UNITED STATES DISTRICT COURT
SOUTHERN DISTRICT OF TEXAS
CORPUS CHRISTI DIVISION

DRAFT COPY

MARC VEASEY, ET AL.,	)	CASE NO: 2:13-CV-00193
	)	
Plaintiffs,	)	CIVIL
	)	
vs.	)	Corpus Christi, Texas
	)	
RICK PERRY, ET AL.,	)	Wednesday, September 3, 2014
	)	(7:59 a.m. to 12:11 p.m.)
Defendants.	)	(1:10 p.m. to 5:49 p.m.)

BENCH TRIAL - DAY 2

BEFORE THE HONORABLE NELVA GONZALES RAMOS,
UNITED STATES DISTRICT JUDGE

Appearances:	See Next Page
Court Recorder:	Genay Rogan
Clerk:	Brandy Cortez
Court Security Officer:	Adrian Perez
Transcriber:	Exceptional Reporting Services, Inc. P.O. Box 18668 Corpus Christi, TX 78480-8668 361 949-2988

Proceedings recorded by electronic sound recording;
transcript produced by transcription service.

APPEARANCES FOR:

Plaintiffs:	**CHAD W. DUNN, ESQ.** **KEMBEL SCOTT BRAZIL, ESQ.** **Brazil and Dunn** **4201 Cypress Creek Parkway, Suite 530** **Houston, TX 77068** **ARMAND DERFNER, ESQ.** **P.O. Box 600** **Charleston, SC 29402** **J. GERALD HEBERT, ESQ.** **Attorney at Law** **191 Somervelle Street #405** **Alexandria, VA 22304** **NEIL G. BARON, ESQ.** **914 FM 517 Rd. W, Suite 242** **Dickinson, TX 77539** **LUIS ROBERTO VERA, JR., ESQ.** **111 Soledad, Suite 1325** **San Antonio, TX 78205** **EMMA P. SIMSON, ESQ.** **Campaign Legal Center** **215 E. Street NE** **Washington, DC 20002**
Mexican American Legislative Caucus, et al.:	**EZRA D. ROSENBERG, ESQ.** **Dechert, LLP** **902 Carnegie Center, Suite 500** **Princeton, NJ 08540-6531** **MARK A. POSNER, ESQ.** **AMY L. RUDD, ESQ.** **LINDSEY COHAN, ESQ.** **JENNIFER CLARK, ESQ.** **Lawyers' Committee for Civil Rights** **1401 New York Ave. NW, Suite 400** **Washington, DC 20005**

APPEARANCES FOR: **(CONTINUED)**

United States of America:	**RICHARD DELLHEIM, ESQ.** **ELIZABETH S. WESTFALL, ESQ.** **ANNA BALDWIN, ESQ.** **PAMELA CARLIN, ESQ.** **U.S. Department of Justice** **950 Pennsylvania Ave. NW** **Washington, DC 20530** **BRUCE I. GEAR, ESQ.** **Department of Justice** **1800 G Street NW** **Washington, DC 20006**
Ortiz Plaintiffs, et al.:	**JOSE GARZA, ESQ.** **7414 Robin Rest Dr.** **San Antonio, TX 78209** **ROBERT W. DOGGETT, ESQ.** **Texas Rio Grande Legal Aid, Inc.** **4920 North IH 35** **Austin, TX 78751** **MARINDA VAN DALEN, ESQ.** **Texas RioGrande Legal Aid, Inc.** **531 E. St. Francis** **Brownsville, TX 78520**
Texas League of Young Voters Education Fund:	**RYAN HAYGOOD, ESQ.** **NATASHA KORGAONKAR, ESQ.** **NAACP Legal Def. and Educational Fund** **40 Rector St., 5th Floor** **New York, NY 10006**
Also present:	**Imani Clark** **DANIELLE CONLEY, ESQ.** **KELLY DUNBAR, ESQ.** **Wilmer Cutler Pickering, et al.** **1875 Pennsylvania Avenue, NW** **Washington, DC 20006**

APPEARANCES FOR: **(CONTINUED)**

Texas Association of Hispanic County Judges and County Commissioners:	**ROLANDO L. RIOS, ESQ.** **115 E. Travis** **Suite 1654** **San Antonio, TX 78205**
Also present:	**ROGER GALVAN, County Commission Calhoun County**
State of Texas:	**JOHN BARRET SCOTT, ESQ.** **Scott, Yung, L.L.P.** **208 N. Market Street** **Suite 200** **Dallas, TX 75202**
	JOHN REED CLAY, JR., ESQ. **LINDSEY E. WOLF, ESQ.** **JENNIFER ROSCETTI, ESQ.** **G. DAVID WHITLEY, ESQ.** **STEPHEN L. TATUM, JR., ESQ.** **STEPHEN R. KEISTER, ESQ.** **Office of the Attorney General** **P.O. Box 12548** **MC001** **Austin, TX 78711**
	ARTHUR D'ANDREA, ESQ. **Office of the Attorney General** **209 W. 14th Street, 7th Floor** **Austin, TX 78701**
	BEN A. DONNELL, ESQ. **Donnell Abernethy Kieschnick** **555 N. Carancahua, Suite 400** **Corpus Christi, TX 78401**
	WHITNEY DEASON, ESQ.

INDEX

Corpus Christi, Texas; Wednesday, September 3, 2014; 7:59 a.m.

(Call to order)

THE COURT: Court calls Cause Number 2-13-193, *Veasey, et al. versus Perry, et al.* Are you ready to proceed, Plaintiffs?

MR. DERFNER: Yes, your Honor.

THE COURT: Defense?

MR. DERFNER: I'm Armand Derfner on behalf of the Veasey Plaintiffs. With me here is Emma Simson. And if Mr. Scott and we could approach the bench for a minute?

THE COURT: Okay.

(Begin bench conference at 7:59 a.m.)

MR. DERFNER: Good morning, your Honor.

THE COURT: You all want to speak in here. Good morning.

MR. DERFNER: You're probably aware that I have severe visual impairment or whatever. In fact, I'm legally blind, although I pretend not to be. And so for that reason, if it's possible, Ms. Simson, who's a lawyer on our team, may be assisting me, especially with some of the exhibits and that sort of thing. And I --

THE COURT: That's perfectly fine.

MR. SCOTT: Absolutely.

MR. DERFNER: And also -- I also -- that means I can't see gestures. If somebody makes a dirty face at me, I

won't recognize it. And if you give me a dirty look, I won't know to stop, so if you -- if I'm doing something I shouldn't, please say something.

THE COURT: I will.

MR. DERFNER: Thank you very much.

(End bench conference at 8:00 a.m.)

THE COURT: All right. So are we ready to proceed with the Plaintiffs' next witness?

MR. DERFNER: Plaintiffs would call Michael Herron, please.

THE COURT: Good morning, sir. Would you raise your right hand?

MICHAEL HERRON, PLAINTIFFS' WITNESS, SWORN

DIRECT EXAMINATION

BY MR. DERFNER:

Q Please give your name and your address and your position, please.

A My name is Michael Charles Herron. My address is 6 Bridgeman Road in Hanover, New Hampshire, and I'm a Professor of Government at Dartmouth College.

Q How long have you been a professor at Dartmouth College?

A I believe this is my eleventh year.

Q And what is your department there?

A The Department of Government.

Q And do you hold -- what is your faculty rank?

A I hold an endowed chair at Dartmouth.
Q And the name of the chair?
A It's the William Clinton Story Remsen 1943 Professor of Government.
Q And what are your teaching or research specialties?
A I teach applied statistics and I teach some game theory as well.
Q And have you published articles in your field?
A Yes.
Q Could you give me some of the journals of some of your recent articles, especially the refereed journals?
A Sure. I regularly publish in referee journals. I published in the -- the general journals in my field, for example, are *American Political Science Review* and *American Journal of Political Science*. And I also have published and regularly do in the specialty journals in my field, for example, *Election Law Journal*.
Q And do you have a specialty in election law or elections and voting?
A I would say it's studying of election administration, voting irregularities, and so forth.
Q And in the course of your work and you research, have you had occasion to deal with analysis of databases and working with large quantities of database and computer records?
A Yes. I regularly do that in the course of my research.

