing so,

NOTE Confidence: 0.82586044

00:16:18.166 --> 00:16:19.870 it makes it impossible to see

NOTE Confidence: 0.82586044

00:16:19.934 --> 00:16:22.046 how your sample actually how you

NOTE Confidence: 0.82586044

00:16:22.046 --> 00:16:23.950 choose your sample might matter.

NOTE Confidence: 0.82586044

00:16:23.950 --> 00:16:25.234 Isibaya standardizing by this

NOTE Confidence: 0.82586044

00:16:25.234 --> 00:16:26.839 fixed value by the population,

NOTE Confidence: 0.82586044

00:16:26.840 --> 00:16:27.980 you've identified a population,

NOTE Confidence: 0.82586044

00:16:27.980 --> 00:16:29.405 and now we're standardizing by

NOTE Confidence: 0.82586044

00:16:29.405 --> 00:16:31.008 that population standard deviation.

NOTE Confidence: 0.82586044

00:16:31.010 --> 00:16:34.097 That will make the role that the

NOTE Confidence: 0.82586044

00:16:34.097 --> 00:16:36.929 sample plays here much more clear.

NOTE Confidence: 0.82586044

00:16:36.930 --> 00:16:37.259 OK,

NOTE Confidence: 0.82586044

00:16:37.259 --> 00:16:39.233 so the fact that we randomized

NOTE Confidence: 0.82586044

00:16:39.233 --> 00:16:41.082 to treatment and control allows

NOTE Confidence: 0.82586044

00:16:41.082 --> 00:16:43.446 us to estimate these dealt these

NOTE Confidence: 0.82586044

00:16:43.446 --> 00:16:45.517 Spectre Delta using some generalized

NOTE Confidence: 0.82586044

00:16:45.517 --> 00:16:47.537 least squares of some sort,

NOTE Confidence: 0.82586044

00:16:47.540 --> 00:16:49.465 and I'm being a little big here

NOTE Confidence: 0.82586044

00:16:49.465 --> 00:16:51.347 because I'm trying to encapsulate

NOTE Confidence: 0.82586044

00:16:51.347 --> 00:16:52.850 cluster randomized randomized

NOTE Confidence: 0.82586044

00:16:52.850 --> 00:16:54.353 block individual randomizer,

NOTE Confidence: 0.82586044

00:16:54.360 --> 00:16:56.260 all like versions of this.

NOTE Confidence: 0.82586044

00:16:56.260 --> 00:16:59.667 OK, so I can do so, I can separate.

NOTE Confidence: 0.82586044

00:16:59.667 --> 00:17:02.320 These are at additive or rather subtractive.

NOTE Confidence: 0.82586044

00:17:02.320 --> 00:17:04.210 I guess the treatment and

NOTE Confidence: 0.82586044

00:17:04.210 --> 00:17:05.722 the control their step.

NOTE Confidence: 0.82586044

00:17:05.730 --> 00:17:07.282 You can separate them.

NOTE Confidence: 0.82586044

00:17:07.282 --> 00:17:09.610 And and through this I can

NOTE Confidence: 0.82586044

00:17:09.693 --> 00:17:12.069 think about statistical power,

NOTE Confidence: 0.82586044

00:17:12.070 --> 00:17:15.857 and for each of these moderador effects.

NOTE Confidence: 0.82586044

00:17:15.860 --> 00:17:16.234 And so,

NOTE Confidence: 0.82586044

00:17:16.234 --> 00:17:18.125 one way you can do that is through the

NOTE Confidence: 0.82586044

00:17:18.125 --> 00:17:19.835 minimum detectable effect size difference.

NOTE Confidence: 0.82586044

00:17:19.840 --> 00:17:21.622 I don't know how common this

NOTE Confidence: 0.82586044

00:17:21.622 --> 00:17:23.440 is used in the sort of.

NOTE Confidence: 0.82586044

00:17:23.440 --> 00:17:24.268 Biostats world,

NOTE Confidence: 0.82586044

00:17:24.268 --> 00:17:27.166 but it's a pretty common metric that's

NOTE Confidence: 0.82586044

00:17:27.166 --> 00:17:29.910 used in cluster randomized trials in.

NOTE Confidence: 0.82586044

00:17:29.910 --> 00:17:31.870 The world I work in,

NOTE Confidence: 0.82586044

00:17:31.870 --> 00:17:34.222 and so it's nice because it's

NOTE Confidence: 0.82586044

00:17:34.222 --> 00:17:35.790 sort of easily interpretable,

NOTE Confidence: 0.82586044

00:17:35.790 --> 00:17:38.527 so this is the smallest affect size

NOTE Confidence: 0.82586044

00:17:38.527 --> 00:17:41.667 that you could for a for a given

NOTE Confidence: 0.82586044

00:17:41.667 --> 00:17:44.020 Alpha level which is affecting this.

NOTE Confidence: 0.82586044

00:17:44.020 --> 00:17:47.328 Msub knew this is.

NOTE Confidence: 0.82586044

00:17:47.330 --> 00:17:49.020 That's like the critical value.

NOTE Confidence: 0.82586044

00:17:49.020 --> 00:17:51.006 This is sort of the smallest

NOTE Confidence: 0.82586044

00:17:51.006 --> 00:17:53.076 true effect that you could detect

NOTE Confidence: 0.82586044

00:17:53.076 --> 00:17:55.442 with the power that you with like

NOTE Confidence: 0.82586044

00:17:55.442 --> 00:17:58.168 80% power for example.

NOTE Confidence: 0.82586044

00:17:58.170 --> 00:18:01.410 And so this is like a general form for this,

NOTE Confidence: 0.82586044

00:18:01.410 --> 00:18:04.326 and So what I'm showing is that its function,

NOTE Confidence: 0.82586044

00:18:04.330 --> 00:18:06.268 can I like move my hands?

NOTE Confidence: 0.82586044

00:18:06.270 --> 00:18:07.215 I don't know.

NOTE Confidence: 0.82586044

00:18:07.215 --> 00:18:09.105 I'm just going to involve a

NOTE Confidence: 0.82586044

00:18:09.105 --> 00:18:10.479 lot of never mind,

NOTE Confidence: 0.82586044

00:18:10.480 --> 00:18:12.643 so it's a function of the variation

NOTE Confidence: 0.82586044

00:18:12.643 --> 00:18:14.688 in the population in that covariate.

NOTE Confidence: 0.82586044

00:18:14.690 --> 00:18:16.640 It's also a function of S,

NOTE Confidence: 0.82586044

00:18:16.640 --> 00:18:18.712 which is you could think of as

NOTE Confidence: 0.82586044

00:18:18.712 --> 00:18:20.370 the sort of covariance matrix

NOTE Confidence: 0.82586044

00:18:20.370 --> 00:18:22.785 of the X is in the sample,

NOTE Confidence: 0.8670643

00:18:22.790 --> 00:18:24.138 so those are different.

NOTE Confidence: 0.8670643

00:18:24.138 --> 00:18:26.360 And then it's a function of N,

NOTE Confidence: 0.8670643

00:18:26.360 --> 00:18:28.628 which is the sample size per cluster.

NOTE Confidence: 0.8670643

00:18:28.630 --> 00:18:30.430 I'm assuming it's constant here.

NOTE Confidence: 0.8670643

00:18:30.430 --> 00:18:32.870 J is the number of clusters and P

NOTE Confidence: 0.8670643

00:18:32.870 --> 00:18:35.740 is the proportion in treatment. So

NOTE Confidence: 0.80064327

00:18:35.740 --> 00:18:38.154 that what is Sigma XK squared? Is?

NOTE Confidence: 0.80064327

00:18:38.154 --> 00:18:40.352 The population SD of effect modifier or

NOTE Confidence: 0.80064327

00:18:40.352 --> 00:18:42.389 the population variance effect modifier?

NOTE Confidence: 0.80064327

00:18:42.390 --> 00:18:43.790 Is the population variance

NOTE Confidence: 0.80064327

00:18:43.790 --> 00:18:45.890 of the effect moderate or modifier?

NOTE Confidence: 0.80064327

00:18:45.890 --> 00:18:48.200 But then your square rooting it so

NOTE Confidence: 0.80064327

00:18:48.200 --> 00:18:50.438 it's going to be gradual scale.

NOTE Confidence: 0.8265885

00:18:52.670 --> 00:18:55.078 OK, so just to give you a couple

NOTE Confidence: 0.8265885

00:18:55.078 --> 00:18:57.150 of special cases where you can

NOTE Confidence: 0.8265885

00:18:57.150 --> 00:18:59.250 sort of parse out some things.

NOTE Confidence: 0.8265885

00:18:59.250 --> 00:19:00.900 So there's been previous work.

NOTE Confidence: 0.8265885

00:19:00.900 --> 00:19:03.532 I meant to include a citation here by

NOTE Confidence: 0.8265885

00:19:03.532 --> 00:19:05.162 Jessica Spy, Brooke and colleagues.

NOTE Confidence: 0.8265885

00:19:05.162 --> 00:19:06.458 That's looking at power

NOTE Confidence: 0.8265885

00:19:06.458 --> 00:19:07.800 for moderate are tests.

NOTE Confidence: 0.8265885

00:19:07.800 --> 00:19:10.110 And so here's 2 cases we have.

NOTE Confidence: 0.8265885

00:19:10.110 --> 00:19:11.882 Site lab, site level,

NOTE Confidence: 0.8265885

00:19:11.882 --> 00:19:13.654 moderate yrs and individual

NOTE Confidence: 0.8265885

00:19:13.654 --> 00:19:15.169 level moderate yrs and.

NOTE Confidence: 0.8265885

00:19:15.170 --> 00:19:17.319 I'm I'm taking basically what they've got,

NOTE Confidence: 0.8265885

00:19:17.320 --> 00:19:19.156 but re tweaking part of it.

NOTE Confidence: 0.851218459999999

00:19:21.190 --> 00:19:23.542 Because I'm factoring out this Sigma

NOTE Confidence: 0.851218459999999

00:19:23.542 --> 00:19:25.837 squared and noting that you can

NOTE Confidence: 0.851218459999999

00:19:25.837 --> 00:19:27.799 actually pull out this thing called

NOTE Confidence: 0.851218459999999

00:19:27.799 --> 00:19:30.165 RXK at the front and the RXK is

NOTE Confidence: 0.851218459999999

00:19:30.165 --> 00:19:31.918 this ratio of the standard deviation

NOTE Confidence: 0.851218459999999

00:19:31.918 --> 00:19:33.862 of the covariate in the sample

NOTE Confidence: 0.851218459999999

00:19:33.862 --> 00:19:35.592 compared to the standard deviation

NOTE Confidence: 0.851218459999999

00:19:35.592 --> 00:19:37.644 of the covariate in the population.

NOTE Confidence: 0.851218459999999

00:19:37.650 --> 00:19:40.395 And So what you can see here is by

NOTE Confidence: 0.851218459999999

00:19:40.395 --> 00:19:42.686 doing that you by rewriting it.

NOTE Confidence: 0.851218459999999

00:19:42.690 --> 00:19:45.426 This way you can see that our XK is

NOTE Confidence: 0.851218459999999

00:19:45.426 --> 00:19:48.293 having just as much of an effect on

NOTE Confidence: 0.851218459999999

00:19:48.293 --> 00:19:50.433 statistical power as things like the

NOTE Confidence: 0.851218459999999

00:19:50.433 --> 00:19:53.682 square root of N or the square root of P.

NOTE Confidence: 0.851218459999999

00:19:53.682 --> 00:19:55.506 These other parameters that most power

NOTE Confidence: 0.851218459999999

00:19:55.506 --> 00:19:57.178 analysis has spent has focused on,

NOTE Confidence: 0.851218459999999

00:19:57.180 --> 00:19:58.008 and that's true.

NOTE Confidence: 0.851218459999999

00:19:58.008 --> 00:20:00.260 You know, in any of these designs.

NOTE Confidence: 0.851218459999999

00:20:00.260 --> 00:20:02.836 Love seeing it in any of these designs.

NOTE Confidence: 0.851218459999999

00:20:02.840 --> 00:20:06.114 RX shows up. OK,

NOTE Confidence: 0.851218459999999

00:20:06.114 --> 00:20:10.020 so if RX is something that matters for power,

NOTE Confidence: 0.851218459999999

00:20:10.020 --> 00:20:12.190 a question will be well.

NOTE Confidence: 0.851218459999999

00:20:12.190 --> 00:20:14.360 What are people doing in

NOTE Confidence: 0.851218459999999

00:20:14.360 --> 00:20:16.096 practice right now right?

NOTE Confidence: 0.851218459999999

00:20:16.100 --> 00:20:18.698 So maybe maybe people are choosing

NOTE Confidence: 0.851218459999999

00:20:18.698 --> 00:20:19.997 fairly heterogeneous samples,

NOTE Confidence: 0.851218459999999

00:20:20.000 --> 00:20:23.464 and So what I've got here is 19.

NOTE Confidence: 0.851218459999999

00:20:23.470 --> 00:20:25.860 This is 19 randomized trials

NOTE Confidence: 0.851218459999999

00:20:25.860 --> 00:20:28.250 in education that we extracted

NOTE Confidence: 0.851218459999999

00:20:28.336 --> 00:20:30.846 information from and we've got.

NOTE Confidence: 0.851218459999999

00:20:30.850 --> 00:20:32.962 So these are box plots of

NOTE Confidence: 0.851218459999999

00:20:32.962 --> 00:20:35.120 values across each of these 19,

NOTE Confidence: 0.851218459999999

00:20:35.120 --> 00:20:37.880 and for each of them I've calculated for

NOTE Confidence: 0.851218459999999

00:20:37.880 --> 00:20:40.457 holding like the US population of school.

NOTE Confidence: 0.851218459999999

00:20:40.460 --> 00:20:42.539 So this is like the US population

NOTE Confidence: 0.851218459999999

00:20:42.539 --> 00:20:44.729 of let's say elementary schools.

NOTE Confidence: 0.851218459999999

00:20:44.730 --> 00:20:47.019 I'm looking at the ratio of this

NOTE Confidence: 0.851218459999999

00:20:47.019 --> 00:20:49.384 moderate are in the sample in these

NOTE Confidence: 0.851218459999999

00:20:49.384 --> 00:20:52.240 studies to the ratio of that to that

NOTE Confidence: 0.851218459999999

00:20:52.240 --> 00:20:54.694 standard deviation in the population OK,

NOTE Confidence: 0.851218459999999

00:20:54.700 --> 00:20:57.304 and then I'm looking at boxplots of

NOTE Confidence: 0.851218459999999

00:20:57.304 --> 00:21:00.396 this and what you can see like do this,

NOTE Confidence: 0.851218459999999

00:21:00.400 --> 00:21:00.737 don't?

NOTE Confidence: 0.851218459999999

00:21:00.737 --> 00:21:03.433 OK, what you can see here is that

NOTE Confidence: 0.851218459999999

00:21:03.433 --> 00:21:05.230 the bar at the bottom.

NOTE Confidence: 0.851218459999999

00:21:05.230 --> 00:21:08.020 Can you see my cursor?

NOTE Confidence: 0.851218459999999

00:21:08.020 --> 00:21:10.369 Can't tell if you guys can see my curse.

NOTE Confidence: 0.851218459999999

00:21:10.370 --> 00:21:10.631 No,

NOTE Confidence: 0.851218459999999

00:21:10.631 --> 00:21:11.936 you can't see my cursor.

NOTE Confidence: 0.851218459999999

00:21:11.940 --> 00:21:12.182 OK,

NOTE Confidence: 0.851218459999999

00:21:12.182 --> 00:21:14.360 so the bar at the bottom there's an R

NOTE Confidence: 0.851218459999999

00:21:14.425 --> 00:21:16.711 X = sqrt 1/2 and then there's a line

NOTE Confidence: 0.851218459999999

00:21:16.711 --> 00:21:18.974 across the top that's like a dashed one.

NOTE Confidence: 0.851218459999999

00:21:18.980 --> 00:21:21.570 That's the R X = sqrt 2.

