Exhibit 11

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
OAKLAND DIVISION

THE APPLE iPOD iTUNES ANTI-TRUST LITIGATION

Lead Case No. C 05-00037

This Document Relates To:
ALL ACTIONS

CONFIDENTIAL - ATTORNEYS' EYES ONLY
VIDEOTAPED DEPOSITION OF JEFFREY M. WOOLDRIDGE, Ph.D.
Monday, January 6, 2014
San Diego, California

Reported By:
Debby M. Gladish
RPR, CLR, CCRR, CSR No. 9803
NCRA Realtime Systems Administrator

Job No. 10009202

UNITED STATES DISTRICT COURT
NORTHERN DISTRICT OF CALIFORNIA
OAKLAND DIVISION

THE APPLE iPOD iTUNES ANTI-TRUST LITIGATION

Lead Case No. C 05-00037

This Document Relates To:
ALL ACTIONS

Videotaped Deposition of JEFFREY M. WOOLDRIDGE, Ph.D., taken on behalf of the Defendant at 655 West Broadway, Suite 1900, San Diego, California, beginning at 10:29 a.m. and ending at 4:26 p.m., on Monday, January 6, 2014, before Debby M. Gladish, RPR, CLR, CCRR, CSR No. 9803, NCRA Realtime Systems Administrator.

APPEARANCES

For the Plaintiffs:
ROBBINS GELLER RUDMAN & DOWD, LLP
BY: BONNY SWEENEY, ESQ.
BY: JENNIFER N. CARINGAL, ESQ.
655 West Broadway, Suite 1900
San Diego, California 92101
(619)231-1058
bonnys@rgrdlaw.com

For the Defendant, Apple Inc.:
JONES DAY
BY: DAVID KIERNAN, ESQ.
BY: AMIR AMIRI, ESQ.
555 California Street, 26th Floor
San Francisco, California 94104
(415)626-3939
aamiri@jonesday.com

APPLE
BY: SCOTT B. MURRAY, ESQ. (TELEPHONIC APPEARANCE)
1 Infinite Loop, MS 169-2NYJ
Cupertino, California 95014
(408)783-8369
scott_murray@apple.com

Also present:

CHRISTOPHER TISA (VIDEO OPERATOR)
APTUS COURT REPORTING
600 West Broadway
Suite 300
San Diego, California 92101
619.546.9151

INDEX

E X H I B I T S

MARKED		PAGE
Exhibit 6	Document titled "Imbens/Wooldridge, Lecture Notes 2, Summer '07, What's New in Economics?" 32 pages	145
Exhibit 7	Document titled "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results," Justin Wolfers, Working Paper 10014, 30 pages	148
Exhibit 8	Document titled "NBER Working Paper Series, What Are We Weighting For?" 29 pages	149

SAN DIEGO, CALIFORNIA

MONDAY, JANUARY 6, 2014, 10:29 a.m.

THE VIDEOGRAPHER: Good morning. We are now on the record. The time is 10:29 a.m. Today's date is January 6, 2014.

My name is Christopher Tisa of Aptus Court Reporting. The court reporter is Debby Gladish with Aptus Court Reporting located at 600 West Broadway, Suite 300, San Diego, California 92101.

This begins the video-recorded deposition of Jeffrey M. Wooldridge, testifying in the matter of the Apple iPod iTunes Anti-Trust Litigation, pending the United States District Court, Northern District of California, Oakland Division, Case Number C 05-00037 YGR, taken at 655 West Broadway, Suite 1900, San Diego, California 92101.

The video and audio recording will take place at all times during this deposition unless all counsel agree to go off the record. The beginning and end of each video recording will be announced.

Will counsel please identify yourselves and state whom you represent.

MR. KIERNAN: David Kiernan on behalf of Apple.

MR. AMIRI: Amir Amiri on behalf of Apple.

MS. SWEENEY: Bonny Sweeney on behalf of the plaintiffs.

MS. CARINGAL: Jennifer Caringal on behalf of the plaintiffs.

THE VIDEOGRAPHER: The court reporter may now swear in the deponent.

JEFFREY M. WOOLDRIDGE,

having been sworn, testified as follows:

THE VIDEOGRAPHER: You may proceed, Counsel.

MR. KIERNAN: Okay.

EXAMINATION

BY MR. KIERNAN:

Q. Good morning, Dr. Wooldridge.

A. Good morning.

Q. Could you state your full name for the record.

A. Jeffrey M. Wooldridge.

Q. Okay. Have you ever been deposed before?

A. No.

Q. Okay. Have you ever testified before?

A. No.

Q. Okay. Do you understand that your testimony today is under oath subject to penalty of perjury?

A. I do.

Q. Okay. Is there any reason that you cannot testify completely and truthfully today?

A. No.

Q. Any substance that you've taken that would impair your ability to testify completely and truthfully?

A. No.

Q. When were you first contacted to do work on this case?

A. December 5th I received an e-mail from Bonny Sweeney.

Q. December 5th, 2014 -- or 2013?

A. December 5th, 2013, yes.

Q. Okay. And what is your assignment in this case?

A. My assignment is to evaluate different claims about how the proper standard error should be computed in the Noll regression analysis --

Q. Okay.

A. -- and whether cluttering is important or not or I should say whether it's valid or not.

Q. And when did you start work in this matter? When did you start to do the work after being first

contacted on December 5th?
A. A week later, December 12th, 2013.
Q. And about how many hours have you put in on this case?
A. Up to writing the dec- -- submitting the declaration or after that as well?
Q. That's a good time. Up through submitting your declaration.
A. Five to six hours.
Q. Okay. And then, after submitting your declaration, how much time have you spent, if any?
A. Probably another ten hours.
Q. Putting aside conversations that you've had with counsel --
A. Yes.
Q. -- including Bonny or anyone else from Robbins Geller, have you discussed this case with anybody else?
A. No.
Q. Have you discussed the case with Dr. Noll?
A. No.
Q. Have you ever had a discussion with Dr. Noll at any time in your life?
A. No, I don't believe we've met.
Q. Okay. Did you have any support staff or any other person who assisted you?

A. No.
Q. Okay. It was just you?
A. Just me.
Q. All right. And how are you being compensated for your work in this matter?
A. Hourly wage?
Q. Uh-huh.
A. $500 an hour.
Q. And how much have you been paid?
A. Nothing.
Q. Okay. And so when you submit an invoice it's going to be between 15 and 16 hours plus whatever work your deposition today time?
A. Yes.
Q. Have you submitted any invoices or any other bill that reflects the hours spent and the amount you are owed?
A. I haven't submitted invoices yet.
Q. Okay. Since submitting your declaration, what work have you done?
A. I've done some simulation work on -- on the properties of clustered standard errors.
Q. Okay. Anything else?
A. And -- and working out some formulas that can explain the simulation findings.

Q. Okay. Anything else?
A. No.
Q. Did you prepare for the deposition?
A. Yes.
Q. And, just briefly, describe what you did to prepare for --
A. I read --
Q. And I don't want to hear any conversations that you had with counsel. You can tell me if you met with counsel, but I don't want to hear what you guys talked about.
A. We -- we did meet over the phone. I read the various reports, the Murphy, Topel report, the Noll rebuttal report, and I reviewed my own declaration.
Q. Okay. Anything else?
A. Reviewed some of my old work on clustering, but . . .
Q. Like old pub- -- publications?
A. Yes, and my book.
Q. And -- and which book, the graduate book or the undergrad book?
A. My graduate book, which -- which is published with MIT Press.
Q. And with respect to the other clustering work, aside from the textbook, do you recall which -- what the

titles were of those works?
A. The main -- the -- the one was the paper I published in the American Economic Review called the Cluster -- Cluster Sampling and Applied Econometrics.
Q. The 2003 paper?
A. Yes, uh-huh.
Q. With respect to the -- let me just see what that is.
MR. KIERNAN: Do I have to hit escape to go up?
THE REPORTER: Yes.
BY MR. KIERNAN:
Q. With respect to the simulation work on the properties of clustered standard errors, was the simulation work done on the standard errors in -- from Noll's regress- -- rebuttal regressions?
A. No. I set up a simplified framework so that the issues would be more transparent, showing what would happen if you took an independent sample and clustered after the fact based on some characteristics. I did that after I wrote my declaration.
Q. And why did you do that work?
A. Because in my declaration I asserted things that seemed self-evident, but thought it would be useful to actually see the -- the simulation findings that

actually showed what I was claiming.

Q. And are you relying on those simulations for any of the opinions that you're giving in this matter?

A. Not necessarily. I guess -- I didn't rely on them in my declaration and so I'll be talking about my -- the opinions in my declaration, which haven't changed.

Q. Okay. So you're not relying on the simulations that you've done after submitting your declaration as a basis for any --

A. No.

Q. -- of the opinions in your report?

A. No, I'm not.

Q. Okay. In -- in the simulations, what was the dataset that you used?

A. Well, the simulation generates data based on some assumptions about what the population distribution is and then draws randomly from using a standard program, such as Stata, to draw random samples from the population.

Q. And the -- when you're referring to the population, what's the dataset for that?

A. When --

Q. That's part of the simulation program?

A. Yes. So when you define a population you

define a distribution such as a normal distribution or a quasi distribution or something like that and then you randomly sample from that -- from a random variable that has that distribution. It's a very common method used to evaluate any kind of estimator that somebody proposes in econometrics or statistics.

MR. KIERNAN: How do I get this going again?

THE REPORTER: Hit the pause button.

MR. KIERNAN: Pause break? Say again?

THE REPORTER: Use your mouse and --

MR. KIERNAN: Oh, I see it.

THE REPORTER: And hit --

MR. KIERNAN: Got it, got it. Thank you.

BY MR. KIERNAN:

Q. I notice in the declaration you note that you've done some consulting work, like you worked for CRA and --

A. Yes.

Q. Okay. Have you done any work, provided any opinions, in antitrust cases or any antitrust matters?

A. Yes.

Q. Okay.

A. Back with the Charles River Associates work I did some econometric work at the request of Frank Fisher on the Kodak Polaroid patent infringement case. And at

one point I also did work -- there was a case where the NBA was suing, I believe, the -- the super station in Chicago, WGN, for showing Chicago Bulls games with Michael Jordan nationwide.

Q. Okay. Any others that you can think of?

A. I wish my memory were better. I -- I did do some more cases for Charles River Associates. There was a case having to do with airline reservation systems, I believe. And that's as much as I can remember.

Q. And have you ever been retained -- excuse me -- to estimate the impact of some conduct on the prices of consumer products?

A. No.

Q. Okay. Have you ever been retained to estimate damages resulting from alleged impact of conduct on consumer prices -- on consumer products?

A. No.

Q. You note in your declaration that you're currently providing consulting work to Industrial Economics, Inc. on a damage assessment.

A. Uh-huh.

Q. And describe that for me.

A. That's through the government, NOAA, for the deep water horizon oil spill.

Q. And what is the work that you're -- the

consulting work that you're doing in connection with that?

A. We're estimating damages from the oil spill based on consumer willingness to pay.

Q. And you said "we." Are there other people involved?

A. Yes.

Q. Okay.

A. It's -- it's a large team.

Q. And what is your role?

A. My role is mainly as the econometrician to think about sampling issues and model estimation issues and how to compute standard errors.

Q. And in that matter have you proposed a model to estimate damages?

A. Yes.

Q. And is it a regression model or --

A. It's a bit --

Q. Strike that.

Why don't you describe the model. I'll start . . .

A. I'm not sure I'm at liberty to do that. I -- I don't know what the protocol is, but I -- I am --

MS. SWEENEY: Yeah, if it's --

THE WITNESS: It -- it --

MS. SWEENEY: If it's confidential --
THE WITNESS: It's confidential.
MS. SWEENEY: I'll just object to form.
THE WITNESS: Sorry. That is confidential information.
BY MR. KIERNAN:
Q. Okay. And -- and just so I have it -- the record clear, even the type of model that you are or not using to estimate damages in the case, your testimony is you cannot describe it for me because of a protective order --
A. Yes.
Q. -- in that matter?
A. Yes. That's correct.
Q. Other than that matter, have you been retained to estimate or consult in estimating damages?
A. No.
MR. KIERNAN: All right. Let me have his report.
Can mark that as -- why don't we mark it as Wooldridge 1 because I don't think we've been doing them sequentially.
(Exhibit 1 marked.)
BY MR. KIERNAN:
Q. Okay. I'm handing you what's been marked as

Wooldridge 1. Can you identify as -- Wooldridge 1 as the declaration -- as your declaration that you submitted in this case?
A. Yes, it is.
Q. Okay. And is it -- does Wooldridge 1 contain all the opinions that you're offering in this matter?
A. Yes.
Q. And contains all the bases for those opinions?
A. Yes.
Q. And does it list all the materials that you relied on?
A. Yes.
Q. Okay. And did you draft Wooldridge 1?
A. Yes.
Q. Did anyone else assist you with drafting it?
A. Counsel read through and made small editorial comments.
Q. Did you review any deposition transcripts?
A. Yes, I did. I read the Noll report, both the initial report and the rebuttal, and the Murphy and Topel reports.
Q. Okay.
MS. SWEENEY: I -- I --
MR. KIERNAN: I'm going to -- I'll clarify.
BY MR. KIERNAN:

Q. What I mean by "deposition transcript" --
A. Oh, deposition --
Q. -- is --
A. Oh, I'm sorry. Not --
Q. Like today we have a deposition and then --
A. I did not read any deposition transcripts.
Q. Okay.
A. I'm sorry. Yes.
Q. And did you review the -- any of the data -- the datasets that Dr. Noll used in his regressions?
A. No, I didn't see the datasets.
Q. Okay. Did you review the documents that Dr. Noll cites in his reports?
A. Um --
MS. SWEENEY: Objection. Overbroad.
THE WITNESS: Did I -- I was familiar with some of the econometrics works that he cited, but I did not review -- he has previous dec- -- declarations listed there. I did not review those or any of the other -- there was a long list, I believe, of documents, and I did not look at them.
BY MR. KIERNAN:
Q. Okay.
A. I had a limited amount of time.
Q. And do you recall that Dr. Murphy and

Dr. Topel, they also listed a number of documents that they considered?
A. Yes.
Q. And did you review any of those?
A. No, I did not.
Q. Did you review a supplemental report that was jointly signed by Drs. Murphy and Dr. Topel?
A. If it's the -- the recent one --
Q. Yes.
A. -- yes, I did.
Q. Did you review the regression equations used by Dr. Noll?
MS. SWEENEY: Objection. Vague and ambiguous.
THE WITNESS: I -- I looked at the equations and the reported standard errors, but I did not evaluate the equations for content.
BY MR. KIERNAN:
Q. Okay. And did you evaluate the -- well, strike that.
On page 2 of Wooldridge 1 of your declaration, at the bottom, you state, "I restrict my comments to issues associated with computing proper standard errors and do not discuss model specification."
Do you see that?
A. Yes.