Q And do you publish -- do your publications also include research which involves that kind of work?
A Yes, regularly.
Q And have you been involved in litigation as a consultant or as a retained expert?
A Yes.
Q And has some of those also involved database analysis and statistical methods?
A Yes.
Q In the area of political science?
A Yes.
Q And with the specialty in elections?
A Yes.
Q Have you been accepted as an expert in any cases in that field?
A Yes.
Q Tell me about that.
A I was accepted as an expert in an election contest in Florida in 2006.
Q And what was that case about?
A It was about a disputed election that occurred in the -- in a congressional district, and one candidate was not pleased with the outcome and challenged it.
Q And what was your role as the -- as an expert?
A I analyzed some voting data for a company that made the

voting machines used in the disputed election.

Q And what was the purpose of your analysis?

A The purpose was to see if the data was consistent with the allegations that the machines were broken or was consistent with an alternative explanation that the particular format that voters used when they voted led to the abnormality that basically caused this challenge.

Q And what kind of analysis or analyses did you perform in that case?

A I had what are called ballot-level (phonetic) records, among other things, which are records that explain how individuals cast their votes, and I studied how patterns in these ballot records varied across counties in Florida.

Q And was that a federal court or state court?

A I believe it was a state court.

MR. DERFNER: Your Honor, I tender the witness as an expert in the field of political science and statistics with a specialty in elections and voting.

MR. SCOTT: Your Honor, with the understanding that the Court has asked that we put off *Daubert* challenges until the end of the trial.

THE COURT: That's fine. I will address it at the end.

MR. ROSENBERG: End of the direct, I thought; not end of the trial.

THE COURT: After each expert.

MR. ROSENBERG: Yes.

THE COURT: Okay. So we're going to proceed and then the Court will address it at the end of his testimony, is my understanding.

MR. DERFNER: Thank you, your Honor.

BY MR. DERFNER:

Q Professor Herron, what were you asked to look at in this case? First of all, who retained you in this case?

A Plaintiffs.

Q Okay. And what were you asked to look at or analyze?

A I was asked to take some data that was supplied to me by the Department of Justice and study it with two goals in mind. One is to assess the number of non-matches in this dataset. The dataset refers to the results of an algorithm, and I was asked to study the output of the algorithm and assess the number of non-matches in it. And then I was asked to perform a racial analysis on the data that I received and try to understand whether there are racial differences in the non-matches and the data that I -- was supplied to me.

Q And when you say "matches," what records were you analyzing? Were those the state team records and other what you call database identification records?

A Well, the data that was sent to me incorporates data from the team file that you just mentioned, along with data from

other records, like from the Department of State -- excuse me, from the Texas Department of Public Safety and federal databases.

Q Okay. And what is team data?

A The team database, as we heard yesterday, is the -- is an election administration database maintained in Texas.

Q Okay. As far as -- you've identified two questions. As far as the number of non-matches, what was your overall result?

A My overall result was that the number of non-matches is up to 800,000 approximately.

Q And as far as a result -- and we'll get into the details later -- but as far as the result on the racial composition of the non-matched group, what was your result?

A Generally speaking, my result is that the possession rates of forms of identification varied by racial group. I was asked to look at three different groups: whites, Hispanics, and blacks. And my results, generally speaking, are that the white identification possession rate is greater than the Hispanic possession rate and the black possession -- ID possession rate. And I found that these differences between the white identification possession rate and the possession rates of those other two minority groups, these differences are statistically significant.

Q Okay. You've seen a report of Dr. Ansolabehere and you heard his testimony.

A Yes. I've skimmed his report and I was in this court yesterday.
Q Okay. And you know what his results were as far as the racial composition of the non-match group?
A Yes.
Q And how did the results you found compare to the results that Dr. Ansolabehere found?
A They're effectively the same.
Q Thank you.

MR. DERFNER: Okay, what I'd like to do before we get into the detail of the numbers, frankly I'd like to know a little bit more -- and maybe other people -- other people may be geniuses in the field, but I'd like to know a little bit more about the process we went through. So with your Honor's indulgence, we'd like to take a little bit of time to go into some detail and understand some of these great concepts like ACS and CVAP and ecological regression and batching and so forth.

THE COURT: That's fine.

MR. DERFNER: Pardon?

THE COURT: That's fine. You can proceed.

MR. DERFNER: Thank you.

BY MR. DERFNER:

Q Professor Herron, you started with a database matching process; is that correct?

A I received data that was the output of an algorithm that's a -- with a database matching process, and the algorithm was carried out by, as we heard yesterday, faces (phonetic) of the Department of Justice and other members of the federal government.

Q Okay. Now, what is a database matching process in -- just in general?

A Well, from the context of this case, in that particular context, a database matching process is an exercise of trying to understand whether individuals in one database -- in this case, the team database, these are registered voters in Texas -- if those individuals also appear on other -- in other databases, for example, a driver's license database maintained by the State of Texas.

Q Okay. And so you're trying to see if individuals in the team database which is, as you said, registered voters, are in some other databases. How do you know which -- what the other databases are that you select?

A Well, in this case, those other databases were selected based on the law, SB14. That law specifies the valid forms of voter identification useable in Texas at the present time, and that law dictated that, for example, the team database should be compared to a list of individuals who have passports, because a passport is a valid form of identification, along with driver's license -- a state driver's license and so forth.

Q Okay. Now, you said there's an algorithm, right? A wonderful word. Tell me what an algorithm is.
A Again, in the context of this case, an algorithm is a set of rules that explains that -- excuse me, that specifies how entries in the team database are compared to entries in the other databases, the -- whether it -- called the identification databases.
Q So what -- so how does that work? Give me a part of the algorithm. What are you looking to do from one database to the other?
A Well, for example, one might observe a record in the team database of a registered voter whose named John Doe. And this record in the team database has lots of other fields. It not only says John Doe's name, it has, you know, his address, mailing address, residential address, zip code, and so forth. And the question is, how do we know if that particular John Doe also appears in say the driver's license database where there's another individual whose name is John Doe, might be spelled slightly differently for example, so the algorithm is an attempt to figure out how the different individuals map together.
Q Okay. So if you had -- if every name in the team database were entered in the same way, the exact same way in let's say the passport database, it'd be pretty easy, wouldn't it?
A Well, we would -- just the fact that the names matched

wouldn't be sufficient because there are lots of people who have repeated names.

Q Okay. So does the algorithm include several different ways of looking for things that might be matches?

A Yes. It includes a number of different ways. Each of these ways is called a "sweep." And a particular sweep takes a record in the team database, as I think Dr. Ansolabehere explained a bit yesterday, takes a record, and from that record forms some sort of identifier using fields in the team database. So, for example, one wouldn't only use a name for the fact the names are repeated. But if one were to do that, one identifier might simply be the name John Doe with spaces removed, for example, and then one would take that record and look in the driver's license database to see if the same identifier appeared. Now, of course, one does not only use names, but that's the idea.

Q So if you have John Doe in the team database and there's somebody named John C. Doe in the DPS database, what do you do to figure out if that's a match or not?

A Well, this is -- this comes to the question of what are called "sweeps." So each sweep in the algorithm that the Department of Justice executed puts together a set of fields. And because the names are repeated, to use that example, there are no sweeps that only function on, say, first name/last name, because that would lead to lots of incorrect matches. So to

know if John Doe in the team database corresponds to, say, John C. Doe, one would also use other identifying information such as addresses or perhaps a Social Security number or a driver's license number if we were talking about a driver's license database and so forth.

Q Could you have a sweep, for example, that said let's pick up all the ones in which the last name, the house number, and the last four digits of the Social Security number are the same?

A Sure. In principle, that's fine.

Q Okay. And so your sweeps are different variations of the data to try to figure out which things are -- which records are for the same people; is that correct?

A Each sweep is a different way of comparing -- a slightly different way of comparing records in the team database to records in an identification database, that's correct.

Q And if you have more than one sweep, what's the purpose?

A Well, the concern in any database matching process is that typographical errors and variations in name spelling and so forth make it difficult to know if one John Doe is the same as another John Doe who might be -- have a middle initial. So in order to get around this problem a database matching algorithm will have multiple sweeps. Some sweeps, for example, might simply ignore middle initials so that if one database doesn't contain middle initials for whatever reason, that doesn't

contaminate the process. So in this particular algorithm that I analyzed, there are 13 different sweeps.