NOTE Confidence: 0.851218459999999

00:21:21.570 --> 00:21:21.885 OK,

NOTE Confidence: 0.851218459999999

00:21:21.885 --> 00:21:24.090 and so you can see that most

NOTE Confidence: 0.851218459999999

00:21:24.090 --> 00:21:26.298 studies are actually below there,

NOTE Confidence: 0.851218459999999

00:21:26.300 --> 00:21:28.490 less heterogeneous than the population there.

NOTE Confidence: 0.851218459999999

00:21:28.490 --> 00:21:30.310 Below this line for one,

NOTE Confidence: 0.851218459999999

00:21:30.310 --> 00:21:32.392 and they're actually far less heterogeneous

NOTE Confidence: 0.851218459999999

00:21:32.392 --> 00:21:35.040 than the than the population there are.

NOTE Confidence: 0.851218459999999

00:21:35.040 --> 00:21:35.403 Actually.

NOTE Confidence: 0.851218459999999

00:21:35.403 --> 00:21:37.944 If you look at these median values,

NOTE Confidence: 0.851218459999999

00:21:37.950 --> 00:21:40.128 many of them are closed 2.5,

NOTE Confidence: 0.851218459999999

00:21:40.130 --> 00:21:42.754 so they are about 1/4 of the variation

NOTE Confidence: 0.851218459999999

00:21:42.754 --> 00:21:45.227 as we're seeing in the population.

NOTE Confidence: 0.851218459999999

00:21:45.230 --> 00:21:49.054 So this gives you a sense that if.

NOTE Confidence: 0.851218459999999

00:21:49.060 --> 00:21:50.164 That there's, uh,

NOTE Confidence: 0.851218459999999

00:21:50.164 --> 00:21:52.040 an opportunity to improve, right?

NOTE Confidence: 0.851218459999999

00:21:52.040 --> 00:21:55.240 Like I could increase power not just by

NOTE Confidence: 0.851218459999999

00:21:55.240 --> 00:21:57.748 increasing my sample size or increasing.

NOTE Confidence: 0.851218459999999

00:21:57.750 --> 00:21:59.160 My sample size in schools or

NOTE Confidence: 0.851218459999999

00:21:59.160 --> 00:22:01.093 my sample size of the number of

NOTE Confidence: 0.851218459999999

00:22:01.093 --> 00:22:02.658 schools which are pretty expensive,

NOTE Confidence: 0.851218459999999

00:22:02.660 --> 00:22:04.788 but I could also increase my power by

NOTE Confidence: 0.851218459999999

00:22:04.788 --> 00:22:07.028 changing the kinds of samples that I select.

NOTE Confidence: 0.8691194

00:22:09.540 --> 00:22:12.156 And so that's where these numbers came from.

NOTE Confidence: 0.8691194

00:22:12.160 --> 00:22:13.790 They should have gone to

NOTE Confidence: 0.8691194

00:22:13.790 --> 00:22:14.768 slightly different order.

NOTE Confidence: 0.8691194

00:22:14.770 --> 00:22:17.059 So the main point is that design

NOTE Confidence: 0.8691194

00:22:17.059 --> 00:22:18.774 sensitivity, the way we think,

NOTE Confidence: 0.8691194

00:22:18.774 --> 00:22:20.504 whether that statistical power or

NOTE Confidence: 0.8691194

00:22:20.504 --> 00:22:22.362 standard errors or whatever framework

NOTE Confidence: 0.8691194

00:22:22.362 --> 00:22:24.540 that there this is proportional in

NOTE Confidence: 0.8691194

00:22:24.603 --> 00:22:26.934 some way to this RX value that we can

NOTE Confidence: 0.8691194

00:22:26.934 --> 00:22:29.284 improve our design sensitivity by

NOTE Confidence: 0.8691194

00:22:29.284 --> 00:22:31.889 choosing a more heterogeneous sample.

NOTE Confidence: 0.8691194

00:22:31.890 --> 00:22:34.042 And so funny, I must have like put

NOTE Confidence: 0.8691194

00:22:34.042 --> 00:22:36.128 this in here twice on accident.

NOTE Confidence: 0.8691194

00:22:36.130 --> 00:22:38.097 So this is the same thing but

NOTE Confidence: 0.8691194

00:22:38.097 --> 00:22:39.769 with a line through it.

NOTE Confidence: 0.8691194

00:22:39.770 --> 00:22:42.050 OK so if once you have that insight

NOTE Confidence: 0.8691194

00:22:42.050 --> 00:22:43.100 that heterogeneity matters,

NOTE Confidence: 0.8691194

00:22:43.100 --> 00:22:44.505 that it's actually something that

NOTE Confidence: 0.8691194

00:22:44.505 --> 00:22:46.554 we can include in our power analysis

NOTE Confidence: 0.8691194

00:22:46.554 --> 00:22:48.598 and that is something that is not

NOTE Confidence: 0.8691194

00:22:48.598 --> 00:22:50.070 actually happening in practice.

NOTE Confidence: 0.8691194

00:22:50.070 --> 00:22:52.212 Then we can start thinking about how

NOTE Confidence: 0.8691194

00:22:52.212 --> 00:22:54.419 we might plan studies differently.

NOTE Confidence: 0.8691194

00:22:54.420 --> 00:22:55.518 OK, so if.

NOTE Confidence: 0.8691194

00:22:55.518 --> 00:22:58.080 So how can we improve statistical power?

NOTE Confidence: 0.8691194

00:22:58.080 --> 00:22:58.379 Well,

NOTE Confidence: 0.8691194

00:22:58.379 --> 00:23:01.070 a lot of the literature as I was saying,

NOTE Confidence: 0.8691194

00:23:01.070 --> 00:23:02.774 is focused on improving power by

NOTE Confidence: 0.8691194

00:23:02.774 --> 00:23:04.360 increasing sample size or instead.

NOTE Confidence: 0.8691194

00:23:04.360 --> 00:23:06.600 But what I'm arguing here is that you

NOTE Confidence: 0.8691194

00:23:06.600 --> 00:23:08.249 could increase instead this ratio.

NOTE Confidence: 0.8691194

00:23:08.250 --> 00:23:09.575 You could increase the variation

NOTE Confidence: 0.8691194

00:23:09.575 --> 00:23:10.900 in your sample choosing more

NOTE Confidence: 0.8691194

00:23:10.950 --> 00:23:12.114 heterogeneous sample annual have

NOTE Confidence: 0.8691194

00:23:12.114 --> 00:23:13.569 more statistical power for test

NOTE Confidence: 0.8691194

00:23:13.569 --> 00:23:15.129 of heterogeneity of moderators,

NOTE Confidence: 0.8691194

00:23:15.130 --> 00:23:17.514 and So what would you do with this?

NOTE Confidence: 0.8691194

00:23:17.520 --> 00:23:19.266 It would mean you know purposefully

NOTE Confidence: 0.8691194

00:23:19.266 --> 00:23:21.109 choosing sites that were more extreme,

NOTE Confidence: 0.8691194

00:23:21.110 --> 00:23:21.929 it might end,

NOTE Confidence: 0.8691194

00:23:21.929 --> 00:23:24.689 and that's easy enough to do in one variable.

NOTE Confidence: 0.8691194

00:23:24.690 --> 00:23:27.876 And I'm going to talk a little bit about

NOTE Confidence: 0.8691194

00:23:27.876 --> 00:23:31.138 how to do that with multiple variables.

NOTE Confidence: 0.8691194

00:23:31.140 --> 00:23:32.216 So with a simple,

NOTE Confidence: 0.8691194

00:23:32.216 --> 00:23:35.029 let's just say we had one single continuous.

NOTE Confidence: 0.8691194

00:23:35.030 --> 00:23:37.298 Moderate are like this is a normal

NOTE Confidence: 0.8691194

00:23:37.298 --> 00:23:38.270 distance normally distributed.

NOTE Confidence: 0.8691194

00:23:38.270 --> 00:23:40.552 This theory would tell us that we

NOTE Confidence: 0.8691194

00:23:40.552 --> 00:23:42.479 should choose half of our sample.

NOTE Confidence: 0.8691194

00:23:42.480 --> 00:23:44.615 We would choose half of our sample

NOTE Confidence: 0.8691194

00:23:44.615 --> 00:23:46.891 from the upper from the upper an

NOTE Confidence: 0.8691194

00:23:46.891 --> 00:23:48.835 lower tails and choosing them from

NOTE Confidence: 0.8691194

00:23:48.898 --> 00:23:50.878 the upper and lower tails were

NOTE Confidence: 0.8691194

00:23:50.878 --> 00:23:53.172 actually getting an RX of sqrt 2.

NOTE Confidence: 0.8691194

00:23:53.172 --> 00:23:55.116 This is actually a rather large,

NOTE Confidence: 0.8691194

00:23:55.120 --> 00:23:57.272 so this is going to create a much

NOTE Confidence: 0.8691194

00:23:57.272 --> 00:23:59.000 more homogeneous heterogeneous sample,

NOTE Confidence: 0.8691194

00:23:59.000 --> 00:24:00.810 thus increasing our statistical power

NOTE Confidence: 0.8691194

00:24:00.810 --> 00:24:02.870 because it's more heterogeneous than the.

NOTE Confidence: 0.8691194

00:24:02.870 --> 00:24:03.830 In the population.

NOTE Confidence: 0.8697402

00:24:07.120 --> 00:24:08.745 Similarly, if we had two

NOTE Confidence: 0.8697402

00:24:08.745 --> 00:24:09.720 correlated normal variables,

NOTE Confidence: 0.8697402

00:24:09.720 --> 00:24:11.670 when we this is, you know,

NOTE Confidence: 0.8697402

00:24:11.670 --> 00:24:14.270 we could imagine getting the corners of this.

NOTE Confidence: 0.8697402

00:24:14.270 --> 00:24:16.545 These are all principles, by the way,

NOTE Confidence: 0.8697402

00:24:16.550 --> 00:24:18.170 straight up from experimental design.

NOTE Confidence: 0.8697402

00:24:18.170 --> 00:24:20.770 If you think if you think about it,

NOTE Confidence: 0.8697402

00:24:20.770 --> 00:24:22.720 there are principles from like 2.

NOTE Confidence: 0.8697402

00:24:22.720 --> 00:24:24.592 You know two factor studies or

NOTE Confidence: 0.8697402

00:24:24.592 --> 00:24:26.283 multi factor studies where you're

NOTE Confidence: 0.8697402

00:24:26.283 --> 00:24:28.611 manipulating and instead I'm just saying

NOTE Confidence: 0.8697402

00:24:28.611 --> 00:24:30.571 instead of manipulating these factors

NOTE Confidence: 0.8697402

00:24:30.571 --> 00:24:32.406 were now measuring these factors.

NOTE Confidence: 0.8697402

00:24:32.410 --> 00:24:33.980 Someplace you could choose them

NOTE Confidence: 0.8697402

00:24:33.980 --> 00:24:35.550 to be extreme design points.

NOTE Confidence: 0.8697402

00:24:35.550 --> 00:24:37.434 It gets a little harder once

NOTE Confidence: 0.8697402

00:24:37.434 --> 00:24:38.376 things become correlated,

NOTE Confidence: 0.8697402

00:24:38.380 --> 00:24:39.950 so when they become correlated,

NOTE Confidence: 0.8697402

00:24:39.950 --> 00:24:41.868 I don't have as much sample available

NOTE Confidence: 0.8697402

00:24:41.868 --> 00:24:43.816 to me because there's just fewer

NOTE Confidence: 0.8697402

00:24:43.816 --> 00:24:45.596 population units in those corners,

NOTE Confidence: 0.8697402

00:24:45.600 --> 00:24:47.340 and so it's going to become

NOTE Confidence: 0.8697402

00:24:47.340 --> 00:24:49.370 increasingly hard as I add variables,

NOTE Confidence: 0.8697402

00:24:49.370 --> 00:24:51.589 it might become harder and harder in

NOTE Confidence: 0.8697402

00:24:51.589 --> 00:24:53.800 order to figure out what these units

NOTE Confidence: 0.8697402

00:24:53.800 --> 00:24:55.960 are that I could be sampling from.

NOTE Confidence: 0.88043493

00:25:01.780 --> 00:25:03.660 So I started thinking about

NOTE Confidence: 0.88043493

00:25:03.660 --> 00:25:05.540 how you would do this,

NOTE Confidence: 0.88043493

00:25:05.540 --> 00:25:07.899 and I realized that there is actually

NOTE Confidence: 0.88043493

00:25:07.899 --> 00:25:10.690 a literature on this in in the world

NOTE Confidence: 0.88043493

00:25:10.690 --> 00:25:12.405 of sort of industrial experiments

NOTE Confidence: 0.88043493

00:25:12.473 --> 00:25:14.189 and industrial experiments,

NOTE Confidence: 0.88043493

00:25:14.190 --> 00:25:16.350 and in psychology people again are

NOTE Confidence: 0.88043493

00:25:16.350 --> 00:25:18.320 thinking about multi factor studies.

NOTE Confidence: 0.88043493

00:25:18.320 --> 00:25:21.386 So they're thinking about things you could

NOTE Confidence: 0.88043493

00:25:21.386 --> 00:25:24.510 better in the experimenters control.

NOTE Confidence: 0.88043493

00:25:24.510 --> 00:25:27.009 But we could instead bout sampling in

NOTE Confidence: 0.88043493

00:25:27.009 --> 00:25:29.829 the same as the same kind of thing.

NOTE Confidence: 0.88043493

00:25:29.830 --> 00:25:31.720 Except that we don't have control

NOTE Confidence: 0.88043493

00:25:31.720 --> 00:25:32.665 over manipulating them.

NOTE Confidence: 0.88043493

00:25:32.670 --> 00:25:34.958 We can find these units and as as

NOTE Confidence: 0.88043493

00:25:34.958 --> 00:25:36.130 an alternative approach,

NOTE Confidence: 0.88043493

00:25:36.130 --> 00:25:38.098 so one of the things we want to

NOTE Confidence: 0.88043493

00:25:38.098 --> 00:25:40.450 do is we want to make sure that

NOTE Confidence: 0.88043493

00:25:40.450 --> 00:25:42.549 we observe the full range of

NOTE Confidence: 0.88043493

00:25:42.549 --> 00:25:44.634 covariate values in the population,

NOTE Confidence: 0.88043493

00:25:44.640 --> 00:25:46.838 so it requires us to actually think,

NOTE Confidence: 0.88043493

00:25:46.840 --> 00:25:47.484 you know,

NOTE Confidence: 0.88043493

00:25:47.484 --> 00:25:49.094 explore the population data and

NOTE Confidence: 0.88043493

00:25:49.094 --> 00:25:51.023 make sure that we can understand

NOTE Confidence: 0.88043493

00:25:51.023 --> 00:25:52.823 what that range of values is.

NOTE Confidence: 0.88043493

00:25:52.830 --> 00:25:54.818 We might need to think carefully about

NOTE Confidence: 0.88043493

00:25:54.818 --> 00:25:56.610 moderators that are highly correlated.

NOTE Confidence: 0.88043493

00:25:56.610 --> 00:25:59.760 It can be very hard to D alias these effects,

NOTE Confidence: 0.88043493

00:25:59.760 --> 00:26:02.329 so if you have two highly correlated

NOTE Confidence: 0.88043493

00:26:02.329 --> 00:26:03.956 moderators. I think about that.

NOTE Confidence: 0.88043493

00:26:03.956 --> 00:26:05.526 I have two highly correlated

NOTE Confidence: 0.88043493

00:26:05.526 --> 00:26:06.429 moderators like this.