Q. And what do you mean by "model specification"?
A. Well, every regression analysis has a dependent variable, which you're trying to explain, and a set of explanatory variables sometimes called independent variables. And different people can have different opinions on what those variables should be. And I was not asked to evaluate that part of Dr. Noll's analysis and so I haven't formed an opinion. I didn't look at the equations with an eye toward did I think this was proper or not.
I was asked to do something fairly narrow, which was evaluate the clustering issue. And that's what I spent my limited time on.
Q. Okay. Okay. So you're not offering an opinion on the model specification?
A. That's correct.
Q. And not offering an opinion on whether he included the correct explanatory variables or what you called the independent variables?
A. That's correct.
Q. Not offering an opinion on whether the regression suffered from omitted-variable bias?
A. Correct.
Q. And no opinion on whether Dr. Noll's regressions estimate or provide -- or produce a reliable

damages estimate?
A. I would need to study what he did in much more depth to -- to -- to comment on that. And I wasn't asked to do that and I -- I formed no opinion on that.
Q. Fair enough. Fair enough. Okay. So you have not formed an opinion on whether Dr. Noll's regressions produces reliable damages estimates?
A. That's correct.
Q. And no opinion on whether the conduct at issue in this litigation impacted iPod prices?
A. Correct.
Q. No opinion on the amount of damages?
A. That's correct.
Q. What is your understanding about -- of what this case is about?
A. Oh, well, there were certain versions of iPods that were installed with software that essentially blocked a competitor's software that allowed downloading music from competing sites other than the iTunes store. But I -- I have to say I focused my attention on the cluster sampling issue. I understood what the data structure was and what the basic question was, and I didn't think in-depth about what the actual antitrust issue is here.
Q. Okay. And what is the basis for your

understanding of what the data structure used by Dr. Noll in his two regressions is?
A. He has transactions level data for shipments of various classes of iPods, along with information about the characteristics of the iPods and the prices at which the transactions occurred, in- -- including when they occurred.
Q. And you understand Dr. Noll has two regressions, one for resellers and then the other --
A. Yes. I -- I did take note of that, yes.
Q. Okay. And the data structure that you just described, is that true for both types of customers, sales to both types of customers or -- let me stop there.
A. The direct sales have that structure and, yeah, I -- I -- I didn't see their both transactions records in the direct sales. There's -- there's perhaps more than one unit sold.
Q. Okay. And I -- when you say shipments of iPods, what are you referring to?
A. Well, I call them transactions, I believe. But a shipment is -- for the -- the purposes of the data analysis what matters is that there's a transaction that happened for a certain kind of iPod on a certain day at a certain price and so it could have been one iPod or it

could have been many iPods.
Q. Okay. So ship- -- when you used the term "shipment" previously you're referring to transaction?
A. Yes.

[redacted]

Q. And in your declaration you state that Professor Noll is using the entire population of transactions. What do you mean by "the entire

Page 25

population"?
A. Well, my understanding is that Apple provided every -- every transaction over this ten- or 11-year period and except for a couple that were dropped due to missing data issues and, I believe, some outlining observations, he has every transaction ever done.
Alternatively, Apple could have said, "Here's a 10 percent random sample of our transaction," and then it would have been a random sample from that population, but instead he has all the transactions.
Q. And all the transactions worldwide or just in the United States?
A. I didn't read it that -- that closely. Sorry. So if --
Q. Okay.
A. -- if it's just in the United States, then it's the pop- -- then that defines the population.
Q. Okay. And so your understanding is that the transactional data for -- used for both regressions contain virtually every iPod sold in the U.S. during the time period?
A. Yes.
Q. If you go to page 13 --
A. Yes.
Q. -- does paragraph 6 list the summary of your

Page 26

opinions in the case?
A. Uh-huh.
Q. Okay. And with respect to the first, "Clustering is inappropriate where, as here, the regressions use the entire population of transactions," is it your opinion that clustering is inappropriate whenever the entire population is used?
A. It -- no. It -- it could be appropriate if, for example, you use information from other units, other transactions in the data as part of an explanatory variable in a transaction for a particular transaction. So -- but Professor Noll did not do that.
Each transaction was its own separate unit and each provides independent information on prices at which these transactions occurred, given the -- given the characteristics of the -- the different iPods and the different time periods when they were purchased.
Q. Going back to the circumstance that you described where clustering could be appropriate when using the entire population of transactions, you say it could be appropriate if, for example, you use information from other units other transactions in a transaction for a particular transaction.
A. I better clarify that.
Q. Please.

Page 27

A. Let me give you -- let me give you a different example. Suppose we had the entire population of students in a state and we wanted to estimate the effect of some intervention on student test scores, if we had that entire population, we would just use a standard regression analysis. If we wanted to do something like test whether there are peer effects, say, within the neighborhood or the school, and we included a variable that measured characteristics of students nearby the other -- the -- the student in question, then that could -- could create a clustering problem.
Q. Okay.
A. But my understanding is that Professor Noll did not do that.
Q. And using the data that's at issue in our case --
A. Uh-huh.
Q. -- and the products that are at issue that are being modeled, can you give me an example? You -- you used the school example. Can you give an example using the --
A. Frankly, I can't even --
MS. SWEENEY: Objection.
You've got to pause for a moment so I can interject my objection.

Page 28

Objection to form. Vague and ambiguous. Incomplete hypothetical.
BY MR. KIERNAN:
Q. Okay. Go ahead.
A. Frankly, I can't think of a reason you would do that. It -- it makes no sense to me to say that a transaction that happened someplace else, some other time period, would somehow have an effect on the price of this particular transaction when -- and -- and the point is Professor Noll didn't do it, so there's nothing to -- to be concerned about here.
Q. So, in your opinion, there's no circumstance under which clustering could be a problem with respect to the data that Dr. Noll used to estimate his regressions?
A. That's correct. Let's -- and let me -- let me expand on that a little bit.
I mentioned that he has, essentially, the entire population of transactions. He -- he could have or Apple could have given him a 10 percent random sample. There're easy ways to generate a random sample from a large population like that. And then he would, really, have had a random sample and the analysis would have clearly been not subject to a criticism of clustering because the -- the -- the observations would

Page 29

have been -- been drawn to represent the population and independently.

And so if you, then, say, well, what if he took an additional 10 percent of the sample, then he would have more data and that would be reflected in the standard errors falling because you're getting more data. And, again, because it's a random sample, there's no reason to cluster and clustering, in fact, only could inflate the standard errors in an artificial way. And by the time you get up to the population -- having the whole population is not a problem. That's a -- that's a good thing. You have more information. You want more data to more precisely estimate the coefficients in the regression model.

So that's why I assumed that a 10 percent random sample wasn't taken because it's better to use more data than -- than less data.

Q. Right. And so your opinion is that if you have the entire population of iPod transactions there's no circumstance under which clustering would be appropriate?

A. That's correct.

Q. And you state, "There can be no cluster sampling problem because there is no sampling."

If there were sampling --

Page 30

A. If --

Q. -- are there circumstances under which clustering would be appropriate dealing with iPod transactions?

A. With iPod transactions? I don't see how -- not -- not the way -- it -- these were sampled transaction by transaction and so there can't be a cluster sampling problem if that's the way the -- the data had been sampled.

Q. Is it your opinion that the resid- -- that the error terms in Professor Noll's regressions are independent?

A. They're not independent ex post after you choose the clusters, and it's very, very simple to see that. The clustering of a -- either the entire population in this case or a random sample create -- artificially creates a problem that isn't there. So the idea is -- and -- and let's take a simple example of this.

Suppose that we wanted to estimate -- and make this simple -- we want to estimate the average, let's say, hourly wage in the population of all hourly wage workers. The way we would do that -- and, of course, that's a -- that's a big population in the United States. And surveys do this, they go out and they

Page 31

collect a random sample of hourly wage workers and they compute the sample average.

The sample average is the simplist example of an OLS estimate, an ordinary least squares estimator. It minimizes the sum of squared deviation. So it is -- it is an example, essentially, of what Professor Noll is doing. Of course, regression analysis is a little more complicated, but -- but let's stick with that.

Suppose that -- so the proper thing to do would be to collect a random sample and you can look at any introductory statistics book and it will show you the formula for the standard error, which is the standard deviation you estimate from your sample divided by the square root of the sample size. But suppose that you -- along with the hourly wage you actually collected information on the person's occupation. So some people are in the service industry, some people are construction workers, some people are computer programmers.

Now, suppose that after you've computed the sample average, you then compute the residuals, which would be everybody's -- every person's hourly wage net of the total average across the entire sample. So what will you find if you do that? Well, if you compute the residuals within the -- the -- the now cluster of

Page 32

workers in the service industry, those, very likely, will be negative on average because service workers earn -- or, let's say, fast-food workers earn a lower hourly wage than the overall average.

If you go to the computer programmers, you're going to find that that residual is positive because on average they earn more than the total average in the population. In fact, the difference is simply -- the average residual is just the average of the hourly wage for computer programmers minus the overall average.

This is exactly what Murphy and Topel do. They, then, do this ex post clustering, and they -- they make a point of saying that the residuals are negative here, they're negative and, you know, bigger here, they're positive here and so on, when this is perfectly predictable by the ex post clustering, but it does mean that you should compute the standard error by clustering the data, we already know how to compute the proper standard error and that's to use the simple formula.

Q. So is it your opinion that using cluster standard errors overstate the true standard errors?

A. Yes.

Q. And your opinion is that they -- in using clustered robust errors in this case -- overstate the standard errors in Professor Noll's two regressions?

Page 33

A. Yes, I do.
Q. Okay. And what's the basis for that?
A. The -- the basis is that we know how to compute the proper standard errors, which is how Professor Noll does it, and the clustering only -- the clustering ex post induces correlation. And so if you add the term that's at the end of the cluster, the cluster robust standard errors, it's positive on average for exactly the -- the reason I just explained, using the simple example of hourly wages, because you've clustered workers, say, by their occupation and you know that workers in certain occupations are going to be correlated with each other because they're in that occupation, so they have either lower than average wages, they might have average wages or higher than average wages, but those averages move together within each cluster.
Q. And are there procedures or any tests that one could perform to test your conclusion that the clustered errors overstate the true standard errors in this case?
A. There aren't tests because the tests are going to -- are going to give you the conclusion that I just said. You don't need to test it because you know what's going to happen ahead of time, that if you cluster on the basis of some feature where the average

Page 34

systematically changes by that feature you will find that there's been cluster correlation because you've induced it by this -- this clustering that wasn't necessary.
Q. You stated that there aren't tests because the tests are going to give you the conclusion I just said.
A. There -- there aren't --
Q. Sounds --
A. -- useful tests. There aren't -- there aren't useful tests.
Q. So there are no tests?
A. That's correct.
Q. Okay.
A. There is theory and there is simulations.
Q. Are there any simulations that you could run to test your hypothesis that clustered robust errors overstate the true standard errors in Professor Noll's regressions?
A. Yes.
Q. And what are those?
A. You can do the exercise that I basically just laid out. In fact, we should make this -- if we were to make this about iPods and simplify the setting you could do the following: Suppose that there are five different kinds of iPods, what, classic, mini, nano, shuffle,

Page 35

touch, let's say, and suppose there's just a single before-after period where harmony was blocked and when it wasn't. So in this situation what would the right analysis be? If you just wanted an estimate of the average damage, then you could simply run a regression of the -- the price or the log price on the before-after dummy and not even actually have to include the -- the kind of -- the kind of iPod if you -- if both were available, both before and after and basically simulate the data so that it -- it is a random sample from the population and then ask what happens if we cluster after the fact on the kind of iPod, whether it's a classic, mini or shuffle or whatever. This is what Murphy and Topel do.
Q. That's your understanding of what they did?
A. They -- they clustered by time period --
Q. Okay.
A. -- and by family of iPod.
Q. And define for me what family.
A. Well, in this case there -- there would be no difference because -- I wanted to -- if I didn't say this -- to simplify things so that there's only one kind of nano -- nano, one kind of touch, one kind of shuffle, so that there's no differences in capacity or any other features like that.