Q Thirteen different sweeps. If I told you that the D. C. case, or what we'll call the Section 5 case, in here had three sweeps and this case had 13 sweeps -- the algorithm in this case had 13 sweeps, what would you say about the likelihood of accuracy in this case compared to that one?

A Well, any database matching process like this -- well, let me -- excuse me. I wasn't involved in the D. C. case.

Q Right.

A I've read a little bit about it.

Q Take my word for it, it was three. I think that's right, isn't it? Okay.

A I was --

MR. SCOTT: I wasn't involved either.

MR. DERFNER: We may be the only two people here.

BY MR. DERFNER:

Q Go ahead, doctor.

A Thank you. In this instance, I think 13 is greater than three. And so that is -- because the sweeps were not as simple as what I'm describing, first name/last name, that's a good thing. So I'm pleased that the number in this -- with the algorithm that I'm studying is greater than three.

Q Okay. And so the algorithm -- part of the algorithm, if I understand it, is the actual sweeps are in there, right?

A Yes. That is specified in the algorithm.
Q Okay. And the algorithm, I assume, has other things in it that tell the computer what to do about this or that?
A Well, as I -- so an algorithm is a set of rules. And one -- in this case, whoever is executing the sweep would write computer code in order to carry out the sweeps efficiently. You don't need a computer. In principle, you could do it by looking, but that would take forever.
Q Okay. And, in fact, weren't there two separate algorithms in this case?
A There were.
Q What were the two?
A Well, one algorithm was called the Plaintiff algorithm, and that was a set of sweeps that I've been discussing. And there's another algorithm that the Defendants used, and that's called the Defendant algorithm. And that was a set of I believe four sweeps, sometimes five, because that number depends on the availability of driver's license numbers.
Q Now when you say, "the Defendant used," weren't both algorithms used and produced for everybody?
A Yes.
Q So rather than used, wouldn't it be fair to say the one algorithm was selected by the Plaintiffs and one was selected by the Defendants, correct?
A Yes.

Q Okay. But both were used for everybody, right?
A Both algorithms were applied to the team database and the complete set of identification databases.
Q And you, for example, got the results of the running of both algorithms, right?
A Yes.
Q Okay. These -- you were running through how many team records in this case?
A Approximately 13.4 million.
Q Okay. Will the algorithm give you precise answers for whether every one of those 13.4 million is a match or not a match?
A The output of the -- of this algorithm in this case was a set of indicators that correspond to team records. So, for example, to return to my John Doe example I gave before, the results of the algorithm would say whether John Doe -- as in the team database, would explain how this individual matched based on every single sweep to all the identification databases. So the data that I received contained flags. I would call those precise. They would say one or zero. And they indicated whether a particular team record matched in a particular sweep, one to 13, to a particular identification database.
Q And tell me again, the identification databases, that's DPS, State Department, Homeland Security, that correspond to

the different types of IDs permitted under SB14?
A Yes. The identification databases were chosen based on what forms of identification are allowed -- at least, that --
Q So are you saying that the results -- well, do -- did the results that you got back, was it a piece of paper that said there are "X" number of matches?
A No.
Q Okay. Did you get back in fact millions and millions of lines of code?
A I received -- I didn't count the exact number. Millions of lines of data, I would say --
Q Sorry.
A -- and the data files reported some information from the team database -- again, the record of Texas registered voters -- along with various files that were the output of the Plaintiff algorithm and also of the Defendant algorithm. So these files -- each file had approximately 13.4 million lines, although there was one exception because of the way that some of the files are handled, and what I received was all those files. So I suppose if you multiply that number by eight or nine, that's how many lines I got.
Q So let me see if I understand. It sounds like -- let's say from the State Department, you got back a -- from the State -- well, you got back records of the State Department showing for each of the 13.4 million names in team, you got in effect

13 results for each of the people in team, with in effect a yes or a no?
A Yes. It was a little bit more subtle than that --
Q I hope.
A -- but that's the right idea, because when the different federal agencies, as explained by Dr. Ansolabehere yesterday, implemented the DOJ algorithm, there was the occasional -- trying to think of the right word -- the occasional hiccup and -- when one agency basically I think ran the algorithm again, so I think I had duplicate data from that agency, strictly speaking, so I had more than 13, but I knew how to combine them to make them effectively 13.
Q Okay. And to count something as a match from -- let's say, a given agency had 13 sweeps for each name in the team database, to count it as a match, how many sweeps did there have to be a yes on?
A So I chose the most conservative number, which is one. So in this instance, the Plaintiff algorithm is actually divided into a set of what are called primary and secondary sweeps. And I haven't really discussed that because it's not particularly --
Q Don't bother.
A Okay. And so if you have the set of sweeps -- we'll just go with that -- as long as one -- as long as an individual in the team database matches on any sweep in any particular

identification database, for example a passport database, then I consider it a match. So the number is one.

Q Okay. Now, in order to analyze these millions of pieces of data that you got back, what did you have to do?

A Well, originally in June when I received the data, it was encrypted, so I used the password supplied by counsel to decrypt it. It came in a hard drive sent to me -- not a hard drive but effectively a storage device -- then I looked at the data to sort of see what it was, and then I started writing computer code that took all the data and uploaded it into a database that I was creating.

Q Okay. And when you say you started writing computer code, how much of that -- how detailed or simple is that?

A I'm used to writing this sort of computer code. I do it regularly. I --

Q How much computer code did you have to write?

A I would say I wrote hundreds of lines. I haven't counted the number -- hundreds of lines.

Q Each line is in a sense an instruction or part of an instruction?

A Yes. So, for example, when I took the passport -- I didn't have the actual passport database, so when I say the -- when I took the passport database, what I mean, I took the output of applying the algorithm to the passport database. When I had that data, I wrote some code to take all those ones

and zeroes and other numbers -- and these databases also had what's called the Texas VUID. Not all the files had this, but they effectively did, I'll go with that. And I wrote my code to sort of take all these pieces of information and -- the term is upload those files into a database.

Q Okay. Now -- and so when you finished writing your code and running the data through that, that's when you got some numbers of how many people were matches, non-matches, etcetera; is that correct?

A Yes. Initially, of course, all I had was all these lines of data and I couldn't tell from those who matched anywhere. It's not enough. So I took the -- I wrote code to upload all the files into my database and then I wrote additional code to manipulate all those results. I had -- one step is to take the different database results that I had and basically link them together, or join them -- that's the technical term -- using what's called the Texas VUID. And that step I had to do first before I could analyze any matched numbers.

Q Okay. Did you -- in writing your code to derive the results from data, did you rely on anybody else, such as Dr. Ansolabehere?

A No. All of the code I used in this process I wrote myself.

Q And so anybody who's analyzing would expect to write his own or her own code?

A I mean, I can't speak to anyone, but I wrote everything myself.
Q Okay. And if the Defendants had anybody analyzing these data, then that person would presumably also be writing their own code, right?
A Unless the -- unless this individual had mine, for example, or someone else's. But, yes, in principle, you can't analyze these data in the form that I receive them without writing computer code to do this other data assembly process that I've just described.
Q Okay. And in terms of -- so that gave you at different stages as we'll find out, it gave you the -- as a starting point, the number of non-matches; is that correct? Or number of matches and non-matches?
A Yes. After the -- I mean, I've sort of summarized a lot of computer code writing here, but yes, after one writes code to upload the data and put the datasets together and then start looking for ones and zeroes in the appropriate places, that gives me numbers of matches and non-matches, that's correct.
Q Okay. Now, is this a perfect way -- is the answer going -- the answer that you got going to be completely perfect in telling whether every single person is a match or a non-match?
A No. I don't think it would be completely perfect.
Q And why is that?
A Because this database matching process, like all processes

in which I have been involved, is subject to some sort of error. And the databases I was working with via the Plaintiff algorithm and the Defendant algorithm are very large, 13.4 million. So it wouldn't be surprising to me that if you were to tell me, and as I saw in court yesterday, that there are some issues of errors that come up when working with a database of that size.

Q And if there are errors like that, does that mean that somebody made a mistake?