NOTE Confidence: 0.88043493

00:26:06.430 --> 00:26:08.296 If I want to estimate and

NOTE Confidence: 0.88043493

00:26:08.296 --> 00:26:09.540 understand moderators of X,

NOTE Confidence: 0.88043493

00:26:09.540 --> 00:26:11.430 if I want to explore X&Z and

NOTE Confidence: 0.88043493

00:26:11.430 --> 00:26:12.960 these are highly correlated,

NOTE Confidence: 0.88043493

00:26:12.960 --> 00:26:15.120 I'm going to really need to make sure

NOTE Confidence: 0.88043493

00:26:15.120 --> 00:26:17.222 I have those off diagonals that are

NOTE Confidence: 0.88043493

00:26:17.222 --> 00:26:20.187 kind of more rare in order to help me

NOTE Confidence: 0.88043493

00:26:20.187 --> 00:26:21.932 separate these effects an understand

NOTE Confidence: 0.88043493

00:26:21.932 --> 00:26:23.358 the unique contribution of each.

NOTE Confidence: 0.88043493

00:26:23.358 --> 00:26:25.368 The other is that if we might

NOTE Confidence: 0.88043493

00:26:25.368 --> 00:26:26.784 have many potential moderators

NOTE Confidence: 0.88043493

00:26:26.784 --> 00:26:28.200 that we're interested in,

NOTE Confidence: 0.88043493

00:26:28.200 --> 00:26:30.027 and so we're going to have to

NOTE Confidence: 0.88043493

00:26:30.027 --> 00:26:31.965 anticipate this in advance and think

NOTE Confidence: 0.88043493

00:26:31.965 --> 00:26:33.800 carefully about sort of compromises,

NOTE Confidence: 0.88043493

00:26:33.800 --> 00:26:36.050 we might need to make here.

NOTE Confidence: 0.88043493

00:26:36.050 --> 00:26:37.390 But also think very carefully,

NOTE Confidence: 0.88043493

00:26:37.390 --> 00:26:39.676 like we're not going to be able to expand

NOTE Confidence: 0.88043493

00:26:39.676 --> 00:26:41.918 this study to have a much bigger sample.

NOTE Confidence: 0.88043493

00:26:41.920 --> 00:26:43.856 So a lot of what I'm trying to

NOTE Confidence: 0.88043493

00:26:43.856 --> 00:26:45.659 operate under the constraint here is,

NOTE Confidence: 0.88043493

00:26:45.660 --> 00:26:47.220 let's not change the sample size

NOTE Confidence: 0.88043493

00:26:47.220 --> 00:26:49.130 if we don't change the sample size,

NOTE Confidence: 0.88043493

00:26:49.130 --> 00:26:50.708 but we instead change the height

NOTE Confidence: 0.88043493

00:26:50.708 --> 00:26:52.340 types of units in our study,

NOTE Confidence: 0.88043493

00:26:52.340 --> 00:26:55.238 how much better can we do?

NOTE Confidence: 0.88043493

00:26:55.240 --> 00:26:55.590 OK,

NOTE Confidence: 0.88043493

00:26:55.590 --> 00:26:58.040 so this leads to a principle found

NOTE Confidence: 0.88043493

00:26:58.040 --> 00:27:00.495 in response surface models called D

NOTE Confidence: 0.88043493

00:27:00.495 --> 00:27:03.015 optimality and so AD optimal design.

NOTE Confidence: 0.88043493

00:27:03.020 --> 00:27:06.512 This is work from the 40s and 60 Forties,

NOTE Confidence: 0.88043493

00:27:06.520 --> 00:27:07.585 50s and 60s.

NOTE Confidence: 0.88043493

00:27:07.585 --> 00:27:10.800 A lot of work here by Walt Kiefer,

NOTE Confidence: 0.88043493

00:27:10.800 --> 00:27:13.128 and a lot of people in

NOTE Confidence: 0.88043493

00:27:13.128 --> 00:27:13.904 industrial experiments.

NOTE Confidence: 0.88043493

00:27:13.910 --> 00:27:16.318 The idea is that you can instead focus

NOTE Confidence: 0.88043493

00:27:16.318 --> 00:27:19.340 on the generalized variance an you want

NOTE Confidence: 0.88043493

00:27:19.340 --> 00:27:21.685 to minimize the generalized variance,

NOTE Confidence: 0.88043493

00:27:21.690 --> 00:27:23.562 which is the determinant.

NOTE Confidence: 0.88043493

00:27:23.562 --> 00:27:25.902 So D is for determinant.

NOTE Confidence: 0.88043493

00:27:25.910 --> 00:27:27.452 And so the design that meets

NOTE Confidence: 0.88043493

00:27:27.452 --> 00:27:29.271 this criteria is one that also

NOTE Confidence: 0.88043493

00:27:29.271 --> 00:27:30.739 conveniently minimizes the maximum

NOTE Confidence: 0.88043493

00:27:30.739 --> 00:27:32.646 variance of any predicted outcome

NOTE Confidence: 0.88043493

00:27:32.646 --> 00:27:34.158 based upon these covariates.

NOTE Confidence: 0.88043493

00:27:34.160 --> 00:27:36.134 So this is great if what you're

NOTE Confidence: 0.88043493

00:27:36.134 --> 00:27:38.749 headed for is trying to make predict

NOTE Confidence: 0.88043493

00:27:38.749 --> 00:27:40.453 individual treatment effects or

NOTE Confidence: 0.88043493

00:27:40.453 --> 00:27:42.499 site specific treatment effects.

NOTE Confidence: 0.88043493

00:27:42.500 --> 00:27:44.276 The nice thing about a method

NOTE Confidence: 0.88043493

00:27:44.276 --> 00:27:46.190 that's been around for a while

NOTE Confidence: 0.88043493

00:27:46.190 --> 00:27:47.870 is that there's been algorithms

NOTE Confidence: 0.88043493

00:27:47.870 --> 00:27:49.220 developed for doing this.

NOTE Confidence: 0.88043493

00:27:49.220 --> 00:27:51.075 Better out Federov win algorithm

NOTE Confidence: 0.88043493

00:27:51.075 --> 00:27:52.930 is widely used and variations

NOTE Confidence: 0.831009269999999

00:27:52.992 --> 00:27:55.240 of it and that these are package that

NOTE Confidence: 0.831009269999999

00:27:55.240 --> 00:27:57.100 there are like statistics package

NOTE Confidence: 0.831009269999999

00:27:57.100 --> 00:27:59.155 already available that do this.

NOTE Confidence: 0.831009269999999

00:27:59.160 --> 00:28:01.398 So in our there's something called

NOTE Confidence: 0.831009269999999

00:28:01.398 --> 00:28:03.942 the ALG design package that is set

NOTE Confidence: 0.831009269999999

00:28:03.942 --> 00:28:05.904 up to actually work through this.

NOTE Confidence: 0.831009269999999

00:28:05.910 --> 00:28:08.388 So designs that we know are optimal.

NOTE Confidence: 0.831009269999999

00:28:08.390 --> 00:28:09.398 In other contexts.

NOTE Confidence: 0.831009269999999

00:28:09.398 --> 00:28:12.300 You know like our designs like Latin squares,

NOTE Confidence: 0.831009269999999

00:28:12.300 --> 00:28:15.140 designs etc all become special cases of this.

NOTE Confidence: 0.831009269999999

00:28:15.140 --> 00:28:18.500 So this is a much more general framework

NOTE Confidence: 0.831009269999999

00:28:18.500 --> 00:28:21.640 that doesn't require as many assumptions.

NOTE Confidence: 0.831009269999999

00:28:21.640 --> 00:28:24.034 OK, so once you start down this

NOTE Confidence: 0.831009269999999

00:28:24.034 --> 00:28:26.719 path you realize too that there are

NOTE Confidence: 0.831009269999999

00:28:26.719 --> 00:28:28.915 some tradeoffs here, so we have.

NOTE Confidence: 0.831009269999999

00:28:28.915 --> 00:28:31.739 You can easily imagine that the design that

NOTE Confidence: 0.831009269999999

00:28:31.739 --> 00:28:34.560 is optimal for an average treatment effect,

NOTE Confidence: 0.831009269999999

00:28:34.560 --> 00:28:36.768 which might be a representative sample.

NOTE Confidence: 0.831009269999999

00:28:36.770 --> 00:28:39.298 That sort of like a miniature of the

NOTE Confidence: 0.831009269999999

00:28:39.298 --> 00:28:41.588 population on covariates is likely not

NOTE Confidence: 0.831009269999999

00:28:41.588 --> 00:28:43.976 optimal for some of these standardized

NOTE Confidence: 0.831009269999999

00:28:44.045 --> 00:28:46.319 effect size differences where we might

NOTE Confidence: 0.831009269999999

00:28:46.319 --> 00:28:48.944 need to oversample in order to estimate,

NOTE Confidence: 0.831009269999999

00:28:48.944 --> 00:28:50.048 estimate, estimate these,

NOTE Confidence: 0.831009269999999

00:28:50.050 --> 00:28:51.610 and so there's another.

NOTE Confidence: 0.831009269999999

00:28:51.610 --> 00:28:53.170 Benefit of this approach,

NOTE Confidence: 0.831009269999999

00:28:53.170 --> 00:28:55.767 which is that you can focus on

NOTE Confidence: 0.831009269999999

00:28:55.767 --> 00:28:57.624 augmentation approach and what that

NOTE Confidence: 0.831009269999999

00:28:57.624 --> 00:28:59.990 means is you can actually say using

NOTE Confidence: 0.831009269999999

00:28:59.990 --> 00:29:02.062 these algorithms better billable 30

NOTE Confidence: 0.831009269999999

00:29:02.062 --> 00:29:04.946 sites or already for I've already got

NOTE Confidence: 0.831009269999999

00:29:04.946 --> 00:29:07.954 30 design run so the language of this

NOTE Confidence: 0.831009269999999

00:29:07.954 --> 00:29:11.249 is these sites become designed runs.

NOTE Confidence: 0.831009269999999

00:29:11.250 --> 00:29:13.280 And I need to select 10 more.

NOTE Confidence: 0.87889427

00:29:15.330 --> 00:29:16.914 Meaning population units, what?

NOTE Confidence: 0.87889427

00:29:16.914 --> 00:29:20.341 10 units can I augment it with that will

NOTE Confidence: 0.87889427

00:29:20.341 --> 00:29:23.210 improve that will make this as D optimal

NOTE Confidence: 0.87889427

00:29:23.210 --> 00:29:25.755 as possible given these constraints,

NOTE Confidence: 0.87889427

00:29:25.760 --> 00:29:28.514 and so instead so we're thinking

NOTE Confidence: 0.87889427

00:29:28.514 --> 00:29:30.913 of population units as possible

NOTE Confidence: 0.87889427

00:29:30.913 --> 00:29:33.769 design runs and sample as design

NOTE Confidence: 0.87889427

00:29:33.769 --> 00:29:36.448 runs that we've chosen to use.

NOTE Confidence: 0.87889427

00:29:36.450 --> 00:29:39.762 OK, so I'm just going to go through

NOTE Confidence: 0.87889427

00:29:39.762 --> 00:29:42.388 an example to talk about this.

NOTE Confidence: 0.87089115

00:29:45.440 --> 00:29:47.456 Don't have a ton more slides

NOTE Confidence: 0.87089115

00:29:47.456 --> 00:29:49.490 I should say so success.

NOTE Confidence: 0.87089115

00:29:49.490 --> 00:29:51.330 OK, so here's an example.

NOTE Confidence: 0.87089115

00:29:51.330 --> 00:29:53.466 The success for all evaluation was

NOTE Confidence: 0.87089115

00:29:53.466 --> 00:29:55.332 an elementary school reading program

NOTE Confidence: 0.87089115

00:29:55.332 --> 00:29:57.217 evaluated between 2001 and 2003.

NOTE Confidence: 0.87089115

00:29:57.220 --> 00:30:00.388 The reason I like to use this example.

NOTE Confidence: 0.87089115

00:30:00.390 --> 00:30:02.476 Is that it's old enough that strangely,

NOTE Confidence: 0.87089115

00:30:02.480 --> 00:30:03.875 they actually published in their

NOTE Confidence: 0.87089115

00:30:03.875 --> 00:30:05.650 paper a list of schools they

NOTE Confidence: 0.87089115

00:30:05.650 --> 00:30:07.624 actually named the schools in their

NOTE Confidence: 0.87089115

00:30:07.624 --> 00:30:09.059 study and characteristics of them.

NOTE Confidence: 0.87089115

00:30:09.060 --> 00:30:11.083 I have other data on other studies

NOTE Confidence: 0.87089115

00:30:11.083 --> 00:30:12.920 where people have shared with me

NOTE Confidence: 0.87089115

00:30:12.920 --> 00:30:14.738 the names of the schools involved,

NOTE Confidence: 0.87089115

00:30:14.740 --> 00:30:17.395 but it's all like I have to keep it

NOTE Confidence: 0.87089115

00:30:17.395 --> 00:30:19.527 secret for the for IR be reason,

NOTE Confidence: 0.87089115

00:30:19.530 --> 00:30:21.595 so that the fact that this is

NOTE Confidence: 0.87089115

00:30:21.595 --> 00:30:23.409 available makes it easier to use.

NOTE Confidence: 0.87089115

00:30:23.410 --> 00:30:25.442 So what I did is I went back

NOTE Confidence: 0.87089115

00:30:25.442 --> 00:30:27.642 and looked at the Common Core of

NOTE Confidence: 0.87089115

00:30:27.642 --> 00:30:29.666 data I identified based upon the

NOTE Confidence: 0.87089115

00:30:29.666 --> 00:30:30.966 study that they were.

NOTE Confidence: 0.87089115

00:30:30.970 --> 00:30:32.400 In the way that they

NOTE Confidence: 0.87089115

00:30:32.400 --> 00:30:33.544 talked about their study,

NOTE Confidence: 0.87089115

00:30:33.550 --> 00:30:35.215 that Title One elementary schools

NOTE Confidence: 0.87089115

00:30:35.215 --> 00:30:38.415 in the US at that time might be a

NOTE Confidence: 0.87089115

00:30:38.415 --> 00:30:40.245 reasonable population to think that

NOTE Confidence: 0.87089115

00:30:40.245 --> 00:30:42.379 they were trying to sample for.

NOTE Confidence: 0.87089115

00:30:42.380 --> 00:30:44.648 Title one schools have at least 40%

NOTE Confidence: 0.87089115

00:30:44.650 --> 00:30:46.498 students on free or reduced lunch and

NOTE Confidence: 0.87089115

00:30:46.498 --> 00:30:48.352 meet a few other characteristics and

NOTE Confidence: 0.87089115

00:30:48.352 --> 00:30:50.697 then they identified in the paper 5

NOTE Confidence: 0.87089115

00:30:50.759 --> 00:30:53.069 variables that they thought were possible.

NOTE Confidence: 0.87089115

00:30:53.070 --> 00:30:54.690 Moderators that would be really

NOTE Confidence: 0.87089115

00:30:54.690 --> 00:30:55.986 important to include here,

NOTE Confidence: 0.87089115

00:30:55.990 --> 00:30:57.934 so they talked about total school

NOTE Confidence: 0.87089115

00:30:57.934 --> 00:30:59.230 enrollment being a factor,

NOTE Confidence: 0.87089115

00:30:59.230 --> 00:31:00.526 racial and ethnic composition

NOTE Confidence: 0.87089115

00:31:00.526 --> 00:31:01.498 of the students.

NOTE Confidence: 0.87089115

00:31:01.500 --> 00:31:03.670 So I'm using that here as the

NOTE Confidence: 0.87089115

00:31:03.670 --> 00:31:05.644 proportion of students that are black

NOTE Confidence: 0.87089115

00:31:05.644 --> 00:31:07.612 and the proportion that are Hispanic

NOTE Confidence: 0.87089115

00:31:07.612 --> 00:31:09.932 and SES meaning and a professor but

NOTE Confidence: 0.87089115

00:31:09.932 --> 00:31:11.894 proportion at free and reduced lunch.

NOTE Confidence: 0.87089115

00:31:11.894 --> 00:31:13.916 And they also talk about Urbanicity

NOTE Confidence: 0.87089115

00:31:13.916 --> 00:31:15.963 because they tried to make sure they

NOTE Confidence: 0.87089115

00:31:15.963 --> 00:31:17.665 had some urban schools in rural

NOTE Confidence: 0.87089115

00:31:17.665 --> 00:31:19.315 schools and some other schools.