Page 36

Q. Okay. And define for me what "family" means with respect to iPods.
MS. SWEENEY: Objection. Vague and ambiguous. Are you asking him -- well, I -- I just don't understand the question.
BY MR. KIERNAN:
Q. If you don't know -- do you know what "family" refers to with respect to iPods?
A. "Family," I believe, refers to not just the type, but also different characteristics of the -- of the iPod --
Q. And -- and --
A. -- capacity and features and so on.
Q. And how did Dr. Murphy and Dr. Topel cluster the standard errors? What's the cluster?
A. They said they clustered by family and -- and quarter.
Q. Okay. And how many clusters do they use?
A. I'm not exactly sure because I don't believe it was apparent from the report or I might have -- I -- I believe it's a few hundred.
Q. Do you know how many observations per cluster?
A. That's another thing that I did not find, but if you take in the one case 2 million observations and if it were 400 clusters, that would be 5,000 per cluster

Page 37

on average.
Q. Are there any -- is an alternative simulation from the one you described -- well, strike that.
So the simulation that you just proposed would be to simplify it by going to the model level rather than the family level?
A. Yes.
Q. Am I -- okay.
A. Uh-huh.
Q. Could you also run simulations using the family level?
A. You could, yes.
Q. Okay.
A. Uh-huh.
Q. And is one preferable to the other or would you run them both?
A. Well, if one had the time, you would want to run a simulation that reflects the particular application, yes.
Q. Okay.
A. But the -- the -- if -- if the data structure had been the simple one that I had proposed, then the only clustering that could have been done is by the class of iPod. And this is the analogue of what Murphy and Topel did in their more complicated situation. They

Page 38

defined more clusters, but that's because there's more time periods and more family -- the -- the family -- the number of families is larger in that case.
Q. Well, the simple structure that you described does not exist; correct?
A. Oh --
Q. That doesn't define the data structure?
A. It doesn't exist for this --
Q. In reality.
A. -- particular application --
Q. Correct.
A. -- but it defines lots of data structures that -- that have been used for intervention analysis, sure.
Q. Right. But not in this case?
A. It has the features, though, because, for example, once you have several thousand observations per cluster, then the simpler setting at least helps you learn something about how clustering can give you very overstated standard errors.
Q. Other than the simulations, are there any other procedures that one could perform to verify or test your claim that clustered robust errors overstate the true standard errors in Dr. Noll's regressions?
A. One could work out the theory, uh-huh.

Page 39

Q. And how would one work that out?
A. You start off with the presumption that -- first -- first, again, starting under the assumption that you have a random sample and then you see what happens after you cluster on a feature which could be family by quarter of -- of observation and -- and show that the -- what the bias in the clustered standard error is relative to the correct one.
The simulation -- it's important also to understand the -- the point of the simulation is that you can actually figure out what the proper standard error is --
Q. Right.
A. -- because you control the data and so you know which standard error is -- is the one that's close to the one you're trying to get. And the standard error that's going to win convincingly is the usual standard error that does not cluster.
Q. Well, isn't the point of the simulations to see which standard error is going to win?
A. That's correct.
Q. Okay.
A. Uh-huh.
Q. Does the OLS standard error --
A. I should -- I should bring something into

Page 40

this. We already know that the original standard error is going to do well because in this situation there are many -- more than 2 million observations in your case, but it would work really well with 1,000 observations, for example, because there's a simple formula that has been derived from theory that has been known for a long time and so we know that that formula is going to work.
The only issue is, is there any bias and what is the nature of the bias in doing the clustering when you don't have to? And so the issue is, really, how much are you going to be wrong by doing the clustering and how -- more to the point, how conservative will the clustering be.
Q. And as you -- as one increases the number of observations using OLS standard errors, can the bias -- would the bias tend to increase or decrease?
A. Oh, the bias will decrease. In fact, the -- as I said, in most applications, once you have a 1,000 or a couple thousand observations the standard error that you compute from OLS, even if they're the so-called heteroskedasticity robust standard records do quite well in those cases. That, actually, raises an interesting point about the clustering, is that once you've decided on the clusters, so in the Murphy and Topel case, they've taken a stand that the clustering should be at

Page 41

the family quarter level, which does raise the question, why not at the month family level or the week family level or the year family level? So how they came up with that clustering, I'm not sure. I don't think it's -- it's ever described. But once you've chosen the clustering scheme, the clustered standard errors will never get smaller. They depend only on the number of clusters you've chosen, not on the number of overall observations, which is peculiar because if you think of standard we should think that information is accumulating in a random sample as we get more and more data and that is what happens.

That's why you see the usual OLS standard errors heading to zero at the rate one over the square to the sample size and the clustered standard errors will just stay constant, given the number of groups that you have. So you're left in the odd situation that having lots of transactions data is viewed as being the same as having not very much transactions data. And that's because you're inappropriately clustering the standard errors.

Q. And just so I understand your opinion, if you keep group size constant --

A. If you keep --

Q. It's the G --

Page 42

A. -- the number of groups constant?

Q. Right.

A. Yes.

Q. Let's call it G.

A. Uh-huh.

Q. So G, using your example, equals ten.

A. Yes.

Q. Okay. If you increase the number of observations per group --

A. Yes.

Q. -- per cluster --

A. Uh-huh. And the data have -- have come from a random sample.

Q. Oh, okay. Well, what if the data comes from the entire population?

A. Well --

Q. What's the impact of using clustered robust standard errors --

A. Well --

Q. -- as the number of observations per cluster increase?

A. Well, the traditional standard error, if you don't -- if you actually act as if you have -- the entire population is zero because you have no -- you have no sampling error in the -- in the estimation.

Page 43

That's why you have to use a different kind of thought experiment, which is -- one of which I laid out in section 5 of my declaration that shows you that, essentially, with the whole population you can argue that the usual standard errors are the -- are the right ones to use and, if anything, they're actually conservative because when you have the whole population, there's a -- a population correction that always reduces the standard errors.

So the -- as you get more and more data, again, if you fix the number of clusters using the entire population is operationally the same as getting more data in the perspective that I laid out in -- in section 5.

Q. And would you expect the OLS standard -- the bias of the OLS standard errors to increase or decrease as you increase the number of transactions per -- number of observations per group, keeping the group constant?

MS. SWEENEY: Objection. Vague and ambiguous. Incomplete hypothetical.

THE WITNESS: Okay. So let me -- let me say this again. The data have been collected by random sampling and so the clusters that have been -- that -- the usual OLS sustained errors ignore the clustering and they properly ignore the clustering because this is an

Page 44

ex post structure that you've imposed on the data after you've collected it. So the clustering that you use has no effect on the traditional OLS standard errors as it should be.

Again, let me give you an example. I mention the -- the hourly wage occupation example. Suppose that in addition to occupation you collected information on highest grade completed, so schooling. Now, if you did exactly the same exercise, if you collect the data -- and, remember, the goal here is to just estimate the average wage in the population, but you say, I have information on schooling, now I'm going to put people into clusters based on the highest grade they've completed, which might be five, ten categories, you're going to find exactly the same phenomenon.

On average, people with lower education are going to have a lower hourly -- hourly wage, and so within that cluster you're going to find correlation. Same thing, people with high levels of education are going to have on average a higher hourly wage. So now you've got occupation and you've got education and there're two different ways of cluster, so which is the right one? You know the answer has to be neither is the right one because you've -- you've collected it via a random sample. The goal is to estimate the population

11:36 a.m.
(Recess.)
(Mr. Murray telephonically joins deposition.)
THE VIDEOGRAPHER: Okay. We're back on the record 11:59 a.m.
MS. SWEENEY: Professor Wooldridge, go ahead and make those clarifications.
THE WITNESS: The first clarification was the amount of hours I spent on the case up to writing the declaration, that was -- when I said five or six hours, that was the actual time writing the declaration. I spent another six hours reading the background material, the Noll report and the Murphy and Topel report.
MR. KIERNAN: Okay.
MS. SWEENEY: Was there one other one?
THE WITNESS: The Noll rebuttal?
MS. SWEENEY: No, I'm sorry. I thought that you were going to clarify two issues of testimony.
THE WITNESS: Oh, oops.
MS. SWEENEY: That's okay.
BY MR. KIERNAN:
Q. All right. Dr. Wooldridge, I was going back through my notes and I -- it wasn't entirely clear to me, under what circumstances could clustering standard errors be appropriate when you have all the -- the --

you have the entire population of transactions.
MS. SWEENEY: Objection. Asked and answered.
THE WITNESS: As I said before, there's only one case that I could think of and that's where ex post you -- you group the observations and then you use information that's computed from other transactions as part of the regression model for a particular transaction.
This would be like taking a sample of students and then computing family income of some peers who live next to them and including that in a regression model, but there's nothing like that done in the analysis by Professor Noll.
BY MR. KIERNAN:
Q. Is that something that could be done with the transactional data for iPods?
MS. SWEENEY: Objection. Incomplete hypothetical. Vague and ambiguous.
THE WITNESS: It could be done, but I'm not sure why anybody would do that.
BY MR. KIERNAN:
Q. Are there any --
MS. SWEENEY: Hold it. Excuse me. Before we go on, did -- did anyone join the --
MR. KIERNAN: Oh, yes.

MS. SWEENEY: -- the deposition?
MR. KIERNAN: Scott Murray from Apple is on the phone.
MS. SWEENEY: Okay.
MR. KIERNAN: Hi, Scott.
MR. MURRAY: Hello.
BY MR. KIERNAN:
Q. As you sit here today, can you think of any circumstances under which one would ex post group the observations of iPod transactions and then use information that's computed from other transactions as part of the regression model for a particular transaction?
A. I can't think of why you would do that because a hedonic price regression is about relating the price of a particular unit to characteristics of that unit. And, of course, prices will change over time as demand in supply, conditions affect prices, but that's the -- the nature of a before-after analysis where you want to account for or control for the characteristics of the particular units that are being transacted.
Q. Earlier you describe two ways, two procedures, that could be used to test the hypothesis that clustering -- that the clustered standard errors by Drs. Murphy and Topel inflate the standard errors compared to

the true precision of the estimates.
A. Yes.
Q. Are there any other procedures that you can think of?
A. No. One -- one has basically two tools at -- at one's disposal when trying to evaluate any kind of statistical procedure. And since standard errors are a measure of precision of estimates. That measure is across different realizations or samples of data. And so you can either do a theoretical calculation, which uses the tools of statistics to account for the fact that we're seeing different realizations of data or you can actually do a simulation which creates different samples or realizations of the data and study the problem that way.
Q. With respect to the theoret- -- you said "theoretical calculation"?
A. Yes.
Q. Okay. Have you done a theoretical calculation that's set forth in your declaration that tests the conclusion or hypothesis that Drs. Murphy's and Topel's clustering standard errors vastly inflate the standard errors compared with the true precision of the estimates?
A. After writing my declaration, I did do a

calculation like that, yes.

Q. Is it in your declaration?

A. No. It happened after my declaration. It was to support what I knew had to be true by thinking through the -- the different kinds of examples.

Q. And have you produced that calculation that you're referring to?

A. Produced it to?

Q. To the lawyers in this case.

A. No.

Q. And are you relying upon those calculations, the theoretical calculations, for the opinions set forth in your declaration?

A. No. What I was relying on was the idea that -- and -- and, again, I have to admit, I scaled the problem down so I could think about it better and thinking about either a few occupational classes or a few classes of iPod and what would happen in that case if you clustered on a characteristic such as occupation or class of iPod after you collected the data, and it became clear that, of course, you would find in some of the clusters there -- the residuals have a below -- an average below zero, in some cases it would be above zero. The weighted average of them has to even out, has to be zero, because you know that you're -- you're

computing the overall population average.

Q. And --

A. And so that --

Q. Go ahead.

A. -- perfectly explained what Murphy and -- Murphy -- Professors Murphy and Topel were finding in their calculation of clustered standard errors and looking at the residuals.

Q. What perfectly explained what Professors Murphy and Topel were finding in the calculation?

A. Well, Professors Murphy and Topel report after they do the clustering based on the family by calendar -- by -- by quarter, that they -- to show that there was a -- a problem that needed to be addressed with clustering, they computed the residuals within each of these clusters and they used as their main piece of evidence that -- or one of the main pieces of evidence that these average residuals were different across the different clusters.

And I said that that is perfectly explained by the fact that there -- the prices are going to be on average different for different families as well as different quarters and that in no way implied that the standard errors had to be computed with -- with clustering.

Q. And what are some of the reasons why the prices would be different for different families?

A. Oh, well, of course different places can have different sales going on. They can have -- this isn't my area of expertise.

I -- as I said, I haven't even looked at the data, and I don't need to to understand that a family decides -- is presented with a price or a reseller is presented with a price, and they make a decision to buy at that price or not.