A Obviously that depends on the type of error. But the algorithm is a set of rules designed to check whether, you know, names in the team database correspond to names in other databases. Those rules aren't a hundred percent perfect. We know that humans make mistakes. There might be data entry errors in unusual places. These things crop up now and then when databases are extremely large -- 13.4 million records for example -- so it doesn't imply that there's an error the way you're describing.

Q Okay. And does it imply that there's any error in the person doing the analysis? In other words if it turns out that some thousands or whatever of people are not correctly listed by this, does that mean that you or some other analyst calculated wrong?

A No.

Q Now, in fact, do you have a --

MR. DERFNER: Your Honor, I think the witness has a handheld calculator. May I ask him to use it for a minute?

THE COURT: Yes.

BY MR. DERFNER:

Q Do you have a calculator, Professor Herron?

A Yes, sir, I do. It's right here.

Q Okay. Well, you heard some testimony yesterday I think about some -- I think it was 22,000 people in some category -- I'm not going to go into detail -- but if there were 22,000 people in some category who may have been mislabeled -- I'll call it mislabeled -- by this process, in other words, called no matches when they were really matches, or vice versa, how would that compare to the total universe that was analyzed?

A Well, if you're asking me to assess the question, is 22,000 large? And the answer to that question depends on sort of what we're comparing it to. So if you were asking me right now, I would, as a first thought, I would just divide 22,000 by I'll say 13.4 million and see what that fraction shows.

Q I'm -- yes, I am asking that.

A You're asking me that, okay. So I'm just going to type in 22,000 by 13.4 million, and I'm going to multiply the result by a hundred, so I get around approximately .16 percent.

Q Okay. So that's between one-tenth and two-tenths of one percent?

A Yes, that is correct.

Q Okay, thank you. Okay, I think we've come about to the end of the -- well, let me ask it this way. So is this the process -- you said there were two questions you were asked. Number one, how many matches and non-matches. And the second question had to do with assigning -- analyzing the racial composition of the list, right? So is the process you've described so far, is that what you did to figure out the number of matches and no matches?

A Yes.

Q So let's then move on to the next step. Well, and by the way -- no, I'll come back to that. You were asked to analyze the match and no match list according to race; is that correct?

A I was asked to make an assessment of the extent to which different racial groups -- three -- whites, blacks, and Hispanics had different match rates or correspondingly non-match rates.

Q Okay. And in one sentence as a summary, how do you do that?

A In one sentence. I -- what I did is I took data on race and ethnicity from the American Community Survey and I combined those data with the matching output data that I've described, and I used that combination to draw some conclusions about racial and ethnic voter identification possession rates.

Q Okay. You used the word "American Community Survey." And were you looking at something called CVAP?

A Yes. The American Community Survey --
Q We'll get to that. I just want a yes or no right now, excuse me.
A Yes.
Q And what were the methods, the statistical methods that you used in doing this analysis? Just give me the names because then we're going to go through these words and names one by one.
A I carried out two standard approaches, standard methods, to my dataset when it combined the American Community Survey data and the matching algorithm data. Those methods, which I've used in my own research and in which plenty of other scholars use regularly, are called homogeneous area analysis and ecological regression.
Q Okay. And how about one more term. In the area analysis, were you looking at census blocks?
A No. I looked at --
Q I'm sorry, go ahead.
A No. I looked at what are called census block groups.
Q Okay. So we'll take a look at that term, too. Okay. Now, tell me what -- you were trying to figure out the racial composition of the no match and the match list, correct?
A Yes.
Q Okay. And were you trying to figure out the CVAP figures for those numbers?

A Well, CVAP stands for citizen voting age population.
Q Okay.
A I wouldn't say I was trying to figure out the --
Q Well, the CVAP, the basis of -- or the type of data that you wanted to figure out, the racial composition.
A Yes. I wanted to -- when I was interested in racial composition, I was only interested in using census data that included information on citizens, because in the State of Texas, you have to be a citizen to vote, and you also have to be 18. So CVAP includes data on eligible voters.
Q And what does CVAP stand for? What is each word in that acronym?
A Citizen voting age population.
Q Okay. So that's what you want to know. And where do you go for that CVAP data?
A Well, the standard place to go in the academic literature is the American Community Survey, so --
Q What is the American Community Survey? Is that a census -- is that census data from the Bureau of the Census?
A Yes, it is. The Bureau of the Census has many products, one of which is the American Community Survey. And I'll just refer to it as the ACS. Another is the decennial census, which many people simply refer to as "the census," but that's slightly misleading because there are lots of censuses.
Q Why didn't you use the decennial census in this research

project you did (indiscernible)
A No, because I need CVAP. I need information on citizens and that's contained in the American Community Survey.
Q So what you would call CVAP, the citizens voter age, you couldn't find that in the decennial census, could you?
A No, correct.
Q Okay. And the ACS, is that -- does that come from a particular time or date? Like the decennial census we know is done as of the decennial year and completed or published I think on April 1st of the year ending in zero or one. What about the ACS?
A The ACS is a -- the one I used, the most recent one, is called the 2008-2012 ACS. It doesn't come out every ten years. Well, it doesn't only come out every ten, is I should say. It's more of a rolling snapshot of the country. That's a simple way to think about it.
Q And so for a researcher, whether in the commercial world or the university world, who wants to know about citizens and voting age, etcetera, is ACS the standard way to find that data?
A Yeah. I would say that in my area, that a standard source is the American Community Survey.
Q Okay. Now, you -- so you're going to use that data to look at census information, correct?
A Yes.

Q Okay. And you mentioned the word "block" -- well, you corrected me when I said "block." Tell me what are the -- what are some of the basic kinds of census units and what do they mean?

A Well, the census defines various geographies that cover the United States. They range from small to big, and the smallest unit is basically a block. So it's not a city block, as some people think, but it's called the census block, and that refers to a particular geographical area. And then there is the next unit that's larger, it's called the block group which contains blocks. I use block groups in my analysis because that was the lowest -- or the smallest sort of units in the American Community Survey that contain CVAP data. So some of the decennial census data is published at the block level. The ACS doesn't go that low. And when I say "low," that just means smaller units. So that's why I used block groups. Another unit that's larger is a census tract. And obviously if one has block group data, you can aggregate up to a tract. But it's better to use the data at the smaller level if you can get it.

Q And partly because when you use a small level, you get more units in the sample, don't you?

A That's one consequence, yes.

Q Okay. And did they used to have something called enumeration districts?

A I -- the geographies used in the census have evolved, but I really don't know.
Q Okay. So -- okay, so now I think we know we have the building blocks. What did you do? You said you had two methods. Let's start with the first one. I think the first one you mentioned was the homogeneous analysis, or something like that, right?
A Yes, that's correct.
Q Tell us what that is.
A So I'll use the word "block group" now that --
Q By the way, at any point along this --

MR. DERFNER: Your Honor, if I may, at some points we believe that it may be useful for Dr. Herron to actually come down and do some drawing or make some -- give us some illustrations to help us see. I'm not sure if that's true now, but it probably will be in the next one.

THE COURT: That's fine.

MR. DERFNER: Okay.

BY MR. DERFNER:

Q So Professor Herron, tell us what that is.
A So in a homogeneous block group analysis, what I do is I look at the block groups that cover or partition Texas. And I know for each block group, from the ACS data, which block groups are, for example, a hundred percent white CVAP. So when I say a hundred percent white, I mean a hundred percent white

CVAP. And I'll -- I won't constantly say CVAP, but that's implied. I can tell which block groups are a hundred percent white and which block groups are a hundred percent black and which block groups are a hundred percent Hispanic. That's what I mean by homogeneous.

Q And it might be -- these groups might be anywhere in the state, right?

A Sure. Would you like me to draw this?

Q Yeah, probably that would be a good idea.

THE WITNESS: Your Honor?

THE COURT: Sure.

THE WITNESS: Thank you.

(Pause)

BY MR. DERFNER:

Q Is that Texas? Where's the panhandle?

A I'm not an expert on drawing. So if, for example, this is --

THE COURT: Is that Texas?

THE WITNESS: This is Texas.

THE COURT: Okay.

(Laughter)

BY MR. DERFNER:

A If this is a state, I might know, for example --

Q Let's call it Texas.

A I'll call it Texas. I'll just draw some circles in here.