NOTE Confidence: 0.87089115

00:31:19.320 --> 00:31:21.240 So I should say in previous work of

NOTE Confidence: 0.87089115

00:31:21.240 --> 00:31:23.172 mine I've used this as an example

NOTE Confidence: 0.87089115

00:31:23.172 --> 00:31:24.955 and then this study actually ends

NOTE Confidence: 0.87089115

00:31:24.955 --> 00:31:26.640 up being a fairly representative

NOTE Confidence: 0.87089115

00:31:26.640 --> 00:31:27.988 sample of the population,

NOTE Confidence: 0.87089115

00:31:27.990 --> 00:31:28.890 which is interesting.

NOTE Confidence: 0.87089115

00:31:28.890 --> 00:31:31.313 Is it because they had no real way

NOTE Confidence: 0.87089115

00:31:31.313 --> 00:31:33.133 of they weren't doing it totally in

NOTE Confidence: 0.87089115

00:31:33.133 --> 00:31:35.182 a way that allowed them to compare

NOTE Confidence: 0.87089115

00:31:35.182 --> 00:31:37.377 this or to choose this in a way,

NOTE Confidence: 0.87089115

00:31:37.377 --> 00:31:39.871 but they did a lot of work to try

NOTE Confidence: 0.87089115

00:31:39.871 --> 00:31:40.990 to be representative,

NOTE Confidence: 0.87089115

00:31:40.990 --> 00:31:43.000 and this is much more representative

NOTE Confidence: 0.87089115

00:31:43.000 --> 00:31:43.670 sample then.

NOTE Confidence: 0.87089115

00:31:43.670 --> 00:31:47.162 I take the modal study is in this domain.

NOTE Confidence: 0.87089115

00:31:47.170 --> 00:31:47.485 OK,

NOTE Confidence: 0.87089115

00:31:47.485 --> 00:31:50.005 So what I did for for this example

NOTE Confidence: 0.87089115

00:31:50.005 --> 00:31:53.129 as I'm comparing for you the actual

NOTE Confidence: 0.87089115

00:31:53.129 --> 00:31:54.945 sample that they selected,

NOTE Confidence: 0.87089115

00:31:54.950 --> 00:31:57.290 so it's always these five moderators.

NOTE Confidence: 0.87089115

00:31:57.290 --> 00:31:59.230 The actual sample selected a

NOTE Confidence: 0.87089115

00:31:59.230 --> 00:32:00.394 representative sample selected.

NOTE Confidence: 0.87089115

00:32:00.400 --> 00:32:03.123 If I instead I use something like

NOTE Confidence: 0.87089115

00:32:03.123 --> 00:32:04.290 stratified random sampling.

NOTE Confidence: 0.87089115

00:32:04.290 --> 00:32:06.390 The optimal sample based upon these

NOTE Confidence: 0.87089115

00:32:06.390 --> 00:32:08.276 five covariates using this ALG

NOTE Confidence: 0.87089115

00:32:08.276 --> 00:32:10.426 design package and then various

NOTE Confidence: 0.87089115

00:32:10.426 --> 00:32:11.286 augmentation allocations.

NOTE Confidence: 0.87089115

00:32:11.290 --> 00:32:16.460 And So what I would do here as I'd say.

NOTE Confidence: 0.87089115

00:32:16.460 --> 00:32:19.108 So if I if I took 41, you know.

NOTE Confidence: 0.87089115

00:32:19.108 --> 00:32:20.980 So if I used 36 sites that were

NOTE Confidence: 0.87089115

00:32:21.044 --> 00:32:23.044 selected with random sampling with

NOTE Confidence: 0.87089115

00:32:23.044 --> 00:32:24.644 stratified random sampling and

NOTE Confidence: 0.87089115

00:32:24.644 --> 00:32:26.824 then I reserved five of them that

NOTE Confidence: 0.87089115

00:32:26.824 --> 00:32:28.280 were selected using D optimality

NOTE Confidence: 0.87089115

00:32:28.280 --> 00:32:30.452 and then I would change, you know,

NOTE Confidence: 0.87089115

00:32:30.452 --> 00:32:31.696 the number of those.

NOTE Confidence: 0.86540186

00:32:31.700 --> 00:32:34.188 So you could see this sort of effect.

NOTE Confidence: 0.86540186

00:32:34.190 --> 00:32:35.655 You know that augmentation would

NOTE Confidence: 0.86540186

00:32:35.655 --> 00:32:38.155 have and then for each of these I

NOTE Confidence: 0.86540186

00:32:38.155 --> 00:32:39.785 calculated a few different statistics.

NOTE Confidence: 0.86540186

00:32:39.790 --> 00:32:41.960 So you can see how this works.

NOTE Confidence: 0.86540186

00:32:41.960 --> 00:32:43.826 So one of them is D.

NOTE Confidence: 0.86540186

00:32:43.830 --> 00:32:45.790 This measure of the optimality.

NOTE Confidence: 0.86540186

00:32:45.790 --> 00:32:48.338 And I'm going to show you relative

NOTE Confidence: 0.86540186

00:32:48.338 --> 00:32:50.620 measures because it's a little easier

NOTE Confidence: 0.86540186

00:32:50.620 --> 00:32:52.490 to see with relative measures.

NOTE Confidence: 0.86540186

00:32:52.490 --> 00:32:53.870 I'm also including B,

NOTE Confidence: 0.86540186

00:32:53.870 --> 00:32:54.905 which is generalizability

NOTE Confidence: 0.86540186

00:32:54.905 --> 00:32:56.580 index that I developed.

NOTE Confidence: 0.86540186

00:32:56.580 --> 00:32:59.280 It ranges from zero to one and one means

NOTE Confidence: 0.86540186

00:32:59.280 --> 00:33:01.616 that the sample isn't exact miniature

NOTE Confidence: 0.86540186

00:33:01.616 --> 00:33:04.390 of the population on these covariates.

NOTE Confidence: 0.86540186

00:33:04.390 --> 00:33:07.384 0 means they like are completely

NOTE Confidence: 0.86540186

00:33:07.384 --> 00:33:09.820 orthogonal to each other and.

NOTE Confidence: 0.86540186

00:33:09.820 --> 00:33:13.572 Chip in its the index is highly related

NOTE Confidence: 0.86540186

00:33:13.572 --> 00:33:17.729 to measures of undercoverage and how and

NOTE Confidence: 0.86540186

00:33:17.729 --> 00:33:20.854 the performance of reweighting methods.

NOTE Confidence: 0.86540186

00:33:20.860 --> 00:33:24.380 And then the mean are meaning the ratio

NOTE Confidence: 0.86540186

00:33:24.380 --> 00:33:27.347 between the the ratio between the

NOTE Confidence: 0.86540186

00:33:27.347 --> 00:33:30.347 standard deviation in the sample and

NOTE Confidence: 0.86540186

00:33:30.438 --> 00:33:33.888 population across these five covariates.

NOTE Confidence: 0.86540186

00:33:33.890 --> 00:33:36.170 OK, so this is what we get out of this,

NOTE Confidence: 0.86540186

00:33:36.170 --> 00:33:38.130 and so I just want to talk through this and

NOTE Confidence: 0.86540186

00:33:38.178 --> 00:33:40.047 I'm happy to answer questions if there's.

NOTE Confidence: 0.86540186

00:33:40.050 --> 00:33:43.354 I know there's a lot going on here.

NOTE Confidence: 0.86540186

00:33:43.360 --> 00:33:46.690 Really wish I could figure out how to do a.

NOTE Confidence: 0.86540186

00:33:46.690 --> 00:33:48.900 Pointer.

NOTE Confidence: 0.86540186

00:33:48.900 --> 00:33:51.528 I don't think I can point out that way.

NOTE Confidence: 0.86540186

00:33:51.530 --> 00:33:52.259 OK, so OK.

NOTE Confidence: 0.86540186

00:33:52.259 --> 00:33:54.989 So what I have going on here is the number

NOTE Confidence: 0.86540186

00:33:54.989 --> 00:33:57.660 of sites randomly selected is left to right?

NOTE Confidence: 0.86540186

00:33:57.660 --> 00:33:59.250 So on the left is the

NOTE Confidence: 0.86540186

00:33:59.250 --> 00:34:00.870 is the D optimal sample,

NOTE Confidence: 0.86540186

00:34:00.870 --> 00:34:03.120 meaning the whole all 41 sites

NOTE Confidence: 0.86540186

00:34:03.120 --> 00:34:05.000 were actually selected using a

NOTE Confidence: 0.86540186

00:34:05.000 --> 00:34:06.540 D optimal algorithm on the.

NOTE Confidence: 0.86540186

00:34:06.540 --> 00:34:08.856 Right is the ideal for the

NOTE Confidence: 0.86540186

00:34:08.856 --> 00:34:10.014 average treatment effect.

NOTE Confidence: 0.86540186

00:34:10.020 --> 00:34:12.540 We've used random sampling to stratified

NOTE Confidence: 0.86540186

00:34:12.540 --> 00:34:14.876 random sampling and just like the

NOTE Confidence: 0.86540186

00:34:14.876 --> 00:34:17.204 the sample an in the bar right right

NOTE Confidence: 0.86540186

00:34:17.273 --> 00:34:19.277 there that like right up there.

NOTE Confidence: 0.86540186

00:34:19.280 --> 00:34:21.800 This Gray vertical bar is the actual

NOTE Confidence: 0.86540186

00:34:21.800 --> 00:34:23.900 study values for each of these.

NOTE Confidence: 0.86540186

00:34:23.900 --> 00:34:26.396 OK so you can see the actual study

NOTE Confidence: 0.86540186

00:34:26.396 --> 00:34:29.057 and then what I've got are three

NOTE Confidence: 0.86540186

00:34:29.057 --> 00:34:31.017 different lines going on here.

NOTE Confidence: 0.86540186

00:34:31.020 --> 00:34:33.176 So one line that's sloping down in

NOTE Confidence: 0.86540186

00:34:33.176 --> 00:34:36.008 solid is the relative D optimality value,

NOTE Confidence: 0.86540186

00:34:36.010 --> 00:34:37.129 so this is.

NOTE Confidence: 0.86540186

00:34:37.129 --> 00:34:39.740 You know the highest value is if

NOTE Confidence: 0.86540186

00:34:39.826 --> 00:34:42.370 it was a D optimal allocation.

NOTE Confidence: 0.86540186

00:34:42.370 --> 00:34:43.658 This is a ratio,

NOTE Confidence: 0.86540186

00:34:43.658 --> 00:34:46.160 and then I've got the B index,

NOTE Confidence: 0.86540186

00:34:46.160 --> 00:34:47.890 which is the generalizability index.

NOTE Confidence: 0.86540186

00:34:47.890 --> 00:34:50.298 Is the other solid line going up,

NOTE Confidence: 0.86540186

00:34:50.300 --> 00:34:51.504 and so, not surprisingly,

NOTE Confidence: 0.86540186

00:34:51.504 --> 00:34:53.009 that's increasing as we get

NOTE Confidence: 0.86540186

00:34:53.009 --> 00:34:54.439 to stratified sampling,

NOTE Confidence: 0.86540186

00:34:54.440 --> 00:34:56.510 so these are going in opposition

NOTE Confidence: 0.86540186

00:34:56.510 --> 00:34:57.545 to each other.

NOTE Confidence: 0.86540186

00:34:57.550 --> 00:34:59.776 Is what I'm saying and then this

NOTE Confidence: 0.86540186

00:34:59.776 --> 00:35:01.340 relative average standard deviation.

NOTE Confidence: 0.86540186

00:35:01.340 --> 00:35:03.070 Is this dotted bar line?

NOTE Confidence: 0.86540186

00:35:03.070 --> 00:35:05.086 So what so the main message of

NOTE Confidence: 0.86540186

00:35:05.086 --> 00:35:07.691 this is that these are going in

NOTE Confidence: 0.86540186

00:35:07.691 --> 00:35:09.363 opposite directions right that?

NOTE Confidence: 0.86540186

00:35:09.370 --> 00:35:11.954 The the sample that is optimal for the

NOTE Confidence: 0.86540186

00:35:11.954 --> 00:35:13.987 average treatment effect is on the right.

NOTE Confidence: 0.86540186

00:35:13.990 --> 00:35:15.922 The sample that is optimal for

NOTE Confidence: 0.86540186

00:35:15.922 --> 00:35:17.988 moderate are effects is on the left,

NOTE Confidence: 0.86540186

00:35:17.990 --> 00:35:20.185 and so there's there's tradeoffs

NOTE Confidence: 0.86540186

00:35:20.185 --> 00:35:22.763 involved in these that what's best

NOTE Confidence: 0.86540186

00:35:22.763 --> 00:35:25.219 for one is not best for the other.

NOTE Confidence: 0.86540186

00:35:25.220 --> 00:35:26.948 But there's other lessons in here,

NOTE Confidence: 0.86540186

00:35:26.950 --> 00:35:28.959 wow, so the the B index is,

NOTE Confidence: 0.86540186

00:35:28.960 --> 00:35:30.520 which is a measure of similarity

NOTE Confidence: 0.86540186

00:35:30.520 --> 00:35:32.130 between the sample and population,

NOTE Confidence: 0.85962015

00:35:32.130 --> 00:35:33.565 is actually not that bad

NOTE Confidence: 0.85962015

00:35:33.565 --> 00:35:34.713 for the optimal sample.

NOTE Confidence: 0.85962015

00:35:34.720 --> 00:35:36.448 So these these the sample is

NOTE Confidence: 0.85962015

00:35:36.448 --> 00:35:37.600 different from the population.

NOTE Confidence: 0.85962015

00:35:37.600 --> 00:35:39.616 You'd have to do some re waiting,

NOTE Confidence: 0.85962015

00:35:39.620 --> 00:35:41.240 but it wouldn't be a tremendous

NOTE Confidence: 0.85962015

00:35:41.240 --> 00:35:43.585 amount of re waiting to be able to

NOTE Confidence: 0.85962015

00:35:43.585 --> 00:35:45.090 estimate the average treatment effect.

NOTE Confidence: 0.85962015

00:35:45.090 --> 00:35:46.764 And so one lesson that you

NOTE Confidence: 0.85962015

00:35:46.764 --> 00:35:48.260 could think of it from.

NOTE Confidence: 0.85962015

00:35:48.260 --> 00:35:50.460 This is if you actually if we designed

NOTE Confidence: 0.85962015

00:35:50.460 --> 00:35:51.998 randomized trials to test moderators,

NOTE Confidence: 0.85962015

00:35:52.000 --> 00:35:53.548 we'd actually be in a pretty

NOTE Confidence: 0.85962015

00:35:53.548 --> 00:35:55.250 good space to test moderators.

NOTE Confidence: 0.85962015

00:35:55.250 --> 00:35:56.930 And to estimate the average treatment effect,

NOTE Confidence: 0.85962015

00:35:56.930 --> 00:35:58.370 it wouldn't be that far off.

NOTE Confidence: 0.85962015

00:35:58.370 --> 00:35:59.555 It wouldn't be.

NOTE Confidence: 0.85962015

00:35:59.555 --> 00:36:01.135 It wouldn't be terrible,

NOTE Confidence: 0.85962015

00:36:01.140 --> 00:36:03.390 and that makes sense because we're

NOTE Confidence: 0.85962015

00:36:03.390 --> 00:36:05.771 covering so much of the population

NOTE Confidence: 0.85962015

00:36:05.771 --> 00:36:08.592 by getting her across a bunch of

NOTE Confidence: 0.85962015

00:36:08.592 --> 00:36:10.609 moderators that we can do so that

NOTE Confidence: 0.85962015

00:36:10.609 --> 00:36:13.390 we can re wait when in a domain in

NOTE Confidence: 0.85962015

00:36:13.390 --> 00:36:15.140 which there's no act extrapolations,

NOTE Confidence: 0.85962015

00:36:15.140 --> 00:36:18.290 we have positive ITI we can re wait next.

NOTE Confidence: 0.85962015

00:36:18.290 --> 00:36:20.994 Another sort of I think finding here is

NOTE Confidence: 0.85962015

00:36:20.994 --> 00:36:24.240 if we look over at the right hand side.