The fact that those prices may be the same for several families is -- does not imply that there is a clustering problem. One can think of many situations where -- where that's true. I give an example in my declaration.

Q. Sure. But could there be a clustering problem?

MS. SWEENEY: Objection. Vague and ambiguous. Incomplete hypothetical.

BY MR. KIERNAN:

Q. Well, you said -- so we can clarify it, you said the fact that they may be the same for several families does not imply that there is a clustering problem, one can think of many situations where that's true. I gave an example in my declaration.

Are there circumstances --

A. Oh --

Q. -- that you can think of where there is a clustering problem or could be?

A. No, not with the way the data have been collected. So in -- in my declaration I use the example that also had prices that you would expect not to -- to vary much, especially within geographic units and within time and that would be looking at the prices of some standardized item at a fast-food restaurant.

The fact that two people might go to the same fast-food chain and pay the same price does not mean that those two observations form a cluster. They're independent draws based on a person's decision to buy or not at that particular price and, in fact, it's -- the fact that there isn't that much variation in the prices, so the -- the example I used was suppose you're trying to -- to decide whether prices are systematically different in poor neighborhoods and -- and what I call nonpoor neighborhoods, the fact that there may be little price variation makes it all the more impressive if you can actually find a difference across the two different kinds of neighborhoods.

And, of course, the fact that there's little price variation means that the variance of the residuals

will be small and that helps with the precision of the estimates. And this is what you find in Professor Noll's calculation where the standard errors there's a small residual variance and there's a lot of observations and so he properly finds small standard errors in his regression analysis.

Q. Is there a point at which the standard errors are so low that would cause an econometrician like yourself to question whether they were accurately calculated?

MS. SWEENEY: Objection. Incomplete hypothetical.

THE WITNESS: All I'm concerned about is given the particular application, the model, the estimation method, the way the data have been collected, has the appropriate method been used or not and with lots of -- lots of observations and with little residual variance, there's no rule of thumb below which the standard errors would have to hit before you -- before you got suspicious. So I would say, no, there isn't -- there isn't some sort of threshold.

I would -- I -- I evaluate these on the -- on the -- on the merits of the modeling exercise and the -- the estimator used and in this particular case on how the sample is -- is obtained or in this case the

population.

BY MR. KIERNAN:

Q. Are there factors that could impact the reliability of the precision of the standard errors that Dr. Noll reports?

A. I can't think of any. A standard error calculation is a fairly straightforward thing in most cases with standard econometric methods such as OLS once you understand how the data have been -- been obtained. I should add, he did make the standard errors robust to heteroskedasticity of unknown form, which means the variance can change in an arbitrary way across transaction and that is the appropriate thing to do.

Q. Are there any authorities, textbooks, public articles that support your conclusion that clustered robust errors overstate the true standard errors when using the entire population of transactions?

A. Actually, this is fairly recent material. I started thinking about this a couple of years ago when I had conversations with two -- two people that I've worked with and we let it go. And since then I've been, after writing the declaration, thinking about the merits of this case, I've worked out a little bit of theory as well as the simulation.

It is commonly thought that clustering is,

quote, harmless if it's not needed, but this is -- this is not true in the context of clustering after you've collected a random sample. If you have collected a cluster sample and you have a large number of clusters and relatively small observations within a cluster, then -- then you can show, as the number of clusters get large, the standard errors will approach the right values, but that's assuming you've collected a cluster sample.

Q. Can you cite to any authorities, textbooks or articles, that support your conclusion that clustered robust errors overstate the true standard errors when using the entire population of transactions?

MS. SWEENEY: Objection. Asked and answered.

BY MR. KIERNAN:

Q. I'd like the actual names of the authorities, textbooks or articles to the extent the question was confusing.

A. Oh, so I said that I -- I don't know of any. I've -- I've worked this out since submitting my declaration.

Q. You mention that at some point you talked to -- did you say two people?

A. Uh-huh.

Q. And who were they?

A. Not -- not for this case.

Q. Understood.

A. Alberto Abadie is an econometrician at Kennedy School of Harvard and Guido Imbens is an econometrician at the Stanford Graduate School of Business. I've co-authored with Guido before and I actually do lectures with him.

Q. And you noted that the three of you let it go. What -- what did you mean by that?

A. Oh, it actually -- we didn't completely let it go. We just all got busy and were working on different things.

Q. And did the three of you or any number of you author a working paper?

A. There's no working paper.

Q. Any drafts of a working paper?

A. No.

Q. Any working paper of a working paper?

A. No.

Q. And are you working with them now -- either one of them, now that you've picked this topic back up?

A. I believe we will pick the topic up, yes.

Q. Okay. Have you talked to them about it?

A. No. We have been actually talking about the other -- just to clarify -- the material on the finite

population analysis that I mentioned in section 5, I should say when you're using the entire population.

Q. All right. Describe for me, as precisely as you can, Apple's pricing strategy for iPods.

MS. SWEENEY: Objection. Vague and ambiguous. Overbroad.

THE WITNESS: Actually, I don't know what their strategy is.

BY MR. KIERNAN:

[redacted]

BY MR. KIERNAN:

Q. Do you know how often Apple reviewed its pricing for particular iPod families?

MS. SWEENEY: Same objections.

THE WITNESS: No.

BY MR. KIERNAN:

Q. If you could state it again because you guys --

A. No.

Q. -- talked over each other.

A. No.

Q. Okay. Is that something you examined?

A. When I read the reports, I focused mainly on the econometric issue of clustering. I -- I really didn't form an opinion or absorb the particular of the pricing strategy.

Q. Do you know what factors Apple takes in account in setting the prices for a typical -- pardon me. Strike that.

Do you know what factors Apple takes into account in setting prices for particular types of iPods?

MS. SWEENEY: Objection. Vague and ambiguous. Overbroad. Compound.

THE WITNESS: Well, I assume -- I assume cost is involved and I assume that the -- the -- the demand for the various features is involved.

BY MR. KIERNAN:

Q. Putting aside your assumptions, do you know what factors Apple actually took into account in setting the prices for any iPod that Dr. Noll examined in this case?

A. No.

MS. SWEENEY: Same objections.

Give me a second to interject my objections.

THE WITNESS: Okay.

BY MR. KIERNAN:

Q. For a particular iPod family, how often did Apple change the price on the iPod?

MS. SWEENEY: Objection. Vague and ambiguous as to time. Overbroad. Compound.

THE WITNESS: I'm not sure.

BY MR. KIERNAN:

Q. Do you have any knowledge of how frequent Apple changed the price of an iPod family?

A. No.

[redacted]

Q. Have you examined how Dr. Noll's regressions controlled for Apple's pricing policies and its impact on Apple's prices for iPods?

MS. SWEENEY: Objection. Vague and ambiguous.

THE WITNESS: I looked at his regressions and noted that there are various features of the families -- the units themselves as you would expect in a hedonic price regress- -- regression.

BY MR. KIERNAN:

Q. Okay. And did you examine whether Dr. Noll included all of the explanatory -- all of the variables

that -- or strike that.

Did you examine whether Dr. Noll controlled for all the factors that Apple considered when it set the price for an iPod family?

A. I didn't form an opinion about that. I wasn't asked to consider the model specification.

Q. How did Apple set the price of iPods sold in Apple retail stores?

A. I don't know that.

Q. Okay. How did Apple set the price of iPods sold on Apple online store?

A. Again, I don't know that.

MS. SWEENEY: And I'm going to belatedly object. Vague and ambiguous. Compound.

BY MR. KIERNAN:

Q. For a particular iPod model, let's say, the iPod nano second-generation, 4 gigabyte, would the price that Apple listed for that iPod be the same at the retail store and the Apple online store?

A. I don't know.

MS. SWEENEY: Same objection.

Sorry. Go ahead.

THE WITNESS: I don't know.

BY MR. KIERNAN:

Q. How did Apple set the price of iPods sold to

Page 77

Q. -- family has a different feature?
A. No. So if you -- again, if you go back to the case of how the clustering was done, it's not true that you have to account for all of those features in order for the usual standard errors collected under random sampling to be valid. So --
Q. Go ahead.
A. -- for the issue of computing the standard errors -- and that -- that's why I'm not commenting on model specification -- no, it doesn't matter that there are some features that may not have been accounted for or some interactions of features or something like that. That's a modeling question. That's not a question about the standard errors.
Q. Well, isn't the --
A. So this --
Q. Go ahead.
A. So this is -- this is a common misperception. If you take -- again, let's just start with a -- a large population and you're going to take a large random sample and you're going to estimate two models. You have Y and you have X1 and X2 and you regress Y on X1 and you regress Y on X1 and X2.
Now, whether you should include X2 or not is a modeling issue. The standard error that you compute in

Page 78

the usual way are valid even in the first regression where you've omitted X2. The fact that you've omitted X2 does not affect the calculation of the standard errors.
Q. And what does it affect?
MS. SWEENEY: Objection.
BY MR. KIERNAN:
Q. What would it affect under that scenario?
MS. SWEENEY: Vague and ambiguous. Incomplete.
THE WITNESS: Well, it --
BY MR. KIERNAN:
Q. I'm not going to use -- let me strike that.
I want to use a hypothetical that you were just using and you said that omitting the variable would not affect the calculation of the standard errors. Would omitting the variable effect anything else in the model?
MS. SWEENEY: Objection. Vague and ambiguous.
THE WITNESS: Well, sure, it could bias the coefficient -- the coefficient and the simple regression.
BY MR. KIERNAN:
Q. And so -- and is that what you referred to in your book as omitted-variable bias?

Page 79

A. Yes.
Q. And is it your testimony that omitted-variable bias -- omitted-variable bias has no impact on the -- on reliably calculating the standard errors for a model?
A. That's correct.
Q. And just to make sure that you and I are on the same page, when you refer to "omitted-variable bias," what -- what are you referring to? Define that for me.
A. Well, you would like to estimate the coefficient on X1, let's call it beta 1, controlling for the effects of X2 and if X2 is correlated with X1 and you leave it out of the regression, then, in general, the estimator of beta 1 will be biased.
Q. The coefficient on the X1 will be biased?
A. That's correct. Yes.
Q. But that will have -- your testimony is that will have no impact on the calculation of the standard errors?
A. That's correct. You'll get a valid standard error and confidence interval for the parameter that you are estimating. It's actually easy to think this through. You just -- you can always write any equation that you're estimating. There's a population version of

Page 80

it. And so you can write Y as a linear function of X1. It may not have the coefficient that you want, but you can always do that and you can always write the model for Y as a function of X1 and X2. Once you have a random sample, the calculation of the standard errors is standard. There's no adjustment that needs to be made because you might have omitted X2.
MS. SWEENEY: Did you want to break for lunch?
MR. KIERNAN: Let me see if I'm done.
Yeah, why don't we do that.
THE VIDEOGRAPHER: This will be the end of DVD No. 1. We're going off the record at 12:41 p.m.
(Recess.)
THE VIDEOGRAPHER: This is the beginning of DVD No. 2. We're going back on the record at 1:49 p.m.
BY MR. KIERNAN:
Q. Okay. Dr. Wooldridge, the error terms in Dr. Noll's regressions, what do they represent?
A. Factors that affect price that we don't observe.
Q. And what --
A. Actually, they -- they can just be viewed as the difference between Y and its expectation conditional on the variables that are included in the regression.
Q. And what would be some reasons why factors

Page 81

that affect price would not be observed?
A. Well, usually there are various factors that can --
MR. MURRAY: Scott here.
BY MR. KIERNAN:
Q. Go ahead. You said usually there are various factors that can . . .
A. Right. So if -- for example, if there are systematic differences across family and calendar year, then those differences would be included in the error term.
Q. And would -- if there are omitted product attributes that impact price, would those be captured in the error term?
A. The -- well, let me -- let me answer that like this: If -- it's not necessarily what is in the error term that is -- that's important. It's basically if you're trying to learn about the coefficient on a particular variable the question is whether you've included enough of the other factors.
So if there -- I mean, as I tried to explain before, the -- the nature of the error term, whether it includes omitted factors or not, does not affect the issue that I was asked to -- to evaluate, which is the clustering issue.