I'm not suggesting that block groups are circular. That's not true in general. I might be able to say something like, well, in Texas I can find in this drawing four block groups that are a hundred percent white. And so they might not -- they might be scattered around as I've drawn, they might be some of them that are closer, for example. But with the census data, I can tell which block groups are of that type. And --

Q Every block -- excuse me -- every block group has a number or an identifier, right? In the census.

A Yes.

Q Okay. Go ahead.

A So to continue your earlier question, I figure out which block groups are white, a hundred percent white, and then I combine them effectively and think about a subset of Texas that is completely or a hundred percent white. And I can do the same exercise for a subset of Texas that is a hundred percent Hispanic. So I might also have some block groups -- and I'll just -- I'll put a little line through here just to make them different. I'll just draw three up there. I might say, well, we have some that are a hundred percent white and some that are a hundred percent Hispanic, and I could similarly do that for ones that are a hundred percent black. I won't draw them up. But in principle, I could do that. And that gives me an area -- or in this case two areas -- in Texas that are homogenous of one racial ethnic group and another area that's homogenous in a

different racial and ethnic group.
Q And so what do you do with that data?
A Well, the nice thing about these areas is that I can -- since -- if I have an area that's a hundred percent white and I observe something occurring in that area, I know that whites did it, and that's because, according to the census -- and that's my data source here -- the only people who live there are white. And when I was asked to make a -- to do a racial ethnic analysis of the match rates, I noticed that in my data, I don't have individual level data on race. So it's not the case, for example, that when I put the data together from the Plaintiffs algorithm that I could say, oh, this individual in the team dataset -- in the team database, this individual is white and so forth. But I didn't have that sort of individual level data.
Q And -- let me just break in. That -- is that individual level data, that's not the function of a statistical analysis (indiscernible) isn't it?
A Well, if I have it, but not in this -- my case, no. So what I do here is I -- is this -- I don't know up front which individuals in my dataset are white, for example. I can look at these areas by census blocks here and say, well, anyone who lives in one of those areas must be white. Why is that true? Because the census says only white people live there. Similarly I can look at these other areas, the attached three -

- I think I said they were Hispanic -- and if you lived in one of those, then you're Hispanic. And that must be true because, according to the census, the only people who live there are Hispanic.

Q So you take that information -- you may have a number of census block groups that are all white, and would it be fair to say you add those together, you see how many -- what percentage of match and non-match theories, and that's one number that you have for the percentage match and non-match among white people in Texas; is that correct? That a fair way?

A Yeah, that's a fair way. So if I look at the five white groups -- five white block groups, excuse me, and I combine them, even though they're not, say, contiguous, that gives me a white big block group, say; and then if I observe that in that block group, the match rate for ID possession is, say, around 95 percent, then it must be true that in that area that the whites -- the white identification rate is 95 percent.

Q And you do the same thing for blacks, that is, the black -- the homogenous, hundred percent black block groups, correct?

A Yes, that's correct.

Q And you do the same thing for this Hispanic?

A Yes, that's correct.

Q And so -- and is that what you call the homogenous block group analysis?

A Yes.

Q Okay. And so that's one way of figuring out what the percentage rate among whites and among blacks and among Hispanics in Texas there is for voter ID, right?
A That's one method that I use and that is commonly used in the literature. It's a -- the homogenous area analysis is the standard in the literature.
Q Okay. But it's not the only method, correct?
A That is correct.
Q Okay. Is there something that's close to it?
A Yes, there is.
Q And what is that?
A You might call --
Q One quick question. And the answer to -- when you get results on the homogenous block group analysis, those answers are reflected in a single percentage number for each race; is that correct?
A That's correct, because the block groups are homogenous.
Q Okay. What's the next method?
A Well, I could take -- rather than look at completely homogenous block groups, I could look at block groups that were almost homogenous. And that -- the method used to study those is called the method of bounce.
(Mr. Derfner/Ms. Simson confer)
MR. DERFNER: Your Honor, we have a couple of demonstrative exhibits that I think may help; although, I think

Dr. Herron is doing just fine on his own. We've drawn some pictures that may help him in explaining this next one which is -- what do you call this one?

THE WITNESS: Well, it's what the field calls the method of bounce.

MR. DERFNER: Method of bounce. Is -- would you call -- is it another -- would a colloquial person like me maybe call it the nearly homogenous block group analysis?

THE WITNESS: Yes.

MR. DERFNER: Okay. May we put the demonstratives up?

BY MR. DERFNER:

Q Okay, so tell us about this method, and feel free to use the picture.

A Okay. So the previous example where I drew the State of Texas involved block groups that were completely homogenous. But let's suppose we wanted to relax the definition of homogeneity and look at block groups that were almost homogenous.

Q And why would you do that?

A Because the number of block groups that are completely homogenous is not very large. It's --

Q And you want a bigger sample.

A Yes.

Q Okay.

A So I would relax the definition -- or this method relaxes the definition of homogeneity and says, well, for example, let's -- rather than looking at block groups that are a hundred percent homogenous, let's look at block groups that are at least 99 percent homogenous. So we might say 99 percent is very close to a hundred, but obviously it's a little bit less. And so you then might take 99 and say also 98 and so forth. And in this picture, I'm looking at a block group that is 90 percent black and -- well, 90 percent black and ten percent white. Obviously I'm using 90 because it's easier to have ten people up here. And you notice in the coloring scheme here that the races and ethnicities of these individuals are different. And what's important to recognize is that if I -- when I look at this in the census data, I can't tell you that this individual person is different. All I know in this block group is that, yes, there are nine blacks and one white. Doesn't matter that I can't tell that if I were there which person is which. That doesn't matter. So I'll continue. Suppose I know in this particular block group -- suppose on this particular block group -- again, there are ten people, nine black, one white -- that two people lack voter ID. So notice I'm combining information from the database matching algorithm with this American Community Survey data. If the algorithm says that two people lack ID and I'm looking at this block group, I notice the following. Since two people don't

have ID, and there's only one white person, I can say something about the black ID possession rate in this block group. So you might say, you know, what is the lowest that the black ID possession rate can be? And if you think about it, the lowest that the black ID possession rate will -- can be will happen when this white person doesn't have an ID, because it must be true that if two people lack ID, that at least one of those persons is black. The reason that at least one has to be black is there aren't enough non-blacks to make up two. So I can write down one-ninth, and I know that at least one-ninth of the blacks here lack ID. That must be the case because there aren't enough non-blacks to make up the number two. Similarly it's possible that both black -- excuse me, that both of the people who lack ID, they could both be black. So I could say one-ninth and two-ninths, and I know that in this illustrative example that the black ID -- excuse me, the black non-possession rate must be between one-ninth and two-ninths. It has to be because there aren't enough non-blacks. So in this case, this person is white. But had that person been Hispanic, the same example would have applied. All we need is that we have a lot of blacks and a small number of non-blacks.

Q So, Dr. Herron, when you draw conclusions from this type of exercise, is your answer reflected in a number the way it was with the completely homogenous block group analysis, or something different?

A No. As this indicates here, it's indicated in a range, because I can't tell you here what the black identification lacking -- the black non-possession rate is. I can only put it in a range. The exercise in statistics is called binding. So I can only -- these are called -- and these are called logical bounds. One says the logical bounds are between one-ninth and two-ninths. And the only thing I can do here is bound, or place in a range, the non-possession rate.

Q Okay. And have you done that in this case, Dr. Herron?

A Yes.

Q And does your report in fact include a showing of this logical bound?

A Yes.

MR. DERFNER: Your Honor, may we post page 53 -- this is Exhibit 769, which is Dr. Herron's amended reported. Page 53 is what we'd like to post.

THE WITNESS: May I sit down?

MR. DERFNER: Yes.

BY MR. DERFNER:

Q And -- so, Dr. Herron, did you use this method -- this -- I'll call it the logical bound method, or the nearly -- call it my term, nearly homogenous block group analysis, did you use this as a second method to analyze the racial composition of the match and no-match list?

A Yes. I wouldn't call it a second method since it's very

related to the homogenous block groups. But, yes, I used the method of bounds, or what you're calling the nearly homogenous method, to analyze the racial composition of the individuals who I identified as non-matches from the Plaintiff algorithm.

Q Now, you gave us an example of a 99 percent block group, that is a block group that is -- or block groups that are 99 percent of one of the three races. Did you go down below 99?