NOTE Confidence: 0.85962015

00:36:24.240 --> 00:36:27.390 If we do, you know the trade off is.

NOTE Confidence: 0.85962015

00:36:27.390 --> 00:36:30.050 If I do select for the average

NOTE Confidence: 0.85962015

00:36:30.050 --> 00:36:30.810 treatment effect.

NOTE Confidence: 0.85962015

00:36:30.810 --> 00:36:32.180 I do get a tremendously,

NOTE Confidence: 0.85962015

00:36:32.180 --> 00:36:34.348 you know I can select for the average

NOTE Confidence: 0.85962015

00:36:34.348 --> 00:36:35.750 treatment effect and do pretty

NOTE Confidence: 0.85962015

00:36:35.750 --> 00:36:37.358 well for the average human effect,

NOTE Confidence: 0.85962015

00:36:37.360 --> 00:36:39.264 but not do so well for that.

NOTE Confidence: 0.85962015

00:36:39.270 --> 00:36:40.050 For the moderators,

NOTE Confidence: 0.85962015

00:36:40.050 --> 00:36:41.610 and so what's ideal for average

NOTE Confidence: 0.85962015

00:36:41.610 --> 00:36:43.055 is definitely not deal for

NOTE Confidence: 0.85962015

00:36:43.055 --> 00:36:44.187 the moderate are tests.

NOTE Confidence: 0.85962015

00:36:44.190 --> 00:36:45.968 So and then the third thing would

NOTE Confidence: 0.85962015

00:36:45.968 --> 00:36:48.278 be if you look at the actual study.

NOTE Confidence: 0.85962015

00:36:48.280 --> 00:36:49.216 As I was saying,

NOTE Confidence: 0.85962015

00:36:49.216 --> 00:36:51.013 they actually did a pretty good job

NOTE Confidence: 0.85962015

00:36:51.013 --> 00:36:52.377 in terms of representativeness.

NOTE Confidence: 0.85962015

00:36:52.380 --> 00:36:54.284 You can see that that top dot,

NOTE Confidence: 0.85962015

00:36:54.290 --> 00:36:56.390 but if you look at the bottom

NOTE Confidence: 0.85962015

00:36:56.390 --> 00:36:58.594 at the other two dots you can

NOTE Confidence: 0.85962015

00:36:58.594 --> 00:37:00.800 see they didn't do so well for.

NOTE Confidence: 0.85962015

00:37:00.800 --> 00:37:03.229 Being able to test these these moderators.

NOTE Confidence: 0.861584369999999

00:37:06.550 --> 00:37:09.142 OK, so in case that was not intuitive

NOTE Confidence: 0.861584369999999

00:37:09.142 --> 00:37:11.420 another way you could look at this

NOTE Confidence: 0.861584369999999

00:37:11.420 --> 00:37:13.830 is to actually just look at what

NOTE Confidence: 0.861584369999999

00:37:13.830 --> 00:37:15.785 these samples these these features

NOTE Confidence: 0.861584369999999

00:37:15.785 --> 00:37:17.755 of these samples would look like.

NOTE Confidence: 0.861584369999999

00:37:17.755 --> 00:37:20.210 So in the top the top row here

NOTE Confidence: 0.861584369999999

00:37:20.210 --> 00:37:21.818 are population distributions.

NOTE Confidence: 0.861584369999999

00:37:21.820 --> 00:37:23.550 Of these five covariates that

NOTE Confidence: 0.861584369999999

00:37:23.550 --> 00:37:24.934 were sort of identified,

NOTE Confidence: 0.861584369999999

00:37:24.940 --> 00:37:27.341 and then at the bottom row is

NOTE Confidence: 0.861584369999999

00:37:27.341 --> 00:37:29.450 actually the study that they had.

NOTE Confidence: 0.861584369999999

00:37:29.450 --> 00:37:32.537 So what their actual sample looked like.

NOTE Confidence: 0.861584369999999

00:37:32.540 --> 00:37:34.808 And then the middle is what AD

NOTE Confidence: 0.861584369999999

00:37:34.808 --> 00:37:36.479 optimal sample would look like.

NOTE Confidence: 0.861584369999999

00:37:36.480 --> 00:37:38.436 And then I've overlaid on here.

NOTE Confidence: 0.861584369999999

00:37:38.440 --> 00:37:39.433 These are values,

NOTE Confidence: 0.861584369999999

00:37:39.433 --> 00:37:43.037 so giving you a sense if R is greater is 1.

NOTE Confidence: 0.861584369999999

00:37:43.040 --> 00:37:45.656 It means the sample is like the same

NOTE Confidence: 0.861584369999999

00:37:45.656 --> 00:37:47.630 standard deviation as in the population.

NOTE Confidence: 0.861584369999999

00:37:47.630 --> 00:37:49.598 If R is greater than one,

NOTE Confidence: 0.861584369999999

00:37:49.600 --> 00:37:51.514 it means I've got more heterogeneity

NOTE Confidence: 0.861584369999999

00:37:51.514 --> 00:37:53.859 in my sample than in my population,

NOTE Confidence: 0.861584369999999

00:37:53.860 --> 00:37:55.495 which improves my ability to

NOTE Confidence: 0.861584369999999

00:37:55.495 --> 00:37:56.803 estimate moderate are effects.

NOTE Confidence: 0.861584369999999

00:37:56.810 --> 00:37:59.753 And So what you see are a few things.

NOTE Confidence: 0.861584369999999

00:37:59.760 --> 00:38:02.488 One is in that the optimal sample is.

NOTE Confidence: 0.861584369999999

00:38:02.490 --> 00:38:04.686 It pushes things towards the extremes,

NOTE Confidence: 0.861584369999999

00:38:04.690 --> 00:38:05.048 right?

NOTE Confidence: 0.861584369999999

00:38:05.048 --> 00:38:06.838 It's pushing them towards the

NOTE Confidence: 0.861584369999999

00:38:06.838 --> 00:38:09.072 extremes to get endpoints which we

NOTE Confidence: 0.861584369999999

00:38:09.072 --> 00:38:10.927 know from basic experimental design,

NOTE Confidence: 0.861584369999999

00:38:10.930 --> 00:38:11.664 improved abilities.

NOTE Confidence: 0.861584369999999

00:38:11.664 --> 00:38:13.499 The other nice thing though,

NOTE Confidence: 0.861584369999999

00:38:13.500 --> 00:38:15.864 is a concern always when you're

NOTE Confidence: 0.861584369999999

00:38:15.864 --> 00:38:17.840 doing experimental design like this

NOTE Confidence: 0.861584369999999

00:38:17.840 --> 00:38:20.157 is that you're going to get your

NOTE Confidence: 0.861584369999999

00:38:20.157 --> 00:38:22.330 highly focused on like a linearity

NOTE Confidence: 0.861584369999999

00:38:22.330 --> 00:38:24.502 assumption that you're going to your.

NOTE Confidence: 0.861584369999999

00:38:24.510 --> 00:38:26.784 Your ideal sample would have a

NOTE Confidence: 0.861584369999999

00:38:26.784 --> 00:38:28.550 strong linearity assumption to it,

NOTE Confidence: 0.861584369999999

00:38:28.550 --> 00:38:31.063 but because you have multiple variables an

NOTE Confidence: 0.861584369999999

00:38:31.063 --> 00:38:33.757 because not all design runs are possible.

NOTE Confidence: 0.861584369999999

00:38:33.760 --> 00:38:34.753 In the population,

NOTE Confidence: 0.861584369999999

00:38:34.753 --> 00:38:37.070 you end up with these middle points

NOTE Confidence: 0.861584369999999

00:38:37.135 --> 00:38:39.295 as well so you don't end up with

NOTE Confidence: 0.861584369999999

00:38:39.295 --> 00:38:40.938 only things on both extremes.

NOTE Confidence: 0.861584369999999

00:38:40.940 --> 00:38:43.355 You end up with some middle points

NOTE Confidence: 0.861584369999999

00:38:43.355 --> 00:38:46.572 which allow you to be able to estimate

NOTE Confidence: 0.861584369999999

00:38:46.572 --> 00:38:48.260 nonlinear relationships as well.

NOTE Confidence: 0.861584369999999

00:38:48.260 --> 00:38:50.178 Me and a Third Point with me.

NOTE Confidence: 0.861584369999999

00:38:50.180 --> 00:38:52.084 You can see that you would just end

NOTE Confidence: 0.861584369999999

00:38:52.084 --> 00:38:54.483 up with a lot more variation and so

NOTE Confidence: 0.861584369999999

00:38:54.483 --> 00:38:55.656 not surprisingly, total students,

NOTE Confidence: 0.861584369999999

00:38:55.656 --> 00:38:56.748 which, again schools studies,

NOTE Confidence: 0.861584369999999

00:38:56.750 --> 00:38:58.268 tend to over represent very large

NOTE Confidence: 0.861584369999999

00:38:58.268 --> 00:38:59.770 schools and large school districts.

NOTE Confidence: 0.861584369999999

00:38:59.770 --> 00:39:01.898 You can see this is a place where

NOTE Confidence: 0.861584369999999

00:39:01.898 --> 00:39:03.646 there would be really a real

NOTE Confidence: 0.861584369999999

00:39:03.646 --> 00:39:05.380 opportunity for a change that in

NOTE Confidence: 0.861584369999999

00:39:05.447 --> 00:39:07.225 the sample this was less than one

NOTE Confidence: 0.861584369999999

00:39:07.225 --> 00:39:09.330 an in the in the optimal sample

NOTE Confidence: 0.861584369999999

00:39:09.330 --> 00:39:11.340 it would be greater than three.

NOTE Confidence: 0.861584369999999

00:39:11.340 --> 00:39:13.908 But you can see this for most of

NOTE Confidence: 0.861584369999999

00:39:13.908 --> 00:39:15.630 these variables that you could.

NOTE Confidence: 0.861584369999999

00:39:15.630 --> 00:39:17.250 You could potentially improve your

NOTE Confidence: 0.861584369999999

00:39:17.250 --> 00:39:19.260 power and ability to estimate things

NOTE Confidence: 0.861584369999999

00:39:19.260 --> 00:39:20.910 related to demographics as well.

NOTE Confidence: 0.861584369999999

00:39:20.910 --> 00:39:23.318 And in my paper I actually show that

NOTE Confidence: 0.861584369999999

00:39:23.318 --> 00:39:25.530 because many of these are proportions,

NOTE Confidence: 0.861584369999999

00:39:25.530 --> 00:39:27.402 you can actually also think about

NOTE Confidence: 0.861584369999999

00:39:27.402 --> 00:39:29.088 student level moderate yrs because

NOTE Confidence: 0.861584369999999

00:39:29.088 --> 00:39:30.608 proportions conveniently like the

NOTE Confidence: 0.861584369999999

00:39:30.608 --> 00:39:32.508 variation in proportions at the

NOTE Confidence: 0.861584369999999

00:39:32.564 --> 00:39:34.280 individual level as a function of

NOTE Confidence: 0.861584369999999

00:39:34.280 --> 00:39:35.774 the proportion at the aggregate.

NOTE Confidence: 0.861584369999999

00:39:35.774 --> 00:39:38.526 And so you can actually kind of workout

NOTE Confidence: 0.861584369999999

00:39:38.526 --> 00:39:41.785 a way to select your samples so that you can.

NOTE Confidence: 0.861584369999999

00:39:41.790 --> 00:39:44.193 Estimate individual affects,

NOTE Confidence: 0.861584369999999

00:39:44.193 --> 00:39:47.397 not just cluster aggregates

NOTE Confidence: 0.861584369999999

00:39:47.397 --> 00:39:49.800 for those variables.

NOTE Confidence: 0.861584369999999

00:39:49.800 --> 00:39:51.949 OK, and so then the final point.

NOTE Confidence: 0.861584369999999

00:39:51.950 --> 00:39:54.278 I just want to make is that the

NOTE Confidence: 0.861584369999999

00:39:54.278 --> 00:39:56.962 other thing that this shows is that

NOTE Confidence: 0.861584369999999

00:39:56.962 --> 00:39:58.992 there's real benefit to augmentation

NOTE Confidence: 0.892888899999999

00:39:59.067 --> 00:40:00.152 so. Maybe? You know,

NOTE Confidence: 0.892888899999999

00:40:00.152 --> 00:40:02.642 maybe I'm not going to be able to

NOTE Confidence: 0.892888899999999

00:40:02.642 --> 00:40:05.183 convince people to go switch to selecting

NOTE Confidence: 0.892888899999999

00:40:05.183 --> 00:40:07.230 their samples based upon extremes.

NOTE Confidence: 0.892888899999999

00:40:07.230 --> 00:40:09.720 But maybe you can convince people

NOTE Confidence: 0.892888899999999

00:40:09.720 --> 00:40:12.030 that they could preserve 5 or 10.

NOTE Confidence: 0.892888899999999

00:40:12.030 --> 00:40:14.610 You know 10% or 25% of their

NOTE Confidence: 0.892888899999999

00:40:14.610 --> 00:40:16.090 sample for D optimality.

NOTE Confidence: 0.892888899999999

00:40:16.090 --> 00:40:16.960 So you choose.

NOTE Confidence: 0.892888899999999

00:40:16.960 --> 00:40:19.465 In this case it would be like choose

NOTE Confidence: 0.892888899999999

00:40:19.465 --> 00:40:21.913 30 of your sites using stratified

NOTE Confidence: 0.892888899999999

00:40:21.913 --> 00:40:24.199 sampling to represent the population,

NOTE Confidence: 0.892888899999999

00:40:24.200 --> 00:40:26.720 and then look for like an additional

NOTE Confidence: 0.892888899999999

00:40:26.720 --> 00:40:29.289 class tenor 11 sites that might be

NOTE Confidence: 0.892888899999999

00:40:29.289 --> 00:40:31.866 more extreme that allow you to make

NOTE Confidence: 0.892888899999999

00:40:31.866 --> 00:40:34.164 sure that you can estimate these.

NOTE Confidence: 0.892888899999999

00:40:34.170 --> 00:40:35.702 These moderate are effects

NOTE Confidence: 0.892888899999999

00:40:35.702 --> 00:40:37.234 that you're interested in.

NOTE Confidence: 0.892888899999999

00:40:37.240 --> 00:40:38.888 And you can see that doing so key

NOTE Confidence: 0.892888899999999

00:40:38.888 --> 00:40:40.599 file with these little lines you can

NOTE Confidence: 0.892888899999999

00:40:40.599 --> 00:40:42.716 see that doing so doesn't have a huge

NOTE Confidence: 0.892888899999999

00:40:42.716 --> 00:40:44.348 effect on the average treatment effect,

NOTE Confidence: 0.892888899999999

00:40:44.350 --> 00:40:45.575 but it does greatly improve

NOTE Confidence: 0.892888899999999

00:40:45.575 --> 00:40:46.800 your ability to test moderators.

NOTE Confidence: 0.87828714

00:40:49.510 --> 00:40:52.750 OK, so just to wrap up my take home

NOTE Confidence: 0.87828714

00:40:52.750 --> 00:40:55.850 points today, I suppose would be that

NOTE Confidence: 0.87828714

00:40:55.850 --> 00:40:59.122 the design of randomized trials has big

NOTE Confidence: 0.87828714

00:40:59.122 --> 00:41:01.737 implications for ability to generalize.

NOTE Confidence: 0.87828714

00:41:01.740 --> 00:41:04.260 And that I think we, I think what I've

NOTE Confidence: 0.87828714

00:41:04.260 --> 00:41:06.547 seen over time is that people who are

NOTE Confidence: 0.87828714

00:41:06.547 --> 00:41:08.359 starting to pay attention to that,

NOTE Confidence: 0.87828714

00:41:08.360 --> 00:41:09.800 and they're starting to think

NOTE Confidence: 0.87828714

00:41:09.800 --> 00:41:11.240 about how populations you know.