Page 82

Q. Okay. Could variables that are unobserved or not measured by the regression impact the calculation of the standard errors?
A. Not when -- not when it's the population or a random sample from the population. I -- so I gave you that example where you had X1 and then you had X1 and then X2, and it is, I believe, common -- some- -- somewhat commonly thought that the omission of X2 can somehow affect the calculation of the standard error for the coefficient on X1, but that's -- that's not true under the sampling scheme that we're talking about, random sampling or knowing the population.
Q. Okay. In the -- let me back up a step or two. How is the standard error calculated? As Dr. Noll -- using Dr. Noll's regression, how did he calculate the standard errors?
A. So he -- you would take the residuals from the regression and from those residuals -- so the -- the basic calculation that you learn in your first econometrics course estimates the variance of the error term by using the sum of squared residuals from the regression and then dividing it by a degrees of freedom correction and then that gets multiplied by the so-called X prime X inverse matrix. And that's the valid -- that's the valid calculation for the variance

Page 83

estimate under the assumption that the variance of the error term doesn't depend on any of the factors you've included in your regression model. And there's an adjustment that allows for that variance to be unrestricted, an unrestricted function of those factors that is a little more complicated than that.
Q. And -- and what is that?
A. It's the so-called Eicker Huber White Estimator, which sometimes is called a sandwich estimator because the way the formula appears where on the inside there's a more general matrix that's estimated that allows for the squared error term to be correlated with the Xs and that's where the robustness to heteroskedasticity comes from.
Q. And in calculating the standard errors, is there an assumption that the error terms in the regression are independent of one another?
A. Yes. Well, the -- the -- the foundation of my book, actually, in -- in the case of random sampling, they're always independent of one another. So when you have a random sample there's no issue about whether they're independent or not because they automatically have to be.
Q. Right. And what about the case when it's not a random sample? Are you applying that assumption that

Page 84

the --
A. Well, if -- for example, if the data were truly collected by a cluster sample so that you were sampling clusters rather than individual units, then there would be some -- then that usual calculation of the heteroskedasticity robust matrix would not be correct, but that is assuming that you have collected -- you have cluster sampling.
Q. Are there any circumstances under which ex post clustering, as you've described it in your report and today, would be appropriate when calculating standard errors when dealing with an entire population of transactions?
A. It -- it depends on the regressors that were included -- the -- the factors that -- that are included in the model and as long as those factors are specific to the individual transaction, then, no. And that's what Professor Noll did.
Believe early I men- -- earlier I mentioned a -- a case where if you sampled students independently, but then after the fact looked for peer effects by looking at, you know, children who live near them and so on, then that sort of addition to the model would or possibly could induce cluster correlation within the errors.

Page 85

Q. With respect to that example -- I'm glad you brought it up -- at what level would you cluster in that example that you just gave?

A. Um, you would define -- typically what you would do is you would have to define the notion of who are the potential peers for a particular student. And so it might be something like the classroom or something like the, you know, school and then you would compute an average once you have defined what the peer group is and then you would include those and so you would cluster at that level.

Q. And what factors would an econometrician consider when deciding at what level -- the level of clustering, whether it was, in your example, the classroom or something else, some other level?

A. Well, ideally in the case where you've actually collected a cluster sample, that determines it for you because you know what the clusters actually are. And the other consideration is if you include explanatory variables that are created by defining clusters, then that would also define the level of clustering.

Q. And -- and going back to your example where the data was not collected by cluster sampling --

A. Uh-huh.

Page 86

Q. -- what factors would an econometrician consider in determining the level of clustering in that scenario?

A. Again, you'd have to first define what the peer group is and then that would determine the level of clustering. So if --

Q. Based on what? Based on what factors? I mean --

A. Well, in -- in --

Q. -- what would an econometrician consider?

A. -- in that particular example, it's -- it actually doesn't have anything to do with being an econometrician. That's the sort of question that the person undertaking the empirical work has to decide, what -- what sort of children do I think affects a particular child's outcome? So it would be up to you to specify ahead of time that it's, you know, the neighborhood or the classroom or something like that.

Q. Okay. And my view of that may be different from some other --

A. That's correct.

Q. -- somebody else's.

So it's a matter of judgment --

A. Uh-huh.

Q. -- of the person running the study?

Page 87

A. Yes. But if there are -- if there are no so-called peer effects included in the equation, then there's no need to cluster.

Q. And -- one second.

If under that scenario there was no peer effect, but the researcher did cluster by the -- by the peer clustering level, as you suggested --

A. Uh-huh.

Q. -- what impact, if any, would that have on the precision of the standard errors?

A. Well, it depends on the -- if you had a large number of clusters with few observations, then the effect might be fairly small. But then you would see that the effect is fairly small. That's the -- the proof, essentially, is in the pudding, is that since you know that the original standard errors are correct, if you do the clustering and it matters a lot, then you've either got an unusual sample or the cluster effects have a bias in them -- I'm sorry -- the clustered standard errors have a bias in them.

Q. And, then, knowing that -- sorry -- you said since you know that the original standard errors are correct.

A. So the situation was --

Q. Yeah.

Page 88

A. -- we randomly sample from the population --

Q. Right.

A. -- and then I said where you might have to cluster is when you create a peer effect. And then I -- I -- maybe I misunderstood you.

Q. No, no.

A. You -- you said, but suppose there is no pure effect and you cluster anyway, well, that can only increase the standard errors on average.

Q. What I was referring to, Dr. Wooldridge, is I want to understand when you stated -- what you meant by the original standard errors are correct and you were referring to --

A. I mean the ones that were obtained via the Eicker Huber White heteroskedasticity robust formula, yes.

Q. By drawing a random sample?

A. Yes.

Q. I asked you are there circumstances under which ex post clustering, as you've described it in your report and today, would be appropriate when calculating standard errors when dealing with an entire population of transactions. Do you recall that question?

A. I do.

Q. Okay. Depends on what regressors were

included, the factors that are included in the model, as long as those factors are specific to the individual transaction, then, no, and that's what Professor Noll did.

What is an example when there are factors that are not included where ex -- that would make ex post clustering appropriate --

A. There --

Q. -- or something to be considered?

A. There -- there aren't any unless you actually create these clusters ex post. So, in other words, you'd have to make a conscious decision that you wanted to -- to include information about other transactions directly in the equation for this particular transaction.

So, again, maybe I'm not -- you would be the cause of the clustering problem because you decided to do that. If you don't decide to do that, then there can't be a clustering problem.

Q. Yeah, I think you and I are talking past one another.

My question is -- maybe I'll ask -- reask the question.

Are there any circumstances where you have the entire population of the data in which ex post

clustering, as you described it in your report and today, would be appropriate when calculating the standard errors?

MS. SWEENEY: Objection. Asked and answered.

THE WITNESS: So, again, if you are using information from the transaction that has been drawn, the answer is, no. Only if you've created a problem where you, essentially, use information from other observations can you create a clustering problem in that setting.

BY MR. KIERNAN:

Q. Ex post clustering, is that a generally accepted term in econometrics?

A. There's something closely related called ex post stratification.

Q. And is it your testimony today that those are the same thing?

A. No, they're not the same thing.

Q. Okay.

A. But they're --

Q. Focusing on ex post clustering, is that a term that is used in the field of econometrics?

A. That's a good question. I'm not sure I could point to a source for that, actually. I basically -- the idea is that for a long time cluster sampling was

understood to be you sample clusters from the population and this notion that you would, essentially, take a random sample and then create clusters after you've, essentially, observed the random sample is, I think, a fairly recent phenomenon and turn -- it turns out to be incorrect.

Q. And what authority can you point to that supports your conclusion that ex post clustering is incorrect, as you just put it?

A. Well, I -- I like to think of myself as an authority on this, and this is an issue that has come up in this particular case, and it's come up in some other areas that I'm aware of. That's why, in fact, I mention the two authors that I've been working with started talking about this problem a couple of years ago. So I guess I and my coauthors are the authorities.

Q. And can -- have you published any peer review articles that support your conclusion that ex post clustering is inappropriate when dealing with an entire population?

A. Not on that specific topic, no.

Q. And are you aware of any peer-reviewed articles or other publications that support the conclusion that ex post clustering is inappropriate when dealing with the entire population?

A. The closest thing, which, again, the -- the whole population versus sampling is a bit of a red herring here because if -- if we had taken a -- a large random sample, then the conclusions would -- would be, essentially, the same. The closest I can think of is my own work on talking about inappropriately clustering a stratified sample.

Q. And where's that work?

A. That's the American Economic review paper, the 2003 paper.

Q. And is it your testimony that that paper states that ex post clustering is inappropriate when dealing with an entire population of transactions?

A. No. It's -- so let's be clear. It's a -- a statement about how clustering, when the data had been collected from a stratified sampling, is inappropriate.

Q. And is the -- was the data in -- that's used by Professor Noll, was that collected from a stratified sampling?

A. No. So a special case of stratified sampling is random sampling and so if he had thrown out, you know, 80 percent of the data and called that a random sample, then it would apply -- apply directly to that case. But, as I said before, when you collect more data, that's only a good thing. So the different -- the

issue of the population versus the sample is irrelevant here for the clustering issue.

Q. So yours and Dr. Noll's opinions with respect to the population are irrelevant to the issues in the case of whether clustering is appropriate?

MS. SWEENEY: Objection. Misstates his testimony. Argumentative.

THE WITNESS: No, I think that -- that's not -- certainly not what I said. The --

BY MR. KIERNAN:

Q. How is it relevant then?

A. How is?

Q. How is -- how is the fact that, in your view, that Dr. Noll had the entire population of iPod transactions relevant to the opinions that you're offering in this case?

MS. SWEENEY: Objection. Asked and answered.

THE WITNESS: I'm not sure how I can answer that differently.

BY MR. KIERNAN:

Q. Well, you stated --

A. So -- so, again -- so --

BY MR. KIERNAN:

Q. -- the issue --

A. Okay. So let me --

MS. SWEENEY: Please don't -- David, don't interrupt.

MR. KIERNAN: He had stopped --

THE WITNESS: Yeah, and --

MR. KIERNAN: -- and I started. So I didn't interrupt him. He actually interrupted me.

THE WITNESS: That was my fault, yeah.

BY MR. KIERNAN:

Q. But go ahead.

A. The -- again, if you take the millions of transactions that are in the population and you took a random sample from those, okay, the clustering on the basis of characteristics that you, you know, observe like the family in the quarter, would be the incorrect thing to do.

As you get more and more data, that doesn't change and so whether you think of that as the entire population or a larger random sample, essentially you -- you get -- you get the same answer.

MS. SWEENEY: The spotlight is on you.

MR. KIERNAN: But the documents are great. We're all laughing because I made a joke.

BY MR. KIERNAN:

Q. Your opinion that there can be no cluster sampling problem because there is no sampling, what

authority supports that opinion?

A. The argument that I just used for you, that --

Q. Okay.

A. -- the --

Q. Any peer-review publication that supports your opinion that there can be no cluster sampling problem because there is no sampling?

MS. SWEENEY: And I'd just like to interject an objection and ask the witness, you can go ahead and finish your prior answer that Mr. Kiernan interrupted.

BY MR. KIERNAN:

Q. And if I did interrupt you, I did not mean to.

A. So, again, the -- when you have a large population you could take a random sample from that in which case there's no justification for the clustering.

Q. Yeah. I understood your argument. What I'm asking is what peer-reviewed authorities can you cite to me just --

A. For the population problem, I -- I can't.

MS. SWEENEY: And, again, I'm going to ask you to stop interrupting the witness. You -- every time he starts to give an answer, you interrupt if you don't like it. You're not entitled to do that. So please don't interrupt the witness anymore.

BY MR. KIERNAN:

Q. In your report you describe the -- section 5 you referred to it a couple of times today -- to the unconfoundedness assumption.

A. Uh-huh.

Q. Do you recall that? And define that for me.

A. That means that the assignment to the treatment in the control group doesn't depend on, in this case, what the price differential would be under two regimes. So think of -- think of it before harmony was blocked and after harmony was blocked and actually just -- you don't have to bring time into it, just two states of the world, harmony is blocked, harmony is not blocked. And then you see some units where that was true and some units where that wasn't true and the idea is that the intervention is independent of what the difference in prices would have been in the two states of the world.

Q. And what circumstances must exist for the unconfoundedness assumption to hold?

A. Well, there -- you could have an experiment where you have a random intervention or you can have a before-after where you have included enough factors so that it's -- the intervention is effectively random after you've included those factors.

Q. And have you examined whether or not with

Page 97

respect to Professor Noll's two regressions the unconfoundedness assumption holds?
A. That's more -- that's a modeling question, so I did not think about that, yes.
Q. And so you're not offering an opinion on whether or not the unconfoundedness assumption applies with respect to Professor Noll's two regressions?
A. That's correct.
Q. If there were important variables that explained prices of iPods that were left out of the regression, could that cause the unconfoundedness assumption to be violated?
A. Yes, it could and -- but I don't have an opinion on whether that's the case here.
Q. When dealing with an entire population, are there methods available to econometricians to account for correlation of error terms when calculating standard errors?
A. Well, like I said, there's no need to do it when you're using only the information from each record in your econometric analysis.
So the -- you know, the fact that you then decide that after you have this entire population that you're going to, essentially, arbitrarily define clusters and then conclude that based on your

Page 98

definitions you're seeing correlation within those clusters does not say that you should use cluster robust inference.
Q. In your textbook you note that "We should not expect good properties of the cluster robust inference with small groups and very large group sizes when cluster effects are left in the error term."
Do you recall saying something along those lines?
A. Yes, uh-huh.
Q. What do you mean by that?
MS. SWEENEY: Can -- can -- can I interject for a moment? Do you have a copy of that?
MR. KIERNAN: Well, he recalls it.
MS. SWEENEY: Yeah, but can you give the page cite?
MR. KIERNAN: I don't have the page cite.
THE WITNESS: Is that from the second edition of my book?
BY MR. KIERNAN:
Q. Yes.
A. Means that if you -- if you use -- if you use cluster sampling and you choose only a relatively small number of clusters with large cluster sizes, there's no theory that says that those cluster robust standard

Page 99

errors have any desirable properties.
Q. Okay. And when you state, "We should not expect good properties of the cluster robust standards," what are the good properties that you're referring to?
A. You would want them to, essentially, be unbiased estimates or even consistent estimates of the actual sampling variances.
Q. Right. So one would be unbiasness?
A. Yes. Well, that's -- so the sampling variances, the -- the usual OLS estimators are unbiased estimators of the -- the usual -- the usual variance estimators for the OLS -- that's a -- the sampling variances for the OLS estimators are unbiased and then we take the square roots of them to get the standard errors. And you can't always -- the -- the cluster robust ones, they're never exactly unbiased, so you talk about approximations and you often talk about what happens as you get more and more data.
And the theory on cluster sampling does not allow for the case where you have a small number of clusters and -- well, having a small number of clusters is a problem in general and certainly when you have a large number of observations per cluster none of the theory applies to that case.
Q. And what number of clusters is -- what

Page 100

threshold is the point at which the number of clusters is no longer a problem?
A. This is a -- an impossible question to answer. It is the question empirical people are most interested in. So assuming that it's appropriate to cluster, it depends on lots of different characteristics of the problem. For example, it depends on how big the cluster sizes are. It depends on the distribution of the observables and unobservables in the population.
So this is something that theory can't easily answer and that's why people do simulations to try to find when the -- the -- the theory of having a large number of clusters seems to work fairly well.
Q. When you referred to "observables" in your last answer, were you -- and "unobservables," what were you referring to?
A. The explanatory variables and -- as the observables and then the error term is the unobservables.
Q. And so that I understand your answer, it's that theory doesn't provide the answer and so there are econometricians that are implying empirical analysis to try to answer that question?
A. Simulation studies, yes.
Q. Is it simulation studies like what Hanson did

and the quarter or the date of the transaction and so on.