A Well, yes. The example I just drew I believe had 90 just because it would be too difficult to draw out that many people. But one can relax the definition of homogeneity as low as one wants. So one can start with a hundred and go to 99 and 98 and so forth and one can just -- one can go lower. In fact, in this figure here, the definition is relaxed down to 90 percent.

Q And are there pluses and minuses as you go down the -- from a hundred?

A Yes.

Q And what are those?

A The lower the cutoff goes -- in other words, the more relaxed one -- the definition for homogeneity becomes, the less informative the analysis becomes. And that's because when a block group is very heterogeneous, in other words, it's very mixed, then it's difficult to know -- it's difficult to say much about the racial behavior in that block group. So in the example we had on the board a few minutes ago, I had the case where there was one non-black and two missing IDs. So we could

say something about the black identification rate. But suppose in that example we had something like two missing IDs and five blacks and five whites. Okay, in that -- so that example block group would have been I guess completely heterogeneous, 50 percent white, 50 percent black. If I had that, then I couldn't say really anything about the possession rates of ID based on this sort of analysis, and that's because if the block group were say five whites and five blocks, it's possible that the black ID possession rate is zero and it's possible that the white ID possession rate is zero. You can't -- I can't tell. This method is informative -- and that's the statistical word here -- typically only when the cutoff is closer to a hundred.

Q So on one hand you want to come down from a hundred to get more block groups in the sample to be more meaningful; is that correct?

A Yes.

Q On the other hand, as you go down further, it becomes less informative, or less precise?

A Yes. That's typical in statistics. As we change one feature of a method of analysis, there are tradeoffs. One thing gets better, one thing gets worse.

Q Okay. So that's what we'll call -- we'll call that method 1A, not quite a second, different method, okay. And you used this method here, correct?

A Yes, it's in my report.

Q Okay. And what's the other method that you used?
A I used ecological regression to study the racial and ethnic composition of the non-matches and the matches in my analysis.
Q Okay. I've been working with ecological regression, or people who do ecological regression -- and I know there are a bunch of them here in the audience -- for 40 years, and I still don't know what it is. I don't expect to learn today, but why don't you tell us a little bit about what that is.
A Yes. Okay, I can do that.
Q And feel free to draw if you want to.

THE WITNESS: Okay, your Honor, may I draw?

THE COURT: Yes.

THE WITNESS: Thank you.

MR. DERFNER: Maybe I will learn.

BY MR. DERFNER:

Q By the way, is ecological regression a standard statistical method?
A In the social sciences?
Q Yes.
A Yes.
Q And, in fact, is it something that the courts, including the Supreme Court, have used again and again?
A Yes.
Q And, in fact -- well, you may not know this, but you may -

- are you aware that the Supreme Court has reversed cases where statistical methods used were not ecological regression?

A I haven't read those cases but I have been told that's the case, yes.

Q Okay. Tell us something about it (indiscernible)

A So suppose I have a plot and on the "X" axis here, I'm going to label it percent white. And I'm going to put some dots on this plot. In this diagram -- I'll just draw five dots -- make that six -- each dot refers to a block group. So the location of these dots describe the percentage white of them. And we know percent white ranges from, say, zero up to a hundred. So --

Q When you say the "X" axis, you mean the horizontal line?

A I mean this axis right here (witness indicates).

Q Okay.

A Okay? So these dots represent the percentage white of a particular block group. And on this "Y" axis -- this is the "Y" axis here (witness indicates) -- I will put -- I'll just write down percent ID, but that's shorthand for percent of individuals who have valid voter identification per Senate Bill 14. So I could make a plot like this in principle with my data. And here you have percent whites and here you have percent ID, and these block groups -- these six block groups that I've drawn, I haven't selected because of homogeneity or nearly homogeneity or something like that. That's not relevant

in this particular example. And then what I could do was I could ask, using a technique called regression analysis, which is a standard technique in the social sciences -- well beyond social sciences actually -- I could say, well, could I fit a line -- and I'll just draw --

Q Let me just stop you there. You said you're not looking for homogenous block groups on this. Are you putting in more or less all of the block groups in the State?

A As many as I can.

Q Thousands?

A Yes, and I mean, I've gone six years, in my report I use, I think, around 11,000.

Q 11,000, and so for this purpose it doesn't matter what the racial composition of the block group is?

A That's right, I don't select based on homogeneity.

Q Okay.

A Or nearly homogeneity.

Q Okay, so keep on going.

A So I might use a statistical technique called "regression analysis" placed a line here what's called fit a line based on the locations at these points. This -- this line would be called an "ecological regression line" as opposed to just a "regression line," because on the X axis and the Y axis here we have percentages in groups. So this is just a regular regression. So the term "ecologic regression" just means a

regression applied to ecological units or what are called "aggregate units," so that's all we have here, a line that I superimposed here, and there is a formula --

Q (indiscernible), terms like called "bivariate analysis" and multi-variate analysis," do those mean anything in this?

A Yes. When I describe it here as a "bivariate ecological regression," I didn't carry this exact thing out in my research for this report, but this is intuitively what's going on, and one can understand ecologic regression in this case even if you have what's called "multi-variate ecological regression."

Q I'm not even going to ask you what those mean. Okay, keep on going.

A So you have this line here, and this line, and there's a mathematical formula that determines that line I -- so one uses computer software to calculate it. It's a weighted ecologic regression that should be the standard way to -- to study these sorts of points and this line summarizes the data. And what I mean by that is you can imagine this line here, it could be further up like this, that's called north (indiscernible) slope or it could be, say, down, negatively sloped, or it could be flat, something like this, and the location of this line and how it's sloped and where it hits this axis right here, that allows me to calculate, to use this line to calculate white, black and Hispanic, in this case just whites, ID possession rates using a large collection of block groups.

Q So are you saying that this line somehow translates into a percentage? In other words, when you cross this line, there is some way that mathematically or algebraic, whatever, calculus, you -- that line turns into a number that says "13.6 or 2.9," something like that?
A Yes, it depends on -- as I was describing here, it depends on whether the line is sloped up or sloped down, and where it hits this axis, and there's standard formulas that say how to do that.
Q That's my question. In other words, turning this line, which is based on where all of these data points are, into the number is a standard method, is that correct?
A Yes.
Q And if two scientists had the exact same data points and put in the formula, they should, in that situation, come up with the exact same answers, shouldn't they?
A If everyone had the exact same data, and was interested in the exact same line, then two people, using the same software, would generate the same answer.
Q Because the formula is the formula, correct?
A Yes.
Q Okay. Now, let me ask you this question:
You put this line through, but sometimes these points will be closer to the line and sometimes the points will be further away from the line. Is there a way that the formula

takes that into account?
A Yes. So I drew this so that the points are somewhat scattered around the line. It could be that some points are right on the line; it could be that some points are not, and that affects the ability of one to draw conclusions from the lines, so the formula is about the precision that comes out of these line estimates, both that, the extent to which these points are clustered about this line or set up, that's part of the formulas that one uses when analyzing regressions.
Q Now, I'm going to ask you about a term I've heard, and I don't need you to explain it much, but what's an "R square?"
A An R square value is a measure of how well these points, the (indiscernible) I've drawn, are clustered about this regression line. So in this case it's a squared correlation, but the way to think about this is that the points could be really tightly clustered, or there could be points that are staying way up here -- oh, excuse me, way up here and way down here, they could be scattered all over the place. The more scattered they are the less they are clustered about the line and they're -- it turns out the lower the R squared value.
Q So -- and you used this method, this ecological regression, to also estimate the percentage black, Hispanic and white for non -- for possession and nonpossession rate of SB 14 photo IDs?
A Yes.