NOTE Confidence: 0.87828714

00:41:11.240 --> 00:41:12.740 What are the populations I would

NOTE Confidence: 0.87828714

00:41:12.740 --> 00:41:14.858 add as a side benefit of this is

NOTE Confidence: 0.87828714

00:41:14.858 --> 00:41:16.454 I've I've watched as people in

NOTE Confidence: 0.87828714

00:41:16.511 --> 00:41:18.527 asking people to scientists to think

NOTE Confidence: 0.87828714

00:41:18.527 --> 00:41:20.219 about what the population is.

NOTE Confidence: 0.87828714

00:41:20.219 --> 00:41:22.253 It actually sometimes make some change

NOTE Confidence: 0.87828714

00:41:22.253 --> 00:41:24.300 with the intervention is because you kind

NOTE Confidence: 0.87828714

00:41:24.300 --> 00:41:26.510 of have to realize like is this is this.

NOTE Confidence: 0.87828714

00:41:26.510 --> 00:41:27.950 If this is the population,

NOTE Confidence: 0.87828714

00:41:27.950 --> 00:41:31.190 is this the right intervention?

NOTE Confidence: 0.87828714

00:41:31.190 --> 00:41:33.790 The second sort of point I would say,

NOTE Confidence: 0.87828714

00:41:33.790 --> 00:41:36.346 is that if we want to sort of estimate

NOTE Confidence: 0.87828714

00:41:36.346 --> 00:41:38.189 and test hypothesis and moderators

NOTE Confidence: 0.87828714

00:41:38.189 --> 00:41:40.797 that we would be wise to actually

NOTE Confidence: 0.87828714

00:41:40.797 --> 00:41:43.533 plan to do so and to think about how

NOTE Confidence: 0.87828714

00:41:43.540 --> 00:41:45.390 to have better design sensitivity

NOTE Confidence: 0.87828714

00:41:45.390 --> 00:41:47.240 and statistical statistical power for

NOTE Confidence: 0.87828714

00:41:47.294 --> 00:41:48.854 doing so instead of waiting until

NOTE Confidence: 0.87828714

00:41:48.854 --> 00:41:51.285 the end and then the last point is

NOTE Confidence: 0.87828714

00:41:51.285 --> 00:41:52.950 just that this augmentation approach

NOTE Confidence: 0.87828714

00:41:52.950 --> 00:41:54.865 indicates that we don't have to

NOTE Confidence: 0.87828714

00:41:54.865 --> 00:41:56.870 be perfect at this like that,

NOTE Confidence: 0.87828714

00:41:56.870 --> 00:41:58.490 we could just, you know,

NOTE Confidence: 0.87828714

00:41:58.490 --> 00:42:01.426 use do this for part of our sample.

NOTE Confidence: 0.87828714

00:42:01.430 --> 00:42:03.550 And we would be better off and then

NOTE Confidence: 0.87828714

00:42:03.550 --> 00:42:05.969 I guess I would say maybe my general

NOTE Confidence: 0.87828714

00:42:05.969 --> 00:42:08.479 philosophy in all of this design is that.

NOTE Confidence: 0.87828714

00:42:08.480 --> 00:42:10.370 What I'm trying to do is to get people

NOTE Confidence: 0.87828714

00:42:10.370 --> 00:42:12.637 to think differently and plan differently,

NOTE Confidence: 0.87828714

00:42:12.640 --> 00:42:13.748 and by doing so,

NOTE Confidence: 0.87828714

00:42:13.748 --> 00:42:15.410 even if you don't succeed 100%,

NOTE Confidence: 0.87828714

00:42:15.410 --> 00:42:17.066 you're better off than you would

NOTE Confidence: 0.87828714

00:42:17.066 --> 00:42:17.894 have been before,

NOTE Confidence: 0.87828714

00:42:17.900 --> 00:42:19.652 and you're now able to be in the

NOTE Confidence: 0.87828714

00:42:19.652 --> 00:42:21.552 realm in which you have positive

NOTE Confidence: 0.87828714

00:42:21.552 --> 00:42:22.605 ITI and heterogeneity,

NOTE Confidence: 0.87828714

00:42:22.610 --> 00:42:24.765 and you're able to actually

NOTE Confidence: 0.87828714

00:42:24.765 --> 00:42:26.058 use statistical methods.

NOTE Confidence: 0.87828714

00:42:26.060 --> 00:42:27.936 To get better estimators at the end.

NOTE Confidence: 0.8810891

00:42:30.570 --> 00:42:33.482 Thank you, this is all my contact

NOTE Confidence: 0.8810891

00:42:33.482 --> 00:42:35.945 information and this is the paper

NOTE Confidence: 0.8810891

00:42:35.945 --> 00:42:38.207 that this talk is really about.

NOTE Confidence: 0.8810891

00:42:38.210 --> 00:42:40.220 I'm happy to answer questions.

NOTE Confidence: 0.86773777

00:42:41.220 --> 00:42:42.620 Thanks so much. Best,

NOTE Confidence: 0.86773777

00:42:42.620 --> 00:42:45.141 I think that's really nice talk and

NOTE Confidence: 0.86773777

00:42:45.141 --> 00:42:47.235 thank you for being so inspiring.

NOTE Confidence: 0.86773777

00:42:47.240 --> 00:42:49.166 And maybe let's open to questions

NOTE Confidence: 0.86773777

00:42:49.166 --> 00:42:51.694 1st to see if we have any

NOTE Confidence: 0.86773777

00:42:51.694 --> 00:42:53.254 questions from the audience.

NOTE Confidence: 0.86638516

00:42:55.280 --> 00:42:58.146 If so, please speak up or, you know,

NOTE Confidence: 0.86638516

00:42:58.146 --> 00:43:00.648 send a chat. Either one is OK.

NOTE Confidence: 0.9289141

00:43:07.090 --> 00:43:08.966 And if not, I can go first.

NOTE Confidence: 0.9289141

00:43:08.970 --> 00:43:12.514 'cause I do have a couple of questions.

NOTE Confidence: 0.9289141

00:43:12.520 --> 00:43:15.330 So, so first of all, I think you

NOTE Confidence: 0.9289141

00:43:15.330 --> 00:43:17.430 know there is a constant tension.

NOTE Confidence: 0.9289141

00:43:17.430 --> 00:43:20.238 Of course, like you know when we work with

NOTE Confidence: 0.9289141

00:43:20.238 --> 00:43:23.047 really large trials in the healthcare system,

NOTE Confidence: 0.9289141

00:43:23.050 --> 00:43:25.820 I think there is a tension between how do we

NOTE Confidence: 0.9289141

00:43:25.892 --> 00:43:28.670 better represent the population of interest?

NOTE Confidence: 0.9289141

00:43:28.670 --> 00:43:30.770 Because we want to get effectiveness

NOTE Confidence: 0.9289141

00:43:30.770 --> 00:43:31.820 information 'cause we're

NOTE Confidence: 0.9289141

00:43:31.820 --> 00:43:33.230 spending millions of dollars.

NOTE Confidence: 0.9289141

00:43:33.230 --> 00:43:35.534 But also I think there is a concern

NOTE Confidence: 0.9289141

00:43:35.534 --> 00:43:38.247 on you know how to really better

NOTE Confidence: 0.9289141

00:43:38.247 --> 00:43:39.895 engage these large clusters,

NOTE Confidence: 0.9289141

00:43:39.900 --> 00:43:42.000 large healthcare systems or large clinics,

NOTE Confidence: 0.9289141

00:43:42.000 --> 00:43:44.010 etc. And so I think.

NOTE Confidence: 0.9289141

00:43:44.010 --> 00:43:45.625 People end up getting convenience

NOTE Confidence: 0.9289141

00:43:45.625 --> 00:43:46.917 samples because that's reality.

NOTE Confidence: 0.9289141

00:43:46.920 --> 00:43:48.922 Even though I do believe that there's

NOTE Confidence: 0.9289141

00:43:48.922 --> 00:43:50.910 so much more to improve because

NOTE Confidence: 0.9289141

00:43:50.910 --> 00:43:53.046 they're spending so much money right.

NOTE Confidence: 0.9289141

00:43:53.050 --> 00:43:54.670 And then in the end,

NOTE Confidence: 0.9289141

00:43:54.670 --> 00:43:56.896 you know they may be answering a

NOTE Confidence: 0.9289141

00:43:56.896 --> 00:43:58.882 different question if they have a

NOTE Confidence: 0.9289141

00:43:58.882 --> 00:44:00.826 very highly selected sample and then

NOTE Confidence: 0.9289141

00:44:00.826 --> 00:44:02.738 people also worry about you know,

NOTE Confidence: 0.9289141

00:44:02.740 --> 00:44:04.522 like you know there are some

NOTE Confidence: 0.9289141

00:44:04.522 --> 00:44:06.300 disparities in their sample selection,

NOTE Confidence: 0.9289141

00:44:06.300 --> 00:44:08.136 so that you're basically not covering

NOTE Confidence: 0.9289141

00:44:08.136 --> 00:44:10.088 you know people with maybe more

NOTE Confidence: 0.9289141

00:44:10.088 --> 00:44:12.104 vulnerable conditions etc in your study,

NOTE Confidence: 0.9289141

00:44:12.110 --> 00:44:14.180 but you wish to answer questions.

NOTE Confidence: 0.9289141

00:44:14.180 --> 00:44:15.281 What is population?

NOTE Confidence: 0.9289141

00:44:15.281 --> 00:44:18.330 So I feel like all of this very,

NOTE Confidence: 0.9289141

00:44:18.330 --> 00:44:20.969 very relevant, at least to my work.

NOTE Confidence: 0.9289141

00:44:20.970 --> 00:44:23.826 And so I really appreciate you know this

NOTE Confidence: 0.9289141

00:44:23.826 --> 00:44:26.617 aspect of how to design styles better.

NOTE Confidence: 0.9289141

00:44:26.620 --> 00:44:29.259 1 one of the questions I have

NOTE Confidence: 0.9289141

00:44:29.259 --> 00:44:30.390 is that generally,

NOTE Confidence: 0.9289141

00:44:30.390 --> 00:44:33.062 you know we may not really know priority

NOTE Confidence: 0.9289141

00:44:33.062 --> 00:44:36.050 what the effect modifiers are in planning.

NOTE Confidence: 0.9289141

00:44:36.050 --> 00:44:38.312 The trial that we may have

NOTE Confidence: 0.9289141

00:44:38.312 --> 00:44:39.820 not enough knowledge amount.

NOTE Confidence: 0.9289141

00:44:39.820 --> 00:44:42.676 So how does that generally come into

NOTE Confidence: 0.9289141

00:44:42.676 --> 00:44:45.270 the discussion in the design stage?

NOTE Confidence: 0.9289141

00:44:45.270 --> 00:44:47.720 Is it the tradition that in educational

NOTE Confidence: 0.9289141

00:44:47.720 --> 00:44:50.589 studies we have a lot of prime knowledge

NOTE Confidence: 0.9289141

00:44:50.589 --> 00:44:52.600 on what these effect modifiers are

NOTE Confidence: 0.862398707692308

00:44:52.600 --> 00:44:54.768 or? No, so I think this is actually

NOTE Confidence: 0.862398707692308

00:44:54.768 --> 00:44:56.783 one of the hardest parts, right?

NOTE Confidence: 0.862398707692308

00:44:56.783 --> 00:44:57.998 Like I just laid out.

NOTE Confidence: 0.862398707692308

00:44:58.000 --> 00:44:59.673 Sort of, if we knew what the

NOTE Confidence: 0.862398707692308

00:44:59.673 --> 00:45:01.139 why zeros and Y ones were,

NOTE Confidence: 0.862398707692308

00:45:01.140 --> 00:45:02.350 this is what we would.

NOTE Confidence: 0.862398707692308

00:45:02.350 --> 00:45:04.286 You know this is that would be optimal,

NOTE Confidence: 0.862398707692308

00:45:04.290 --> 00:45:05.736 but I could be wrong on

NOTE Confidence: 0.862398707692308

00:45:05.736 --> 00:45:06.710 what those are, right?

NOTE Confidence: 0.8577501

00:45:09.130 --> 00:45:10.394 And I don't know.

NOTE Confidence: 0.8577501

00:45:10.394 --> 00:45:11.822 I mean, I think so.

NOTE Confidence: 0.8577501

00:45:11.822 --> 00:45:13.670 There's sort of what I call the

NOTE Confidence: 0.8577501

00:45:13.731 --> 00:45:15.447 usual suspects in education,

NOTE Confidence: 0.8577501

00:45:15.450 --> 00:45:17.655 which are like race class and gender,

NOTE Confidence: 0.8577501

00:45:17.660 --> 00:45:19.418 which are really more of concerns

NOTE Confidence: 0.8577501

00:45:19.418 --> 00:45:20.990 about disparity or about closing

NOTE Confidence: 0.8577501

00:45:20.990 --> 00:45:22.715 achievement gaps in various ways.

NOTE Confidence: 0.8577501

00:45:22.720 --> 00:45:24.460 And so those in depth and

NOTE Confidence: 0.8577501

00:45:24.460 --> 00:45:26.706 urbanicity I would add seems to be

NOTE Confidence: 0.8577501

00:45:26.706 --> 00:45:28.406 something that people often like.

NOTE Confidence: 0.8577501

00:45:28.410 --> 00:45:30.300 What add into that as characteristics.

NOTE Confidence: 0.8577501

00:45:30.300 --> 00:45:31.880 Those are the ones that

NOTE Confidence: 0.8577501

00:45:31.880 --> 00:45:33.144 people most often use.

NOTE Confidence: 0.8577501

00:45:33.150 --> 00:45:34.725 But the and those are

NOTE Confidence: 0.8577501

00:45:34.725 --> 00:45:35.985 available in population data,

NOTE Confidence: 0.8577501

00:45:35.990 --> 00:45:37.625 which is the other thing

NOTE Confidence: 0.8577501

00:45:37.625 --> 00:45:38.933 that your limit your.

NOTE Confidence: 0.8577501

00:45:38.940 --> 00:45:40.914 A real limiter is what is available

NOTE Confidence: 0.8577501

00:45:40.914 --> 00:45:43.430 in the population, sure.

NOTE Confidence: 0.8577501

00:45:43.430 --> 00:45:46.240 What I gather is more likely to be a moderate

NOTE Confidence: 0.8577501

00:45:46.309 --> 00:45:48.997 are or something like baseline achievement,

NOTE Confidence: 0.8577501

00:45:49.000 --> 00:45:49.348 right?

NOTE Confidence: 0.8577501

00:45:49.348 --> 00:45:52.480 So if my outcome is achievement then I would,

NOTE Confidence: 0.8577501

00:45:52.480 --> 00:45:54.526 I would think that what the

NOTE Confidence: 0.8577501

00:45:54.526 --> 00:45:56.309 achievement is baseline in any

NOTE Confidence: 0.8577501

00:45:56.309 --> 00:45:58.049 of these places would matter.

NOTE Confidence: 0.8577501

00:45:58.050 --> 00:46:00.126 That's harder to get an education.

NOTE Confidence: 0.8577501

00:46:00.130 --> 00:46:02.218 I mean that information from places,

NOTE Confidence: 0.8577501

00:46:02.220 --> 00:46:04.516 so there's been some work trying to

NOTE Confidence: 0.8577501

00:46:04.516 --> 00:46:06.400 equate tests across across states.

NOTE Confidence: 0.8577501

00:46:06.400 --> 00:46:08.140 I guess that they do.

NOTE Confidence: 0.8577501

00:46:08.140 --> 00:46:08.490 Sometimes

NOTE Confidence: 0.8494228

00:46:08.490 --> 00:46:10.962 they use gain scores just to subtract off

NOTE Confidence: 0.8494228

00:46:10.962 --> 00:46:13.010 that baseline achievement, right? They

NOTE Confidence: 0.851321227777778

00:46:13.010 --> 00:46:13.910 do. Yeah, exactly,

NOTE Confidence: 0.851321227777778

00:46:13.910 --> 00:46:16.392 but the problem is that like if you

NOTE Confidence: 0.851321227777778

00:46:16.392 --> 00:46:18.646 wanted to use state tests or something,

NOTE Confidence: 0.851321227777778

00:46:18.650 --> 00:46:20.778 there are different tests in every state,

NOTE Confidence: 0.851321227777778

00:46:20.780 --> 00:46:22.694 and so there's all of these

NOTE Confidence: 0.851321227777778

00:46:22.694 --> 00:46:24.730 equating issues that go in with it.