BY MR. KIERNAN:

Q. And with respect to the school example --

A. Uh-huh.

Q. -- if I got all the students within a state --

A. Uh-huh.

Q. -- and I got the big data file, wouldn't it be analogous to what you just described in that I would learn of the school and the district from that same dataset?

MS. SWEENEY: Objection to form. Vague and ambiguous. Incomplete.

THE WITNESS: Yes. And, in fact, that's why I included that example in my declaration is because once you've gathered the information on the students, the fact that you also learned about their school and their school district does not mean you should then group them into clusters based on their school or their district.

BY MR. KIERNAN:

Q. But that example in your report -- just to make sure I understand it -- aren't you, what you're referring to right now, is when you randomly sample the students at the first stage?

A. Yes.

Q. Yeah. Okay. Differ- -- I have a different hypothetical.

A. Okay.

Q. You pull all the students' information first --

A. Uh-huh.

Q. -- and then from there you learn about -- you learn from the dataset you have the entire population of students --

A. Uh-huh.

Q. -- and then you learn from the dataset, the schools and districts and so forth. How is that different from what you describe on page 3?

MS. SWEENEY: Objection. Vague and ambiguous. Incomplete.

THE WITNESS: As a practical matter I don't think it's different.

BY MR. KIERNAN:

Q. Okay. How does one determine -- or strike that.

What methods does an econometrician apply to determine if something is ex post clustering?

A. Oh, well, you have to know -- there are -- so if you have a sample of data, then, of course, you know because after you've drawn the observations you can see

what you're clustering on. In the case of a whole population you know because the clustering is essentially arbitrarily defined.

Q. Any other factors that one would consider if ex post factor clustering occurred when a whole population is being used? You named the clustering is arbitrarily defined. Any other factors?

A. Well, in -- again, since the distinction between the population and the sample is really one of number of observations here, if you can determine it based on a random sampling thought experiment, then you can determine it in the population as well.

So, in other words, if I took a random sample of the transactions and then clustered on the basis of family and quarter, then the same sort of clustering would be inappropriate with the entire population.

Q. And what factors does an econometrician consider in determining whether the clustering was, in your words, essentially arbitrary?

A. Well, because -- because the observations don't have a cluster structure.

Q. And how do you determine that? How does -- what are the factors that an econometrician considers to determine whether, in your words, there is -- they do not have a cluster structure?

MS. SWEENEY: Objection. Asked and answered. Vague and ambiguous.

THE WITNESS: You have to have a population of clusters and then from that you know what the cluster structure is.

BY MR. KIERNAN:

Q. And what -- what factors do you consider to determine whether you have a population of clusters?

A. Well, it, again, depends on -- the population is -- of clusters is defined once you have determined the sampling scheme. So in other words, if you're sampling schools, clusters of schools.

Q. On page 6 of your declaration -- let me know when you get there.

A. Yes.

Q. It's roughly a quarter of the way down, "Clustering is a property of how the data are collected and has nothing to do with how much variation there is in the underlying population variable or variables."

A. Uh-huh.

Q. Do you see that?

A. Uh-huh.

Q. And what authority supports that statement?

A. Well, it's -- the -- I've given several examples. It's -- the authority is basically that you

can't really anticipate every time somebody is going to get something wrong like this, so I tried to explain through examples that if you -- if you have a variable that doesn't change very much in a population and you draw a random sample from it, that has -- how much variation there is in the population has nothing to do with whether you have to treat those observations as being from a cluster sample.

MR. KIERNAN: Okay. Move to strike as nonresponsive.

MS. SWEENEY: And I disagree with that characterization.

BY MR. KIERNAN:

Q. For the statement, "Clustering is a property of how the data are collected and has nothing to do with how much variation there is in the underlying population variable or variables," stated on page 6 of Wooldridge 1, please cite for me authority that supports that proposition.

A. I can't give you a citation for that.

Q. And is it your testimony today that that statement is generally accepted in the field of econometrics?

A. Yes, I believe it would be.

Q. And can you cite to any authority, any peer-

reviewed publications, that support your testimony that that is generally accepted in the field of econometrics?

A. I -- I'm sorry. Could you repeat that? That sounded like the same question to me.

Q. Slightly different. And that is, can you cite to any authority, including a peer-reviewed publication, that support your testimony that your statement on page 6 is generally accepted in the field of econometrics?

A. Well, here's what I can do: I can point to the literally thousands, if not tens of thousands, of papers published in empirical economics that never discusses whether the variation and the dependent variable has any bearing on whether to cluster or not. So I would think it would show up somewhere if that were actually an issue.

So many labor economists have done analyses with all kinds of response variables, including, as I said, variables such as do you have a job or not, and that has much less variation because it's a 01 variable than if you look at their annual earnings. And the fact that one is much less variable than the other has nothing to do with whether you should cluster the data.

Q. And it's your testimony that labor econometricians, it's generally accepted, that they

would not use clustering in the scenario that you just described?

A. If -- if you ran- -- if you had a random sample, I certainly hope not. Again, you could only cluster if after you obtained your data you define some clusters to -- to cluster on and that would lead to an increased bias in your standard errors.

Q. Okay. Name for me the ones that you can recall, as you sit here, the literally thousands, if not tens of thousands, of papers published empirical economics that never discuss whether the variation in the underlying population variable or variables --

A. Well, I certainly --

Q. -- is relevant with respect to clustering?

MS. SWEENEY: Yeah, I'm going to object. That's sort of a ridiculous question. He's not going to sit here and identify tens of thousands of articles for you.

BY MR. KIERNAN:

Q. Could you list, as you sit here today, tens of thousands of articles?

MS. SWEENEY: Objection. Improper --

THE WITNESS: As I sit here right now, no.

BY MR. KIERNAN:

Q. Okay. Name -- name five for me.

A. Well, look, I follow empirical work and --

Q. Just name five.

MS. SWEENEY: Objection.

MR. KIERNAN: Okay.

MS. SWEENEY: Asked and answered. Argumentative.

BY MR. KIERNAN:

Q. Name one.

MS. SWEENEY: Harassing the witness. Stop now.

THE WITNESS: Angrist -- Angrist and Krueger, their paper on estimating the effects of schooling on wages.

BY MR. KIERNAN:

Q. Okay. Any others you can think of?

A. Melon's paper on evaluating a job training program, an AER paper. The thing is that they don't discuss this issue because it doesn't come up.

Q. What issue doesn't come up?

A. The fact that there's -- how much variation there is in their data, whether that leads to a clustering problem or not.

MR. KIERNAN: Let's take a short break.

MS. SWEENEY: Are you almost done?

THE VIDEOGRAPHER: Going off the record at

2:43 p.m.

(Recess.)

THE VIDEOGRAPHER: Okay. We're back on the record at 3:04 p.m.

BY MR. KIERNAN:

Q. Dr. Wooldridge, do you agree with the statement, "It is probably a sensible rule to at least consider the data as being generated as a cluster sample whenever covariates at a level more aggregated in the individual units are included in an analysis"?

A. That sounds like something I wrote.

Q. And do you agree with it?

A. Actually, I would want to re-examine that statement in light of this sort of recent research that I've done. It certainly is -- when you have variables that are defined at, like, a school district level and you have schools, you may or may not have to cluster. But then when you do it or you don't do it, you can see the answer, assuming that you have a large number of clusters with relatively few units per cluster. But it may be too -- it may be too conservative to do that.

Q. And, as you sit here today, do you stand by your statement that "It's a sensible rule to at least consider the data as being generated as a cluster sample whenever covariates at a level more aggregated than the

individual units are included in an analysis"?

A. If you -- I mean, ideally you would know the level of which the data were clustered, so that's -- that's a case -- that's a conservative approach to the problem where you might not know how the data were generated.

Q. And so today do you stand by your statement in your book?

A. Actually, I would -- as I said, in light of these sort of new findings on clustering random samples, I would actually want to revisit that to see whether that's a -- a useful thing to do or not. And it has to do with -- it's always going to have to do with the number of clusters that you have and the number of observations per cluster.

Q. And what would you want to examine in analyzing or considering whether or not to revise the statement in your book that we've been discussing?

A. Well, I would want to work out the theory for what happens when you assume the data are generated from a random sample, but you include covariates that are defined at a higher level.

Q. Anything else?

A. Simulation.

Q. In Professor Noll's analysis of the iPod

transactional data, did he include covariates at a level more aggregated than the individual unit transactions?

A. Well, each -- each variable -- no, each transaction is defined by its characteristics.

Q. Okay. So you're understanding is that Professor Noll's analysis of the iPod transaction data, he did not include covariates at a level more aggregated than the individual transactions? Is that your understanding?

A. No. He included time effects and he included attributes of the products.

Q. Okay. And is it your testimony that time effects --

A. Those are --

Q. -- and attributes of products are covariates at a level more aggregated than the individual unit transactions?

A. They're not more aggregated. An example would be to say that because some people have the same level of education that -- that how -- is somehow a variable that's aggregated -- that's defined at a more aggregated level because you can put everybody into a class of education.

Q. Okay. And either we're not communicating well -- well, let me -- I just want to ask this to make

sure I have this clear.

Did Professor Noll's analysis of iPod transactional data include covariates at a level more aggregated in individual iPod transactions?

A. Not the way I would define them, no.

Q. And when you say "define them," what is the "them" referring to?

Q. Did you examine Professor Noll's analysis to determine whether he included covariates at a level more aggravated than the individual units that are included in the analysis?

A. Did I examine the regressions? I did look --

Q. For that purpose.

A. I did look at the variables that were included

in the regression, yes.

Q. Do you agree that in some cases you can define the clusters to allow additional spacial correlation?

A. Spacial correlation? Well, spacial correlation has to do when you have -- is usually a feature where you have large geographical units and you don't have random sampling.

Q. So, for example, if you think of sampling fourth grade classrooms and you're concerned about correlation in student performance not just within the class, but also within the school, then you could define the clusters to be the schools?

A. Not if you take a random sample of fourth graders from the population of fourth-grade classrooms. That would be an example of what I'm calling ex post clustering because you would then be, essentially, looking at the school that the classroom came from and making that your cluster when you already have a random sample of fourth-grade classrooms so you don't need to do anything further.

(Exhibit 2 marked.)

BY MR. KIERNAN:

Q. Okay. Let me hand you what's been marked as Wooldridge 2.

And I'll represent to you, Dr. Wooldridge,

that this is a copy of the section on Clustering Sampling, 20.3, from Chapter 20 of your textbook Econometric Analysis of Cross Section and Panel Data, Second Edition.

Do you recognize the section on Cluster Sampling?

A. Yes, I do.

Q. Okay. And if you turn to page 864 --

A. Okay.

Q. -- and this is under Section 20.3.1 --

A. Uh-huh. Okay.

Q. -- and in this section you're discussing an example which a random sample fourth-grade classrooms is drawn in the state and the common factor affecting students in a given classroom is the characteristics of the teacher.

A. Uh-huh.

Q. Is that right?

A. I'm sorry. Where are you seeing that?

Q. In this section where you're using the sample of fourth-grade classroom, my under- --

MS. SWEENEY: Go ahead and take a second to read it.

MR. KIERNAN: Yeah.

THE WITNESS: Are you looking at the back?

Where are you -- I'm sorry.

BY MR. KIERNAN:

Q. I'm looking at Clustering Sampling.

A. This page.

Yeah, I might rethink that now.

Q. Why is that? Well, rethink what?

A. In other words, whether you actually have to cluster at the school level --

Q. So in --

A. -- after doing --

Q. -- your textbook -- oh, sorry.

A. After doing more recent analysis, yes.

Q. Okay. In your textbook you propose clustering when calculating the --

A. So this is a conservative thing to do, yes. That doesn't mean that if you have good reason not to do it, that you -- that you should still do it.