Q Okay. Is this a third method you used, or it's a second and a half?
A Yes.
Q Okay. Okay, I think -- here's the form, thank you. Thank you, Mr. (indiscernible).
Now tell me, what's the advantage -- let's go back to homogeneous block group analysis, what's the advantage of that?
A With homogeneous blocks -- excuse me, block groups, one can look at a collection of block groups and know what happened in them because, according to the census data -- when I say "know what happened to them," I mean know what the different racial rates of behaviors are, in this case ID possession, because the block groups are completely homogeneous, so we don't have to talk about uncertainty in a -- if I have a block group that's 100 percent Hispanic, for example, and the ID possession rate in that block group is 90 percent, I don’t have to say, "Well, maybe it's the non-Hispanics who are -- don't have IDs" because there aren't any.
Q Okay, and what's the disadvantage of these -- of the homogeneous or logical bounds in that -- by the way, I take it the logical bounds are simply something that are not quite as precise (indiscernible), for example?
A I wouldn't say it's not quite as precise although --
Q Or wasn't precise, okay --
A So -- maybe you could say that. If there are -- there's

more -- yeah, that's right, there are more uncertainty, there's more uncertainty when you relax the definition of homogeneity, the bounds get wider, that's what that means, so the advantage is that what these bounds are what are called "accounting identities," and that means when you have a set of logical bounds it must be true that the behavior you are studying, in my case, ID possession rate, falls within those bounds.

Q So the advantage of a homogeneous block group analysis or it's relative to the (indiscernible) bounds analysis if it were pretty confident or very confident that these people had that percentage, correct?

A Well, I don't use the word "confident" here, but I would say is that these -- is that the bounds, the logical bounds that I have described, must capture the quantity of interest. So, yes, I am confident, but it's stronger than that.

Q Okay. What's the disadvantage?

A The --

Q -- the limitation of this data?

A The disadvantage is that the homogeneous and the nearly homogeneous block group analysis relies on a smaller subset, the number of block groups relies on a small subset, and that subset simply depends on the geography of, in my case, Texas. I can't -- I can't move people around in Texas and say let's have blocks of homogeneous white ones and blocks of homogeneous black ones, they are what they are, and so, you know, the

advantage is precision; the disadvantage is a smaller number of block groups.

Q Okay. And is it also a disadvantage that, in fact, the behavior or facts related to people in what I'll call homogeneous (indiscernible) segregated favorably might be different from the behavior of people in, I call it more multi-racial areas?

A Yes, that's possible also.

Q Okay. And, finally, it's a little bit technical, but how many block groups in Texas -- you said there was thousands (indiscernible).

A There are approximately 15,000 overall. Some of them, as Dr. Ansolahehere here noted, after you don't have any within them, (indiscernible) highway, so around, I believe, 15,000.

Q And about how many did you use for your logical -- or how many -- not did you use, how many turned up in your logical bounds analysis?

A Well, that number depends on how low I want to make the cutoff. I believe if I go down to 96 percent the number is around 800.

Q 800, just a block of --

A I believe so, approximately.

Q And as I recall, (indiscernible) I think this is in your report, there is only several hundred thousand people?

A I don't remember the number off the top of my head.

Q Okay. Now, so that's an advantage and a little bit of limitation, disadvantage of this homogeneous and logical bounds type of analysis.

What about the ecological regression, what's the advantage there?

A So the advantage is that one can use a much larger group of block -- a much larger collection of block groups. I don't have to specifically only use places that are very homogeneous, and that's nice because the sample size then gets larger, and also I don't have to be concerned that maybe homogeneous block groups are different in other ways, so that's the advantage of that approach.

Q Okay, and what's the disadvantage?

A The disadvantage is that it's -- it's a statistical model and that -- I say it's a statistical model as opposed to a -- a calculation that bounds quantities which means there's more uncertainty, in general, that comes out of a regression model than that comes out of a bounds analysis, so it's just the typical tradeoff in statistics. You make your sample size bigger, which is good, but there are costs as well, and that's -- that tradeoff is, again, very common in statistical analysis.

Q So what happens when you use both methods (indiscernible)?

A Well, what I look for in my own academic research, and what I looked for here is when I take these different methods,

two and a half, do they give similar or different results? That's the question I'm looking for.

Q Okay, and did they?

A No.

Q I'm sorry. Similar results?

A Oh, excuse me. They gave very similar results.

Q That's what -- okay. Okay. Okay. I think we now know maybe more than we ever wanted to know, but as much as we have to know for this case about the process.

Let's go to the results, let's go to what actually happened. How did this begin for you? Did you get some data from the database matching process?

A Yes, as I described earlier, I received that encrypted hard drive and wrote a lot of code to take the data and put it altogether.

Q And you got that around the 1st of June of this year?

A I believe I got it either May 30th or the 1st of June, yes.

Q Okay. And you then did some analysis and you did a report, correct? And that was handed in on June 27?

A I don't remember the exact date, but I wrote a report applying the methods that I have described here to the data that I received on approximately June 1.

Q And then it turned out that there was a problem with the data, correct?

A I received an affidavit to the effect that when the Texas DPS sent their data to the Department of Justice, that due to some computer code error, I'm not sure what it meant, that approximately 2.8 million records were not included and so that was -- yes, that was the error that I was alerted to sometime, I believe, in June or July.

Q Okay. And then, in fact, it turned out that even the 2.8 million was not correct, is that right?

A I believe -- I was told by Counsel, and I believe I heard yesterday in Court that the correct number was somewhere in the neighborhood of 3.1 million.

Q Okay. And so you got some new records, correct?

A At that point some time, I believe in July, I received an updated data set from the Department of Justice, and this data set incorporated -- they took all of the affidavits I received, incorporated the missing approximately 3.1 million records from the Texas DPS data.

Q And you analyzed those data, in effect, three times that you've gotten the data, and you reported in your report on how many no matches you found at that point, is that correct?

A Well, after I got the new data I basically just wrote the report again. I had to update most of the numbers in it based on the fact that my data set had changed.

Q And what was the number of no matches you found at that point?

A This was my -- I guess my late July number?
Q Yes.
A That was approximately 619,000.
Q That's 619,000 no matches, correct?
A That is correct.
Q Okay. And thereafter you got some more information about more no matches, or more people from the Department of Justice, is that correct?
A That's correct.
Q And what was that data?
A So I --
Q What was those data? I guess data is a pleural word.
A So I believe on August 7th or August 8th I received from Counsel a list of approximately 183,000 of voter IDs, these are IDs that indicate -- they correspond to Texas registered voters, and I was told that these IDs, for the most part, I think it's (indiscernible) because I -- the vast majority of these correspond to individuals whose licenses -- driver's licenses had been surrendered, and this is the first time that I had heard about a data source that -- that indicated who did -- who had surrendered a driver's license.
Q Okay. And at this point the deadlines were fast approaching, is that correct?
A Yes.
Q In other words, even once -- once the new data came in

from the State, 2.8 million, 3.1 million, we were in the middle of the summer with data -- with deadlines fast approaching, and very little time to do anything else, is that a fair -- a fair assessment?

A Yes.

Q Okay. And so were you asked -- were you told to assume that all 183,000 people should be put in the no match category?

A Well, Counsel told me that it was their understanding that of this list of approximately 183,000 individuals, that almost all of them had surrendered driver's licenses, but there was some uncertainty if a small group was dead. And I didn't know if that meant dead and a surrendered license, and I'm not sure that Counsel knew either, but if they were dead then, obviously, they shouldn't be part of my analysis at all, but Counsel didn't know and I didn't know how to take this 100 -- list of 183,000 and divide it into those two groups, so the only thing that I was advised to do, which I think was the right choice, is to treat all of those individuals as having surrendered licenses. And so I conducted another analysis where I took the nonmatched individuals who were a product of the July analysis, that's the post-3.1 million issue, and I subtracted the 183,000 approximately individuals from my match list and moved them to the no match list.

Q Okay. And so going on that assumption that all 183,000 should be in the no match list, what was the total number of no

matches that you came up with in your final version of your printed report due by the deadline?

A Approximately 800,000.

Q Okay. So you're not saying that you know that there were 800,000 matches, but that it could be up to that number subject to how many of those people might be dead?

A That's correct.

Q Okay. Did you analyze the 800,000 by race?

A Yes. I conducted the homogeneous and nearly homogeneous and ecological regression analyses on the 800,000 -- well, on the data set characterized by 800,000 nonmatches, yes.

Q Okay. And are the results in your report?

A Yes, on Pages 53 and 54.

Q Okay. And I'm going to --

(Counsel confer)

Q Okay, we're going to post -- we'd like to post Page 53 which is the example we used earlier for the logical bounds.

Is that where you reported your analysis of the homogeneous and logical bounds analysis?

A Yes.

Q Of the 800,000?

A Yes.

Q Okay.

(Counsel confer)

Q Tell us what this -- tell us what this chart or this

picture means?