NOTE Confidence: 0.851321227777778

00:46:24.730 --> 00:46:26.746 My guess is that implementation is another

NOTE Confidence: 0.851321227777778

00:46:26.746 --> 00:46:29.240 one that people often come up with is

NOTE Confidence: 0.851321227777778

00:46:29.240 --> 00:46:30.500 like something with implementation.

NOTE Confidence: 0.851321227777778

00:46:30.500 --> 00:46:32.528 Now this is tricky because implementation

NOTE Confidence: 0.851321227777778

00:46:32.528 --> 00:46:34.178 is coming after assignment and

NOTE Confidence: 0.851321227777778

00:46:34.178 --> 00:46:35.666 so it's really like a mediator.

NOTE Confidence: 0.851321227777778

00:46:35.670 --> 00:46:37.500 But if you think about often,

NOTE Confidence: 0.851321227777778

00:46:37.500 --> 00:46:38.524 if you think implementation

NOTE Confidence: 0.851321227777778

00:46:38.524 --> 00:46:40.502 may be part of what is leading

NOTE Confidence: 0.851321227777778

00:46:40.502 --> 00:46:42.058 to treatment effect variation,

NOTE Confidence: 0.851321227777778

00:46:42.060 --> 00:46:44.588 then you can kind of think well what.

NOTE Confidence: 0.851321227777778

00:46:44.590 --> 00:46:45.922 Affects implementation and so

NOTE Confidence: 0.851321227777778

00:46:45.922 --> 00:46:47.920 people can sometimes think a little

NOTE Confidence: 0.851321227777778

00:46:47.978 --> 00:46:49.743 more carefully about what affects

NOTE Confidence: 0.851321227777778

00:46:49.743 --> 00:46:51.155 implementation like Oh well,

NOTE Confidence: 0.851321227777778

00:46:51.160 --> 00:46:51.902 it's probably.

NOTE Confidence: 0.851321227777778

00:46:51.902 --> 00:46:54.499 You know, it's probably easier to implement

NOTE Confidence: 0.851321227777778

00:46:54.499 --> 00:46:56.696 this in schools that are like this.

NOTE Confidence: 0.851321227777778

00:46:56.700 --> 00:47:00.536 Then schools that are not like that.

NOTE Confidence: 0.851321227777778

00:47:00.540 --> 00:47:02.658 You might try to find various

NOTE Confidence: 0.851321227777778

00:47:02.658 --> 00:47:04.360 measures of this for the

NOTE Confidence: 0.8922803

00:47:04.360 --> 00:47:07.768 implementation that sounds more like a.

NOTE Confidence: 0.8922803

00:47:07.770 --> 00:47:09.674 It's sort of a version of multiple

NOTE Confidence: 0.8922803

00:47:09.674 --> 00:47:11.305 treatments, and it's a violation of

NOTE Confidence: 0.8922803

00:47:11.305 --> 00:47:12.392 the suitable condition, probably.

NOTE Confidence: 0.8922803

00:47:12.392 --> 00:47:13.480 Yeah, yeah, exactly yeah.

NOTE Confidence: 0.85544264

00:47:13.480 --> 00:47:15.384 So I mean so it it gets.

NOTE Confidence: 0.85544264

00:47:15.390 --> 00:47:17.016 It gets tenuous. Yeah, I don't.

NOTE Confidence: 0.85544264

00:47:17.020 --> 00:47:18.916 I don't have this is, you know,

NOTE Confidence: 0.85544264

00:47:18.916 --> 00:47:21.060 this is like I when I first started

NOTE Confidence: 0.85544264

00:47:21.121 --> 00:47:22.729 doing this work I was like,

NOTE Confidence: 0.85544264

00:47:22.730 --> 00:47:24.606 well assuming moderate yrs and assume a

NOTE Confidence: 0.85544264

00:47:24.606 --> 00:47:26.270 population moving on as a statistician.

NOTE Confidence: 0.85544264

00:47:26.270 --> 00:47:27.896 But actually those are the two

NOTE Confidence: 0.85544264

00:47:27.896 --> 00:47:29.301 hardest things when working with

NOTE Confidence: 0.85544264

00:47:29.301 --> 00:47:30.676 people in planning these trials

NOTE Confidence: 0.85544264

00:47:30.676 --> 00:47:32.250 is thinking about what they are.

NOTE Confidence: 0.85544264

00:47:32.250 --> 00:47:33.876 I'll give you an example though.

NOTE Confidence: 0.85544264

00:47:33.880 --> 00:47:36.645 Uh, like a positive case which was.

NOTE Confidence: 0.85544264

00:47:36.650 --> 00:47:39.086 I was part of designing something called

NOTE Confidence: 0.85544264

00:47:39.086 --> 00:47:41.430 the National Study of learning mindsets,

NOTE Confidence: 0.85544264

00:47:41.430 --> 00:47:43.310 which is we randomly sampled

NOTE Confidence: 0.85544264

00:47:43.310 --> 00:47:45.480 100 high schools in the US,

NOTE Confidence: 0.85544264

00:47:45.480 --> 00:47:48.424 and then we randomly and then the students.

NOTE Confidence: 0.85544264

00:47:48.430 --> 00:47:49.036 There were.

NOTE Confidence: 0.85544264

00:47:49.036 --> 00:47:50.854 Ninth graders were in the study

NOTE Confidence: 0.85544264

00:47:50.854 --> 00:47:53.004 and so 9th graders were randomly

NOTE Confidence: 0.85544264

00:47:53.004 --> 00:47:55.224 assigned to either using a computer

NOTE Confidence: 0.85544264

00:47:55.285 --> 00:47:57.375 based intervention to a growth

NOTE Confidence: 0.85544264

00:47:57.375 --> 00:47:59.047 mindset intervention or something

NOTE Confidence: 0.85544264

00:47:59.047 --> 00:48:02.152 that was not growth mindset that was

NOTE Confidence: 0.85544264

00:48:02.152 --> 00:48:05.240 just sort of control condition and.

NOTE Confidence: 0.85544264

00:48:05.240 --> 00:48:07.464 And in doing that we had the social

NOTE Confidence: 0.85544264

00:48:07.464 --> 00:48:09.124 psychologist I was working with had

NOTE Confidence: 0.85544264

00:48:09.124 --> 00:48:11.353 a lot of questions like we had a

NOTE Confidence: 0.85544264

00:48:11.353 --> 00:48:13.153 lot of hard questions about these

NOTE Confidence: 0.85544264

00:48:13.153 --> 00:48:15.195 moderators and they had a lot of

NOTE Confidence: 0.85544264

00:48:15.195 --> 00:48:17.169 theories about what they might be like.

NOTE Confidence: 0.85544264

00:48:17.170 --> 00:48:19.170 So we oversampled like we.

NOTE Confidence: 0.85544264

00:48:19.170 --> 00:48:21.782 Looked at for example,

NOTE Confidence: 0.85544264

00:48:21.782 --> 00:48:25.047 proportion of students that are

NOTE Confidence: 0.85544264

00:48:25.047 --> 00:48:28.117 minorities in the school and then.

NOTE Confidence: 0.85544264

00:48:28.120 --> 00:48:30.066 And when we started we wanted to

NOTE Confidence: 0.85544264

00:48:30.066 --> 00:48:31.897 stratify on that as well as school

NOTE Confidence: 0.85544264

00:48:31.897 --> 00:48:33.738 at a measure of sort of school

NOTE Confidence: 0.85544264

00:48:33.738 --> 00:48:34.890 achievement as well,

NOTE Confidence: 0.85544264

00:48:34.890 --> 00:48:38.066 and so we needed to be able to

NOTE Confidence: 0.85544264

00:48:38.066 --> 00:48:40.948 cross these in a way in order to.

NOTE Confidence: 0.85544264

00:48:40.950 --> 00:48:43.353 In order to D alias these trends and so

NOTE Confidence: 0.85544264

00:48:43.353 --> 00:48:45.630 that they could estimate one without,

NOTE Confidence: 0.85544264

00:48:45.630 --> 00:48:46.254 you know,

NOTE Confidence: 0.85544264

00:48:46.254 --> 00:48:47.502 without estimating with separated

NOTE Confidence: 0.85544264

00:48:47.502 --> 00:48:50.000 from the other. So a lot of it.

NOTE Confidence: 0.85544264

00:48:50.000 --> 00:48:50.876 So I mean,

NOTE Confidence: 0.85544264

00:48:50.876 --> 00:48:53.430 in some places people are much more better,

NOTE Confidence: 0.85544264

00:48:53.430 --> 00:48:54.363 much better theoretically.

NOTE Confidence: 0.85544264

00:48:54.363 --> 00:48:55.296 Thinking about this,

NOTE Confidence: 0.85544264

00:48:55.300 --> 00:48:57.309 I think some fields are better at

NOTE Confidence: 0.85544264

00:48:57.309 --> 00:48:58.927 thinking about these mechanisms than

NOTE Confidence: 0.85544264

00:48:58.927 --> 00:49:01.540 other fields are, but yeah, it's really hard.

NOTE Confidence: 0.85544264

00:49:01.540 --> 00:49:02.401 It's really hard.

NOTE Confidence: 0.85544264

00:49:02.401 --> 00:49:04.410 So my my other other than my

NOTE Confidence: 0.85544264

00:49:04.476 --> 00:49:06.222 like standard set, you know,

NOTE Confidence: 0.85544264

00:49:06.222 --> 00:49:07.777 race, class and gender is,

NOTE Confidence: 0.85544264

00:49:07.780 --> 00:49:11.028 I often ask people to to think about.

NOTE Confidence: 0.85544264

00:49:11.030 --> 00:49:12.734 Watch what variables might just be

NOTE Confidence: 0.85544264

00:49:12.734 --> 00:49:14.154 related to other things, right?

NOTE Confidence: 0.85544264

00:49:14.154 --> 00:49:15.290 That if you could.

NOTE Confidence: 0.85544264

00:49:15.290 --> 00:49:17.486 If you can think of it as like I

NOTE Confidence: 0.85544264

00:49:17.486 --> 00:49:19.060 ultimately want to test moderators

NOTE Confidence: 0.85544264

00:49:19.060 --> 00:49:20.968 that I don't really know exactly

NOTE Confidence: 0.85544264

00:49:21.025 --> 00:49:21.820 what they are,

NOTE Confidence: 0.85544264

00:49:21.820 --> 00:49:24.084 but I need to get variation in them,

NOTE Confidence: 0.85544264

00:49:24.090 --> 00:49:25.794 and that means probably by getting

NOTE Confidence: 0.85544264

00:49:25.794 --> 00:49:26.930 variation in something else.

NOTE Confidence: 0.85544264

00:49:26.930 --> 00:49:29.486 I'm going to get variation in those as well,

NOTE Confidence: 0.85544264

00:49:29.490 --> 00:49:31.150 so.

NOTE Confidence: 0.85544264

00:49:31.150 --> 00:49:33.472 The size of your site, you know,

NOTE Confidence: 0.85544264

00:49:33.472 --> 00:49:34.094 I think,

NOTE Confidence: 0.85544264

00:49:34.094 --> 00:49:36.910 is one place where you know an education.

NOTE Confidence: 0.85544264

00:49:36.910 --> 00:49:38.938 You can see that everybody's in

NOTE Confidence: 0.85544264

00:49:38.938 --> 00:49:39.952 very large sites.

NOTE Confidence: 0.85544264

00:49:39.960 --> 00:49:42.680 And So what if we increase the variation?

NOTE Confidence: 0.85544264

00:49:42.680 --> 00:49:44.370 The variation of district size

NOTE Confidence: 0.85544264

00:49:44.370 --> 00:49:45.384 and school size?

NOTE Confidence: 0.85544264

00:49:45.390 --> 00:49:47.442 It seems like has to increase

NOTE Confidence: 0.85544264

00:49:47.442 --> 00:49:49.459 variation of some other things as

NOTE Confidence: 0.83546096

00:49:49.460 --> 00:49:50.489 well. Agreed, agreed.

NOTE Confidence: 0.83546096

00:49:50.489 --> 00:49:53.301 Yeah, I think another aspect why I so

NOTE Confidence: 0.83546096

00:49:53.301 --> 00:49:55.395 appreciate like the aspect of effect

NOTE Confidence: 0.83546096

00:49:55.395 --> 00:49:58.377 modifiers is that it really is a way to

NOTE Confidence: 0.83546096

00:49:58.377 --> 00:50:00.326 move forward with information from Co.

NOTE Confidence: 0.83546096

00:50:00.326 --> 00:50:02.414 Buryats and then when we talk

NOTE Confidence: 0.83546096

00:50:02.414 --> 00:50:04.940 on 80 in a randomized study,

NOTE Confidence: 0.83546096

00:50:04.940 --> 00:50:06.425 we often ignore covariates and

NOTE Confidence: 0.83546096

00:50:06.425 --> 00:50:08.502 then just hold that the unadjusted

NOTE Confidence: 0.83546096

00:50:08.502 --> 00:50:10.618 analysis provides unbiased estimates,

NOTE Confidence: 0.83546096

00:50:10.620 --> 00:50:12.888 even though that may come with

NOTE Confidence: 0.83546096

00:50:12.888 --> 00:50:14.022 a larger variation.

NOTE Confidence: 0.83546096

00:50:14.030 --> 00:50:16.690 So by really talking about effect modifiers,

NOTE Confidence: 0.83546096

00:50:16.690 --> 00:50:18.580 we somehow incurve those information,

NOTE Confidence: 0.83546096

00:50:18.580 --> 00:50:20.854 but perhaps even in the estimation

NOTE Confidence: 0.83546096

00:50:20.854 --> 00:50:22.370 of the average affect,

NOTE Confidence: 0.83546096

00:50:22.370 --> 00:50:23.890 which can increase precision.

NOTE Confidence: 0.83546096

00:50:23.890 --> 00:50:24.650 So yeah.

NOTE Confidence: 0.85203755

00:50:26.800 --> 00:50:27.070 Yeah.

NOTE Confidence: 0.79443485

00:50:30.360 --> 00:50:32.112 Yeah I haven't questioned,

NOTE Confidence: 0.79443485

00:50:32.112 --> 00:50:34.740 so I actually have two questions,

NOTE Confidence: 0.79443485

00:50:34.740 --> 00:50:36.930 so seems you're you're interested

NOTE Confidence: 0.79443485

00:50:36.930 --> 00:50:38.678 in both individual level

NOTE Confidence: 0.79443485

00:50:38.680 --> 00:50:41.310 an cluster level moderators right? When

NOTE Confidence: 0.79443485

00:50:41.310 --> 00:50:43.500 you have cluster level moderators,

NOTE Confidence: 0.79443485

00:50:43.500 --> 00:50:45.690 how does that work with

NOTE Confidence: 0.79443485

00:50:45.690 --> 00:50:47.002 the augmentation design?

NOTE Confidence: 0.79443485

00:50:47.002 --> 00:50:48.754 'cause you mentioned that

NOTE Confidence: 0.79443485

00:50:48.754 --> 00:50:50.506 in the orientation design,

NOTE Confidence: 0.79443485

00:50:50.510 --> 00:50:54.158 you might want to pick like.

NOTE Confidence: 0.79443485

00:50:54.160 --> 00:50:56.290 10 or 30% of the sides.

NOTE Confidence: 0.79443485

00:50:56.290 --> 00:50:58.065 An kind of like choose

NOTE Confidence: 0.79443485

00:50:58.065 --> 00:50:59.486 them samples from those.

NOTE Confidence: 0.79443485

00:50:59.486 --> 00:51:01.922 But how do you choose those 3%?

NOTE Confidence: 0.79443485

00:51:01.922 --> 00:51:03.908 You choose those third percent with

NOTE Confidence: 0.79443485

00:51:03.908 --> 00:51:06.228 respect to the cluster level modelers.