Q. Where do you state it's a conservative -- conservative thing to do and if you have a reason not to do it, you shouldn't do it? Where is that in your textbook?

A. Well, we -- as I mentioned, this is a learning process, right? So after examining these problems with this ex post clustering, I now know that it's a conservative thing to do. So, yes, I would probably --

I would rewrite this section a bit if I -- if there's a third edition coming.

Q. And did you make the same recommendation in the first edition of your textbook, Econometric Analysis of Cross Section Panel Data?

A. Actually, that's -- I can't remember. It was a fairly extensive revision of the book.

Q. And you state, "After examining these problems with this ex post clustering," are you referring to in connection with this case?

A. No, just in general. Just the theory that I've worked out and that, as I mentioned, my coauthors and I had been working on.

So, in other words, when you realize that if you are randomly sampling any unit, whether it's an individual student or a fourth-grade classroom, it's actually -- it's certainly conservative to compute the cluster robust standard errors and it's not going to be very costly in a case like this if you don't have to because you have a large number of clusters with relatively small cluster sizes. But if you do it in the case where you don't -- where your cluster sizes are very large and so there's no theory to tell you that those clusters standard errors are going to settle down to the usual ones, then I would be more careful here and

argue that you -- you shouldn't necessarily do it because your inference could be much too conservative.

Q. Okay.

A. This is the -- the point of -- there's another section in here on stratified sampling. And it's the same sort of argument that you can always do something that's very conservative, but you may not learn much and if you can do something else that is -- is actually providing the proper standard errors, then you should do that.

Q. And have you reached an ultimate conclusion of whether to revise the paragraphs in your Chapter 20.3.1?

A. Yes. I probably -- I think I would revise them in light -- in fact, I would add a section on ex post clustering.

Q. And what -- what authorities or peer-reviewed papers would you cite in support of your new section in your textbook?

A. Well, a lot of this book is actually based on original research, so I probably wouldn't. If I finish the work with my co-authors, then I would cite that.

Q. And, as of today, you have not completed that work?

A. The theory is -- is essentially finished --

Q. And --

A. -- and some -- and some simulation results.

Q. Have you completed the simulation result -- work?

A. Simulation work is -- it's hard to decide where to stop, but there's -- there's a fair amount of simulation work.

Q. Okay. Have you completed the empirical work with respect to your theory on ex post clustering?

A. So by "empirical" do you mean with an actual dataset or do you mean the simulations?

Q. The simulations. I know this morning you were describing the simulations as the empirical work like --

A. Oh.

Q. -- what Hansen was doing.

A. Okay. So just to be clear, when economists say "empirical work," they usually mean data that's been collected from the real world --

Q. Sure.

A. -- as opposed to generating. So if you mean the simulations experiments, is it completed? Well, you can always -- you can always vary parameters and see how things change when you vary parameters, but the simulations predict the theory quite well.

Q. And have the simulations been peer reviewed?

A. No.

Q. If you turn to page 868, you have an example 20.3. Just tell me when you get there.

A. Yes.

Q. And this is Cluster Correlation in Teacher Compensation. Do you see that?

A. Uh-huh.

Q. And "The data set is in BENEFITS.RAW, includes average compensation, at the school level, for teachers in Michigan."

Do you see that?

A. Yes.

Q. And do you understand that that data include the entire population?

A. Yes.

Q. Okay. And then in your textbook you state, "We view this as a cluster sample of school districts, the schools within districts representing the individual units"; is that accurate?

A. It's an example, yes.

Q. Okay. And this is an example of what you described on page 864?

A. Actually, this is -- yeah, this is just an example to, essentially, create a -- what could have been a cluster sample so that they can see what happens when you -- when you do cluster at a -- at a more

aggregate level, but -- and you can see that the standard errors do go up, so it's a conservative thing to do.

Q. And in example 20.3, the -- the data was not collected using cluster sampling; isn't that correct?

A. That's correct.

Q. You collected the entire population?

A. Actually, it's not the entire population, no. It's -- it's a subset of the districts from the state of Michigan. It contains a lot of them, but not all of them.

Q. How many did you exclude?

A. 500 and -- probably about --

Q. Does eight come to mind?

A. Eight? No, I think it's got to be more than that. I think there are currently 500 and 55 -- 18 or 20 or something like that, yeah.

Q. Okay. So virtually all the school districts?

A. It -- very close, yes. Where you have -- yeah. So G equals 537 when these are elementary schools, I believe. So you have a large G with relatively few schools per -- so few observations per cluster.

Q. Now, are you familiar with --

A. And this, by the way, actually fits with the

theory that I -- that the standard errors would go up by a fair amount.

Q. And are you -- other than theory, are you relying upon anything to support that the standard errors would go up by a large amount? Any --

A. The simulations.

Q. Other than the ones that you -- that you've been working on recently, are there other simulations that you're relying upon for that statement?

A. I don't know of any simulations that ask the question what would happen if you simulated -- if you clustered a random sample based on characteristics that you draw along with the main variable, yes. Usually when -- when properties of the clustered standard errors are evaluated via simulation, they are actually cluster samples that have been drawn like in Chris Hansen's work, for example, or the DueFlow, et al., paper in the QJE.

Q. You state in your report that clustered standard errors are not justified with, say, ten clusters and 200 observations per cluster.

Do you recall that page 5 of your report?

A. Uh-huh.

Q. And I notice you don't have a citation there. Can you cite to me --

A. I should --

Q. -- the authority --

A. I should have put a citation in there. I'd probably have to go look that up. There's a paper by, I think it's Mitchell Peterson. It's in a financial journal. I'd have to go -- Journal of Financial, maybe it's Journal of Finance.

Q. And what is a -- the conclusion or opinion set forth in that paper with respect to this issue?

A. That clustering can be effective for computing standard errors when you have a large number of clusters and not too many units per cluster. That's actually a paper that the setting is a bit different because it's -- it's panel data, so there is a time dimension, and so they're considering the case of clustering both in the cross-sectional dimension in some cases and the time series and others and then across both dimensions and others.

I could probably -- I -- I'd have to think. I -- I mentioned this before, that there's no sort of absolute rule of thumb that you can use, but the theory certainly doesn't allow for that kind of configuration. If you think that thinking of G heading off to infinity with the number of observations per cluster fixed is a good thought experiment, it isn't because the -- the

large sample theory that's obtained from letting G get large is not going to be very relevant for that particular structure.

Q. So are you relying upon Hansen's paper for the statement that clustered standard errors are not justified with ten clusters in 200 observations?

A. He considers a similar configuration. I don't know if it's exactly that configuration. In fact, he may have fewer observations per cluster and shows that they don't work as the theory -- as -- as the large cluster theory says they should.

Q. When was the last time you reviewed Hansen's paper, 2007 paper?

A. Ah, it has been a little while. Well, actually, no. I -- I just looked at it the other day, but now I can't remember what the -- yes. So I'd have to go back and look at that more carefully.

Q. Did you review it before submitting this report? In connection with drafting this report, did you review it?

A. I reviewed -- I -- I reviewed the lecture notes that I've written that refer to his report, his paper.

Q. Did you review any of the simulations that were included in tables 1 through 4 in his -- in his

paper?

A. I thought I looked at them, yes.

Q. Okay. And is it your testimony that Chris Hansen stated that the higher the ratio of observations per cluster to number of clusters the more poorly clustered standard errors --

A. I'm not sure he said that, no.

Q. Is it your testimony that his paper supports that conclusion in your report?

A. I'm not sure it supports that statement. He does -- he does show that the performance deteriorates as for a fixed number of clusters you get more observations per cluster, yes.

Q. What deteriorates?

A. The --

Q. Performance of what? Oh, sorry.

A. Sorry. The performance of the standard errors.

Q. Under what -- under what calculation? So you recall Hansen does OLS clustered and random. Under which does it perform more poorly as you increase the number of observations --

A. Well --

Q. -- per cluster as you keep number --

A. All that's --

Q. -- number of clusters constant?
A. All that's relevant for this is the OLS versus clustering.
Q. Right.
A. Right. And clustering -- well, no, but, see, in the -- in the OLS case, there -- if you're talking about computing the usual standard errors, he has built cluster correlation into his simulations -- that was my comment earlier -- whereas my simulation said, suppose we take a random sample and then we do clustering, he's built that into his analysis so that there is cluster correlation. And so neither of them works very well when you have a small number of clusters and many observations per cluster.
Q. Your testimony is that's what Chris Hansen states in his paper --
A. That's what --
Q. -- that as you increase -- that as you increase the number of observations per cluster, keeping cluster -- the number of clusters constant that OLS performs more poorly and so does the clustered standard errors?
A. Yes.
MS. SWEENEY: And I -- I don't think the court reporter caught the witness's first answer because the

examiner, again, talked over him.
Do you want to repeat what you had said before, before David interrupted you, if you can remember it.
THE WITNESS: Oh, the -- I -- I think it was about -- talking about three different versions of the standard errors, did you say, OLS, clustered, and you said something about random --
BY MR. KIERNAN:
Q. Random effects.
A. -- random effects. So, yeah, that -- the issue here is the OLS versus the clustered standard errors and the OLS does poorly, but that's because cluster correlation has been built into the simulation.
Again, it's a -- it's a different situation because he's dealing with panel data and so the cluster correlation is actually what we call serial correlation in the errors across time.
Q. Now, Dr. Wooldridge, you cite to Hansen -- Hansen's paper as supporting your opinions in this case; right?
A. Yeah. Uh-huh. Yes.
Q. Okay. Even though he's using panel data?
A. Yes, because the panel data -- the panel data introduces correlation in the cluster. It introduces it

of a specific form, but it still has the same effect that it induces correlation within a unit in the cluster.
MR. KIERNAN: Let's just do this.
I will hand you what is Exhibit 3. Is that right?
(Exhibit 3 marked.)
BY MR. KIERNAN:
Q. Can you identify Exhibit 3?
A. Yes.
Q. And what is this?
A. This is the -- this is Chris Hansen's 2007 paper on clustering.
Q. And is this the paper that you're relying upon in your declaration, Wooldridge 1?
A. Somewhat, yes.
Q. Okay. And if you turn to pages 612 to 615 --
A. Uh-huh.
Q. -- you'll see the simulations that you've been discussing.
A. Uh-huh.
Q. And if -- pardon me.
Okay. I just want to walk through this. So if I look down at table one, N equals 10 and T equals 10, is it your understanding that N represents a number

of clusters?
A. Yes.
Q. And T represents number of observations per cluster?
A. Yes.
Q. Okay. And then column four, that's the target standard error?
A. Yes.
Q. Okay. And then the difference between two and four -- columns two and four is how well the estimator is computing the standard errors?
A. Yes.
Q. And, therefore, the difference shows you the bias of the estimated standard errors?
A. Yes.
Q. Okay. And if we look down at row -- where it says B. Random Effects --
A. Right.
Q. -- and let's consider the case where the intergroup correlation is high, so it's 0.9.
A. Okay. Uh-huh.
Q. Okay. And if we look at OLS -- compare OLS to clustering, which performs better?
A. The clustering --
MS. SWEENEY: Objection. Vague and ambiguous.

I'm sorry. Go ahead.
BY MR. KIERNAN:
Q. Go ahead, Dr. Wooldridge.
A. The clustering performs better because the cluster effect is left in the error term for OLS, so this is data that had been generated with a cluster affect.
Q. And then in table two Chris Hansen keeps the number of clusters constant at ten; correct?
A. Uh-huh.
Q. And he increases the observations to 50 observations per cluster; correct?
A. Yes.
MS. SWEENEY: I'm sorry. Where are you?
MR. KIERNAN: I'm on table two.
MS. SWEENEY: Any particular place in table two?
MR. KIERNAN: No.
MS. SWEENEY: No? Okay. Sorry.
MR. KIERNAN: That's all right.
BY MR. KIERNAN:
Q. Okay. And in looking at the same case where the intergroup correlation is high, so it's 0.9 under Random Effects, here this shows that when the number of observations per cluster increase, keeping the number of

clusters constant, OLS performs even worse; isn't that correct?
A. The OLS standard error performs worse, correct.
Q. Okay. And how about the cluster standard errors?
A. It -- they're performing a little better.
Q. Okay.
A. And that's -- but in this data-generating mechanism, the cluster is of a form where it's serial correlation that's dying out over time, not a -- not what you would get if you were using randomly sampled data and then clustering after you've randomly sampled.
Q. And then if you turn to page, look at Table three, now Dr. Hansen increases the number of clusters or group, from ten to 50.
A. Uh-huh.
Q. But uses ten observations per cluster. Do you see that?
A. Yes.
Q. And then using the same case where the intergroup correlation is at 0.9 --
A. Uh-huh.
Q. -- which performs better, the OLS or the cluster?