NOTE Confidence: 0.8195881

00:51:06.230 --> 00:51:09.070 You could do it with respect to either.

NOTE Confidence: 0.8195881

00:51:09.070 --> 00:51:11.660 You can do it with respect there

NOTE Confidence: 0.8195881

00:51:11.660 --> 00:51:13.834 because it depends the way you

NOTE Confidence: 0.8195881

00:51:13.834 --> 00:51:15.814 enter them into the model so.

NOTE Confidence: 0.8542276

00:51:17.850 --> 00:51:20.390 You work out so you can work out that if

NOTE Confidence: 0.8542276

00:51:20.453 --> 00:51:23.120 if I'm interested in the individual level.

NOTE Confidence: 0.8542276

00:51:23.120 --> 00:51:25.122 Moderate are that that what I need

NOTE Confidence: 0.8542276

00:51:25.122 --> 00:51:27.445 to do is I need the I actually

NOTE Confidence: 0.8542276

00:51:27.445 --> 00:51:29.624 need to include as a covariate the

NOTE Confidence: 0.8542276

00:51:29.624 --> 00:51:32.420 interaction between like X an 1 -- X.

NOTE Confidence: 0.8542276

00:51:32.420 --> 00:51:33.970 That's what I included here

NOTE Confidence: 0.8542276

00:51:33.970 --> 00:51:34.900 as the covariates.

NOTE Confidence: 0.8542276

00:51:34.900 --> 00:51:37.021 I'm 'cause I want to increase the

NOTE Confidence: 0.8542276

00:51:37.021 --> 00:51:38.310 variation within sites right?

NOTE Confidence: 0.8542276

00:51:38.310 --> 00:51:40.790 And so you could do it either way,

NOTE Confidence: 0.8542276

00:51:40.790 --> 00:51:42.030 because what it's doing,

NOTE Confidence: 0.8542276

00:51:42.030 --> 00:51:43.580 what the augmentation approach does?

NOTE Confidence: 0.8542276

00:51:43.580 --> 00:51:45.883 Is it assess is how much variation

NOTE Confidence: 0.8542276

00:51:45.883 --> 00:51:48.599 you have in those 30 sites already.

NOTE Confidence: 0.8542276

00:51:48.600 --> 00:51:52.128 And then it looks for possible design runs,

NOTE Confidence: 0.8542276

00:51:52.130 --> 00:51:53.843 meaning other samples.

NOTE Confidence: 0.8542276

00:51:53.843 --> 00:51:57.840 Other places that would greatly improve that.

NOTE Confidence: 0.8542276

00:51:57.840 --> 00:51:59.358 And it just it doesn't algorithmically,

NOTE Confidence: 0.8542276

00:51:59.360 --> 00:51:59.999 which is nice.

NOTE Confidence: 0.8542276

00:51:59.999 --> 00:52:02.026 The that I would say I should add an

NOTE Confidence: 0.8542276

00:52:02.026 --> 00:52:03.478 extra benefit of this is concerned

NOTE Confidence: 0.8542276

00:52:03.478 --> 00:52:05.219 with all of this sample recruitment

NOTE Confidence: 0.8542276

00:52:05.219 --> 00:52:06.724 is that there's non response.

NOTE Confidence: 0.8542276

00:52:06.730 --> 00:52:08.000 You're never going to get,

NOTE Confidence: 0.8542276

00:52:08.000 --> 00:52:10.540 you know it's not like I can just say like.

NOTE Confidence: 0.8542276

00:52:10.540 --> 00:52:11.046 Here's your.

NOTE Confidence: 0.8542276

00:52:11.046 --> 00:52:12.817 Here's your like 40 sites go ask

NOTE Confidence: 0.8542276

00:52:12.817 --> 00:52:14.598 them and they're going to say yes,

NOTE Confidence: 0.8542276

00:52:14.600 --> 00:52:16.310 but with the augmentation approach if

NOTE Confidence: 0.8542276

00:52:16.310 --> 00:52:18.149 somebody says no you can like throw

NOTE Confidence: 0.8542276

00:52:18.149 --> 00:52:19.963 that out and then go look for it

NOTE Confidence: 0.8542276

00:52:19.963 --> 00:52:21.463 like what's the next best alternative

NOTE Confidence: 0.8542276

00:52:21.463 --> 00:52:24.076 so you can keep kind of iterating.

NOTE Confidence: 0.8542276

00:52:24.076 --> 00:52:24.540 So,

NOTE Confidence: 0.85717666

00:52:24.540 --> 00:52:26.170 so in our current application,

NOTE Confidence: 0.85717666

00:52:26.170 --> 00:52:30.114 I think the attributes are all cluster level.

NOTE Confidence: 0.85717666

00:52:30.120 --> 00:52:31.332 Information right summary statistics

NOTE Confidence: 0.85717666

00:52:31.332 --> 00:52:32.849 yeah yeah, well that's what

NOTE Confidence: 0.85039884

00:52:32.850 --> 00:52:34.042 I have right here.

NOTE Confidence: 0.85039884

00:52:34.042 --> 00:52:36.207 That's in the in the slides but

NOTE Confidence: 0.85039884

00:52:36.207 --> 00:52:38.037 I didn't include in here though

NOTE Confidence: 0.85039884

00:52:38.037 --> 00:52:40.186 is like you could it's but it's

NOTE Confidence: 0.85039884

00:52:40.186 --> 00:52:42.544 in the paper is you could also do

NOTE Confidence: 0.85039884

00:52:42.544 --> 00:52:44.054 this with individual level only.

NOTE Confidence: 0.85039884

00:52:44.060 --> 00:52:46.058 For proportions mean because just because

NOTE Confidence: 0.85039884

00:52:46.058 --> 00:52:48.298 the proportion works out that you can get.

NOTE Confidence: 0.85039884

00:52:48.300 --> 00:52:50.112 You can think about this with

NOTE Confidence: 0.85039884

00:52:50.112 --> 00:52:51.675 the same statistics you would

NOTE Confidence: 0.85039884

00:52:51.675 --> 00:52:53.145 get at the cluster level.

NOTE Confidence: 0.85039884

00:52:53.150 --> 00:52:55.302 You can't get the variation you can with

NOTE Confidence: 0.85039884

00:52:55.302 --> 00:52:57.387 a normal like a continuous variable.

NOTE Confidence: 0.85039884

00:52:57.390 --> 00:52:58.642 I can't get the.

NOTE Confidence: 0.85039884

00:52:58.642 --> 00:53:00.520 I don't have the standard deviation.

NOTE Confidence: 0.85039884

00:53:00.520 --> 00:53:02.510 Insights I can't do that,

NOTE Confidence: 0.85039884

00:53:02.510 --> 00:53:04.850 right, right?

NOTE Confidence: 0.85039884

00:53:04.850 --> 00:53:06.150 Also, the other question

NOTE Confidence: 0.85039884

00:53:06.150 --> 00:53:07.778 is about so it seems

NOTE Confidence: 0.8673171

00:53:07.780 --> 00:53:09.400 like all these designs are

NOTE Confidence: 0.8673171

00:53:09.400 --> 00:53:11.028 under the assumption that you're

NOTE Confidence: 0.8673171

00:53:11.030 --> 00:53:12.650 interested in all the moderators,

NOTE Confidence: 0.8673171

00:53:12.650 --> 00:53:14.600 like equally like meaning that you're

NOTE Confidence: 0.8673171

00:53:14.600 --> 00:53:16.876 not like you don't have like primary

NOTE Confidence: 0.8673171

00:53:16.876 --> 00:53:18.505 moderators that you're interested in

NOTE Confidence: 0.8673171

00:53:18.505 --> 00:53:20.454 estimating the moderate effect on and

NOTE Confidence: 0.8673171

00:53:20.454 --> 00:53:23.046 then you have a couple of them that.

NOTE Confidence: 0.8673171

00:53:23.050 --> 00:53:24.680 I mean, if you you

NOTE Confidence: 0.8673171

00:53:24.680 --> 00:53:26.630 can. So I mean, what's great?

NOTE Confidence: 0.8673171

00:53:26.630 --> 00:53:29.177 I mean, I think about this like this area

NOTE Confidence: 0.8673171

00:53:29.177 --> 00:53:31.564 is that it's been so richly developed

NOTE Confidence: 0.8673171

00:53:31.564 --> 00:53:34.038 in this other sort of design runway

NOTE Confidence: 0.8673171

00:53:34.038 --> 00:53:36.467 is that you can actually add weights.

NOTE Confidence: 0.8673171

00:53:36.470 --> 00:53:38.621 So you can say like I'm more like or

NOTE Confidence: 0.8673171

00:53:38.621 --> 00:53:40.884 more interested in this variable than

NOTE Confidence: 0.8673171

00:53:40.884 --> 00:53:42.942 that variable, and it will focus.

NOTE Confidence: 0.8673171

00:53:42.942 --> 00:53:45.330 You know it will focus on one

NOTE Confidence: 0.8673171

00:53:45.330 --> 00:53:46.950 variable over the other.

NOTE Confidence: 0.8673171

00:53:46.950 --> 00:53:48.696 Because you Can you imagine like

NOTE Confidence: 0.8673171

00:53:48.696 --> 00:53:50.439 that ask like that D matrix.

NOTE Confidence: 0.8673171

00:53:50.440 --> 00:53:51.588 The determinant of S.

NOTE Confidence: 0.8673171

00:53:51.588 --> 00:53:53.640 You could just add weights into that.

NOTE Confidence: 0.8673171

00:53:53.640 --> 00:53:55.558 So if you add weights into that

NOTE Confidence: 0.8673171

00:53:55.558 --> 00:53:58.049 then you can start looking at the

NOTE Confidence: 0.8673171

00:53:58.049 --> 00:54:00.054 determinant of that weighted version.

NOTE Confidence: 0.8673171

00:54:00.060 --> 00:54:01.758 Right, so you would add weights

NOTE Confidence: 0.8673171

00:54:01.758 --> 00:54:03.456 in that matrix and optimize that.

NOTE Confidence: 0.8673171

00:54:03.456 --> 00:54:04.020 Yeah exactly,

NOTE Confidence: 0.8673171

00:54:04.020 --> 00:54:05.679 if you add weight so that some

NOTE Confidence: 0.8673171

00:54:05.679 --> 00:54:07.470 of the Kobe rates are getting

NOTE Confidence: 0.8673171

00:54:07.470 --> 00:54:08.826 more weight than others.

NOTE Confidence: 0.8543322

00:54:10.400 --> 00:54:13.179 So I guess just maybe more precisely,

NOTE Confidence: 0.8543322

00:54:13.180 --> 00:54:15.556 I think the D optimality criteria.

NOTE Confidence: 0.8543322

00:54:15.560 --> 00:54:17.545 Shouldn't that be the X

NOTE Confidence: 0.8543322

00:54:17.545 --> 00:54:19.530 transpose V universe in general?

NOTE Confidence: 0.8543322

00:54:19.530 --> 00:54:21.515 Just because you're working with

NOTE Confidence: 0.8543322

00:54:21.515 --> 00:54:23.103 clustered randomized studies so

NOTE Confidence: 0.8543322

00:54:23.103 --> 00:54:25.119 that the outcome correlation is

NOTE Confidence: 0.8543322

00:54:25.119 --> 00:54:27.069 somehow included in that variance?

NOTE Confidence: 0.8543322

00:54:27.070 --> 00:54:29.296 Is that what the algorithm is

NOTE Confidence: 0.8543322

00:54:29.296 --> 00:54:31.840 trying to get in general for?

NOTE Confidence: 0.8543322

00:54:31.840 --> 00:54:34.250 Yeah yeah.

NOTE Confidence: 0.81006

00:54:34.250 --> 00:54:36.690 Inverse, yeah, it's the X prime X inverse,

NOTE Confidence: 0.81006

00:54:36.690 --> 00:54:38.220 which is the covariance. Yeah,

NOTE Confidence: 0.81006

00:54:38.220 --> 00:54:40.660 but but really not so it you don't.

NOTE Confidence: 0.81006

00:54:40.660 --> 00:54:42.358 You don't need to have the

NOTE Confidence: 0.81006

00:54:42.358 --> 00:54:44.010 variance matrix of the outcome.

NOTE Confidence: 0.8557798

00:54:45.570 --> 00:54:47.930 Exactly, you don't need to have the outcome,

NOTE Confidence: 0.8557798

00:54:47.930 --> 00:54:49.689 it's all about the inputs, right?

NOTE Confidence: 0.8557798

00:54:49.689 --> 00:54:51.363 But that's, which is why you

NOTE Confidence: 0.8557798

00:54:51.363 --> 00:54:53.244 can do it in advance, right?

NOTE Confidence: 0.8557798

00:54:53.244 --> 00:54:55.596 So it's all about the Android just nicely.

NOTE Confidence: 0.8557798

00:54:55.600 --> 00:54:56.782 You can leverage population

NOTE Confidence: 0.8557798

00:54:56.782 --> 00:54:58.258 data that you have totally.

NOTE Confidence: 0.8557798

00:54:58.258 --> 00:55:00.722 And again, I assume in all of this

NOTE Confidence: 0.8557798

00:55:00.722 --> 00:55:02.295 that like there's measurement error

NOTE Confidence: 0.8557798

00:55:02.295 --> 00:55:04.108 and that you know you can just

NOTE Confidence: 0.8557798

00:55:04.108 --> 00:55:05.891 sort of assume that like you're

NOTE Confidence: 0.8557798

00:55:05.891 --> 00:55:08.025 not going to get it exactly right,

NOTE Confidence: 0.8557798

00:55:08.025 --> 00:55:09.975 but my baseline comparison is always

NOTE Confidence: 0.8557798

00:55:09.975 --> 00:55:12.408 what are we doing now versus what could

NOTE Confidence: 0.8557798

00:55:12.408 --> 00:55:14.869 we be doing an like frankly anything.

NOTE Confidence: 0.8557798

00:55:14.870 --> 00:55:17.093 Any you know it looks to me like we

NOTE Confidence: 0.8557798

00:55:17.093 --> 00:55:18.836 have fairly homogeneous samples and

NOTE Confidence: 0.8557798

00:55:18.836 --> 00:55:21.915 that any effort we can make to increase

NOTE Confidence: 0.8557798

00:55:21.915 --> 00:55:24.075 that heterogeneity is an improvement.

NOTE Confidence: 0.88261515

00:55:28.050 --> 00:55:30.786 So, well, I think we're about the hour,

NOTE Confidence: 0.88261515

00:55:30.790 --> 00:55:33.016 but let's see if we have any

NOTE Confidence: 0.88261515

00:55:33.016 --> 00:55:34.890 final questions from the audience.

NOTE Confidence: 0.80667937

00:55:40.450 --> 00:55:42.858 Alrighty, if not, I think you know I'm,

NOTE Confidence: 0.80667937

00:55:42.860 --> 00:55:44.995 I'm sure if you have any questions

NOTE Confidence: 0.80667937

00:55:44.995 --> 00:55:46.838 that petition will we have to

NOTE Confidence: 0.80667937

00:55:46.838 --> 00:55:48.278 answer them offline by email?

NOTE Confidence: 0.80667937

00:55:48.280 --> 00:55:50.065 So thanks so much. Again, bath.

NOTE Confidence: 0.80667937

00:55:50.065 --> 00:55:52.630 It's really nice to have you and thanks to

NOTE Confidence: 0.80667937

00:55:52.698 --> 00:55:55.194 everybody for attending or see all of you.

NOTE Confidence: 0.80667937

00:55:55.200 --> 00:55:58.056 Hopefully after the break so have

NOTE Confidence: 0.80667937

00:55:58.056 --> 00:55:59.958 a great holiday. See you later.

NOTE Confidence: 0.77599597

00:56:02.150 --> 00:56:03.646 Totally not master connect

NOTE Confidence: 0.77599597

00:56:03.646 --> 00:56:04.762 alright, thanks again.

NOTE Confidence: 0.77599597

00:56:04.762 --> 00:56:08.917 Talk to you later. Bye take care.