A. The cluster.
Q. And then in table four, Dr. Hansen runs a simulation again, keeping the number of clusters constant at 50, but doubles the number of observations per cluster.
Do you see that?
A. Yes.
Q. And what impact does that have on the performance on the OLS estimator?
A. It deteriorates.
Q. And what impact does that have on the cluster robust estimator?
A. So in this case it gets a little better.
Q. Roughly -- do you recall roughly how many variables Dr. Noll had in his regressions?
A. Well, it's two -- two pages' worth. So I don't know. Maybe -- maybe it's more than two pages' worth. Must be 50 or 60, something like that.
MR. KIERNAN: Okay. I'm going to hand you what is his rebuttal report. Okay? I'm going to mark this as Exhibit Wooldridge 4.
MS. SWEENEY: So which one are you marking as 4? Just the --
MR. KIERNAN: Just the -- yeah, these, but I've provided the --

MS. SWEENEY: Okay. Thank you.
MR. KIERNAN: -- entire rebuttal at your request.
MS. SWEENEY: Okay.
(Exhibit 4 marked.)
BY MR. KIERNAN:
Q. And you can take a moment to see that Exhibit 4 is the -- Dr. Noll's reseller and direct consumer regressions, the reports in Exhibit 3 of his rebuttal report.
Do you see that?
A. Yes.
Q. And, roughly, how many variables does Dr. Noll include in his regressions?
A. Let's look. It's more like 100, roughly.
MS. SWEENEY: Do you want him to count them on this page?
THE WITNESS: Eighty, something like that.
BY MR. KIERNAN:
Q. Okay.
A. Yeah.
Q. And then -- just a second. I lost my copy.
MR. KIERNAN: Oh, there it is. Thank you.
BY MR. KIERNAN:
Q. And you'll notice that the vast majority are

statistically significant at the one percent level.

A. Yes.

Q. Okay. Is it -- is it unusual to find this level of significant over so many variables in a regression?

MS. SWEENEY: Objection. Overbroad. Vague and ambiguous.

THE WITNESS: It's unusual to have 2 million observations.

BY MR. KIERNAN:

Q. And not my question.

A. And to have -- and to have -- so -- so is it unusual? The -- often we don't have good explanatory variables for micro-type outcomes. So if this were a wage equation and we only had, you know, a dozen characteristics of people to explain their wage, then I would expect much more residual variance. But if we had two million observations, we can still get quite small standard errors and statistical significance.

I mentioned some work earlier by Angrist and Krueger and Angrist has also done work with Bill Evans, using five percent census data and, yeah, you -- you get small standard errors when you have large datasets like that.

Q. And if you look at Exhibit 3B -- so we were

looking at 3A -- 3B's the direct sales regression --

A. Uh-huh.

Q. -- and if you look at the regression output, Dr. Noll reports that every single coefficient is statistically significant at the one percent level.

Do you see that?

A. Yes.

Q. Do you find it unusual or is it unusual that a regression with this many variables could -- or strike that.

If you have a regression with close to 80 variables, could it all be considered near perfect variables for explaining iPod prices?

MS. SWEENEY: Objection. Vague and ambiguous.

THE WITNESS: Yes. I'm not sure what you mean by "near perfect." If you mean that statistically significant at a low significance level, then, again, I'm not surprised with this very large sample size.

This is the difference between -- so without commenting on the coefficients, the difference between practical significance and statistical significance. If you have -- even if you have really small coefficients with enough data you can drive the standard error to be close to zero and so, no, it's not surprising to me.

BY MR. KIERNAN:

Q. Okay. And if you look at -- if you look at Exhibit 3B and let's take the harmony2 variable, do you know what that refers to?

A. I believe there was a second version of the harmony software and that's what that indicates.

Q. Okay.

A. It's a 01 variable indicating when harmony2 was released, I believe.

Q. And what would be the T statistic on the harmony2 variable in the regression represented in 3 here?

A. You're asking me to do calculations that I'm not very good at, so maybe you have done it for me.

Q. How would you calculate the --

A. Oh, I would --

Q. -- T statistic?

A. Yeah, I would take the coefficient estimate and divide by the standard error. Actually, I would get it from the Stata output probably because that would be -- yes.

Q. And if you had a T statistic, let's say, 25, what would that tell you? How would you interpret that, a T statistic of 25?

A. It's a strong rejection that the coefficient is equal to zero.

Q. And if the T statistic -- strike that.

A. These are much bigger than that.

Q. I know.

A. But that's, again, with 36, 37 million observations, it's rare that one has a dataset with that many observations. And, as I said, it's -- it's also because these are such good predictors of price you have little residual variance to explain.

Q. And how do you know that they are such good predictors of price? What are you basing that statement on?

A. The R-squared.

Q. Anything else?

A. No, because you could have statistical significance even at a very high level and not have necessarily a high adjusted R-squared. There's nothing that says there has to be any particular relationship between those.

As I mentioned in the Angrist and Krueger work, they have a fair amount of residual variance left over, but because they're using the 5 percent census, they still get large T statistics.

(Exhibit 5 marked.)

BY MR. KIERNAN:

Q. Let me hand you what's been marked as

Page 145

Wooldridge 5. Can you identify Wooldridge 5 for me?
A. This is a set of lecture notes for the basis of a set of lectures that Guido Imbens and I gave at the National Bureau of Economic Research in the summer of 2007.
(Exhibit 6 marked.)
BY MR. KIERNAN:
Q. Okay. I will hand you what's been marked as Wooldridge 6.
Can you identify Wooldridge 6 for the record?
A. This is also a set of lecture notes. This -- I should have said, the first set is on estimating average treatment effects under unconfoundedness and these are lecture notes on linear panel data models for the series of lectures.
Q. In your report on page 5, if you look at the second full paragraph --
A. Uh-huh.
Q. -- about two-thirds of the way down you state, "The higher is the ratio" -- "The higher the ratio of observations per cluster to number of clusters, the more poorly clustered standard errors behave."
Do you see that?
A. Yes.
Q. What support do you have for that statement?

Page 146

A. Well, I would have to go back and look at the papers that I've read that have done simulations on this. They don't -- this is a general statement. So you could -- you could find -- in other words, general patterns. You could find specific simulations for given number of clusters and observations. This might not be true, but it's -- generally -- it's a general statement about the patterns you would observe across lots of simulations as you get more and more observations per cluster.
Q. And, as you sit here today, can you identify any work that supports that statement?
A. Well, I -- as I -- any -- any published work? No. I've done my own simulations that -- that show the standard errors become -- are -- are conservative when you -- when you get -- well, with -- with a fixed and relatively small number of clusters, yes. In -- in the cases I have looked at very conservative, but that's a -- that's an issue of inappropriate clustering. There's a separate issue of how would the clustered standard errors behave if -- if you had a cluster sample where you just had a small number of clusters with many observations per cluster.
Q. And you state, "The higher is the rate -- "The higher is the ratio," that should read, "The higher the

Page 147

ratio."
And at what point -- at what ratio does the clustered standard errors start performing more poorly?
A. You can't -- this is a question you can't really know because it depends so much on the specifics of the simulation. So it would depend whether it's a panel dataset or a true cluster sample. It would depend on whether you've applied clustering to a random sample. There are all these things that would come into play.
Q. Okay. And in this paragraph, you're describing that if clustering were legitimate in this case --
A. Yes.
Q. -- and you're stating that it's not legitimate --
A. Right.
Q. -- but you're assuming here that even if it were legitimate, there's another problem which is the ratio of observations to clusters per -- to observations per cluster could be too high.
A. Uh-huh.
Q. Is that something that you've examined in this case?
A. Not in this specific case, no.
Q. So you're stating here it could be a

Page 148

problem --
A. Uh-huh.
Q. -- but you haven't reached an opinion on that?
A. That's -- that's correct.
MR. KIERNAN: Let's go off the record, give me five minutes.
MS. SWEENEY: Okay.
THE VIDEOGRAPHER: Going off the record at 3:53 p.m.
(Recess.)
THE VIDEOGRAPHER: Back on the record at 4:07 p.m.
(Exhibit 7 marked.)
BY MR. KIERNAN:
Q. I will be handing you what has been marked as Wooldridge 7.
And Wooldridge 7 is the paper by Justin Wolfers, "Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results."
Do you recognize this paper?
A. I know of this paper, yes.
Q. In fact, it was the subject of research in the paper that you did with --
A. Yup.
Q. -- Solon and Haider?

THE WITNESS: No.
BY MR. KIERNAN:
Q. Okay. And are you aware -- or do you have any understanding of the datasets that are used by econometricians in antitrust cases involving price fixing?
MS. SWEENEY: Objection. Compound. Overbroad. Foundation.
MR. KIERNAN: Didn't like that one.
THE WITNESS: No, I'm not -- I'm familiar with the dataset that Professor Noll analyzed, at least the description of it.
BY MR. KIERNAN:
Q. In antitrust cases involving where there's allegations by plaintiffs that some conduct impacted price, is it unusual for the econometricians in such cases to have the entire population of data?
MS. SWEENEY: Objection. Foundation. Compound. Overbroad. Vague and ambiguous.
THE WITNESS: I -- I don't know.
BY MR. KIERNAN:
Q. Is it your opinion that in antitrust cases where the parties are attempting to estimate the impact of the challenge conduct on pricing when they're using the entire population of transactions that clustering

due to common factors -- accounting for clustering due to common factors is never appropriate?
MS. SWEENEY: Objection. Vague and ambiguous and compound. Overbroad.
THE WITNESS: Again, the way I think about this is if you had a large population of data, then you could randomly sample and still have a large number of observations.
The original -- the usual calculation of the standard errors without clustering would be appropriate and as you get more and more data you will find that the standard errors shrink to zero. And that's what I think the appropriate thing that -- the appropriate finding is.
BY MR. KIERNAN:
Q. Have you -- have you reviewed any text, publications on guidelines or recommendations for proving damages in an antitrust case?
A. I was sent Chapter 6 of the book Proving Antitrust Damages --
Q. And do you --
A. -- by the American Bar Association.
Q. And do you recall the recommendation in that text on whether to account for a clustering due to common factors?

A. Yes. It recommended clustering, but that's in the context where -- or it should -- there should have been qualifications that the clustering should be done when it's actually appropriate. The clustering, based on essentially an arbitrary, you know, partitioning of the data after you've looked at it is not appropriate. So there should have been a qualifier in there, I believe.
Q. So you disagree with the proposal in the ABA guidelines?
A. I think it's overly broad, yes. It doesn't discuss the issue at all of taking a random sample and then clustering on some characteristics.
Q. And is one possible reason for that is because in most antitrust cases parties are dealing with the entire population of transactions, the prices from the defendants?
A. I don't believe so.
MS. SWEENEY: Objection.
THE WITNESS: I'm sorry.
MS. SWEENEY: Foundation. Vague and ambiguous.
Sorry.
THE WITNESS: I don't believe so, no. For the same reason that I have talked about over and over again

because you could if you wanted to take a large random sample and then you would know that this clustering ex post, as I've called it, is the inappropriate thing to do.
MR. KIERNAN: I'm not going to take two more minutes.
Mark the transcript attorneys' eyes only per the protective order.
Last chance. Not going to ask anything?
That's all I have.
MS. SWEENEY: We don't have anything.
THE VIDEOGRAPHER: Stipulations.
THE REPORTER: Handling of the original? Who will handle the original transcript?
MS. SWEENEY: What have we been doing?
THE REPORTER: Do you want to go off the record?
MR. KIERNAN: Yeah, yeah.
MS. SWEENEY: Yeah, let's go off the record.
MR. KIERNAN: Yeah. Let's go off the record because he doesn't need to hear this.
THE VIDEOGRAPHER: Okay. This concludes the video portion of the deposition. Two DVDs were made. We're going off the record at 4:26 p.m.
(Deposition concluded at 4:26 p.m.)

I, the undersigned, a Certified Shorthand Reporter of the State of California, do hereby certify:

That the foregoing proceedings were taken before me at the time and place herein set forth; that any witnesses in the foregoing proceedings, prior to testifying, were duly sworn; that a record of the proceedings was made by me using machine shorthand, which was thereafter transcribed under my direction; that the foregoing transcript is a true record of the testimony given.

Further, that if the foregoing pertains to the original transcript of a deposition in a federal case, before completion of the proceedings, review of the transcript [ ] was [ ] was not requested.

I further certify I am neither financially interested in the action nor a relative or employee of any attorney or party to this action.

IN WITNESS WHEREOF, I have this date subscribed my name.

Dated: January 10, 2014

Debby Gladish

Debby M. Gladish
RPR, CLR, CCRR, CSR No. 9803
NCRA Realtime Systems Administrator

DECLARATION UNDER PENALTY OF PERJURY

Case Name: The Apple iPod iTunes Anti-Trust Litigation
Date of Deposition: 01/06/2014
Job No.: 10009202

I, JEFFREY WOOLDRIDGE, PH.D., hereby certify under penalty of perjury under the laws of the State of _____ that the foregoing is true and correct.

Executed this _____ day of _____, 2014, at _____.

_____

JEFFREY WOOLDRIDGE, PH.D.

DEPOSITION ERRATA SHEET

Case Name: The Apple iPod iTunes Anti-Trust Litigation
Name of Witness: Jeffrey Wooldridge, Ph.D.
Date of Deposition: 01/06/2014
Job No.: 10009202

Reason Codes: 1. To clarify the record.
2. To conform to the facts.
3. To correct transcription errors.

Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____

DEPOSITION ERRATA SHEET

Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____
Page _____ Line _____ Reason _____
From _____ to _____