A So this chart shows the results of the homogeneous and nearly homogeneous block group analysis, so you can see --

Q Shows both the homogeneous and the logical bounds, right?

A Yes.

Q Both of those two which are related, in effect?

A Yes.

THE WITNESS: Your Honor, may I use this pointer?

THE COURT: Yes.

THE WITNESS: So you can see on this axis right here that it's the X axis in this figure, it shows the homogeneity cutoff as it goes from 100 down to 99, down to 98, all the way down to 90. And here you have what's written as the "Identification for" -- excuse me, the "Identification possession percentage," this is akin to the match rate, okay?

So -- and there are three colors of figures here, there are -- and you can see in the Legend, if that's legible, up here we have this cloud of points describes the white identification percentage, it happens to be that those are higher; and in the middle we have data on the Hispanic identification percentage; and here we have data on the black identification percentage. Now I didn't choose them to be ordered this way, they just are in the data, that's the way things came out.

The top cloud of points, which is sort of pale white,

this is based on an analysis of homogeneous or very homogeneous block groups with respect to white. So this dot right here is the identification percentage in a collection of block groups that is completely white, and that rate is around 95 percent.

BY MR. DERFNER:

Q That's the dot, right?

A That's correct, that dot right there.

And then you can see, as I'm going to sort of move a little bit down and I'll cutoff, you can see here that as the homogeneity percentage decreases, so this is in the definition of homogeneity, you start to get bounds and, you know, these are akin to this one-ninth and two-ninths figure that I drew earlier, these bounds captured -- capture the range of possible values for white identification percentage, and those are what I called the logical bounds or they are accounting identity bounds.

And here is the exact same sort of analysis for Hispanics, and you can see here that this dot is around 90 percent, and then this is the dot for the Hispanic identification possession rate in block groups, there are at least 99 percent and so forth.

And here's the same situation for blacks.

Q Okay. Let's turn now to the ecological regression of the 800,000 people.

Now you described earlier the general method of

ecological progression. Tell us a little more specifically, what do you have to do to get to the ecological regression?
A Well, I put together --
Q What did you have to do in this case?
A Yes, I understand. I put together my data sets, and then I used common software, which is standard in the fields, to estimate this regression.
Q Okay. Now when you say "put together your data sets," what does that mean?
A Well, the exercise I went through earlier. I took the data that the Department of Justice sent me and I manipulated it with computer code, and I put it altogether.
Q That's -- well, that's the -- that's the results from match/no match process.

But what about -- you referred earlier to census block groups. Did you do something called "Geocoding?"
A Yes. In fact, this relates to the homogeneous previously -- the homogeneous block group analysis as well. One thing I had to do was try to figure out where the various individuals in the Texas -- registered voters in Texas lived, and so the exercise is called "Geocoding" or "Geolocating." And I needed to Geocode the locations of each individual so that I could know if each individual lived in a particular block group.
Q So you're using a method that's standard, and the mathematical formulas are standard in this ecological

regression, is that correct?
A Yes, that's correct.
Q But the input, such as the Geocoding, the selection of the block groups and some labeling of those, that you did, right?
A Yes.
Q Or --
A We're using the census data.
Q You used the census data. But you had to actually code for everything, and that -- and did you rely on anybody else, such as Dr. Ansolahehere to do that?
A I didn't rely on him for coding any addresses, no.
Q And if he coded addresses for his ecological regression, he didn't rely on you, did he?
A Not to my knowledge.
Q Okay. So is it possible that there would be slight differences between the output of results that the two of you got?
A Yes.
Q And is that achievable? Would that be akin, for example, to asking two people to measure this room and they both have the same tape measure, but two people are likely to wind up at least a couple of inches different from each other in their measurement?
A Yes.
Q Okay.

(Counsel confer)

MR. DERFNER: We have as demonstrative exhibit, your Honor, that we'd like to post, and this is -- well, we'll post it in the last ==

BY MR. DERFNER:

Q Dr. Herron, are you familiar with this chart?

A Yes.

Q And is this a chart that reflects data that you found and data that are in your report?

A Yes.

Q Okay. Tell us what this chart shows.

A This chart is a summary of the ecological regression that I carried out. I fit some lines and then used standard formulas from those lines to estimate the white identification percentage rate, and then correspondingly the white non-identification percentage rate. Of course, those two rates have to sum to 100, so if approximately 96.3 percent of whites have IDs then it must be true that approximately 3.7 percent don't, so one can easily talk about possession rates or nonpossession rates, those are two sides of the same coin. So here we have the results that -- explained in the context of nonpossession rates, so you see it's 3.7 percent approximately for whites, 8.3 percent for Hispanics approximately, and 11.5 percent approximately for blacks.

Q Okay, and tell me, what is the approximate ratio, feel

free to use the calculator if you like, what is the approximate ratio of the nonpossession rate for Hispanic to white?

A I would say it's around twice, roughly speaking.

Q Okay. And, again, what is the approximate ratio of the nonpossession rate among -- compared from blacks compared to whites?

A I would say it's approx. -- in the range of 3.

Q Three to one?

A Yes, three.

Q Okay. Now I'm going to post another chart, and this is another demonstrative, and this one has, on the left-hand side, Dr. Herron's figures, in fact, identical to what we just saw. And on the right-hand side are Dr. Ansolahehere's figures that were part of one of the charts that was posted yesterday, basically Dr. Ansolahehere's ecological regression results for the population. You remember that Dr. Ansolahehere had separate charts for people under 65, or listed as sub-categories, but for his overall figure, this was the chart.

And, Dr. Herron, how would you compare the numbers that you found and the numbers that Dr. Ansolahehere found as shown in that chart?

A I would say in the context of this process where I mean the data base match, manipulations of the data, the slightly different approaches we took, I would say that -- I would call these numbers effectively the same.

Q And this is even though you each independently performed the analysis yourselves, is that correct?
A Yes. I didn't consult with Dr. Ansolahehere's regression analysis when I carried out mine.
Q Okay. And --
(Counsel confer)
Q Okay, now, and this last -- the set on the left was what was reported in your report, correct?
A Yes, that is correct.
Q Your amended report?
A That is correct.
Q Okay. And after your amended report was issued, because it had to be by the deadline, did you then do some further checking on your numbers because you got a -- you were informed of what Dr. Ansolahehere's final no match number was?
A Yes, I did that.
Q Okay. And when was that? Your report was handed in on a Friday. When did you do that analysis?
A I believe I -- I believe I sent my report to Counsel on Thursday --
Q Okay.
A It was due on Friday, and --
Q That was nice to have it ahead of time.
A And the Monday following I received more information about Dr. Ansolahehere's nonmatch list; in fact, I was sent a file by

Counsel that contained that no match list.
Q And was that 786,000 names on his no match list?
A Yes, that was -- so at that point I didn't have Dr. Ansolahehere's report, and the data that I received on the Monday after the Thursday when I sent my report in, contained, among other things, the non -- one could infer from it the nonmatch list that Dr. Ansolahehere used, I believe, to generate this figure on the right.
Q And what was the number that was --
A Again, approximately 786,000.
Q Okay. And did you, as a further check, take this 786,000 number and do a racial analysis of that number?
A Yes, I took his nonmatch list and carried out the same sort of analysis that I carried out in my data, and that means I looked at the logical bounds for nearly homogeneous block groups and I looked at the homogeneous block groups, and I estimated an ecologic regression using his -- his nonmatch list as if I -- that sort of an exercise, as if I didn't even have my own, so I did his separately.
Q Now, the difference between what you were working with, which is the -- we'll call it the up to 800,000 number, and his 786,000 number, what's the difference between them?
A Around 14,000.
Q And what percentage is that of the -- of your no match numbers?

A So I'll calculate it. This -- I'm going to divide 14,000 divided by 800,000, then multiply that by 100, it's around 1.75 percent.
Q Okay, so less than 2 percent difference?
A Yes.
Q Okay. And would a lay person tend to say, "Well, if it's only less than 2 percent difference, I'm going to assume that the racial percentages don't change," is that what a lay person might say?
A In principle.
Q Okay, but an expert like you wouldn't say that, would you?
A Well, I'd want to check.
Q In fact, I tried to get you to say it looks the same, doesn't it and did you agree?
A I suppose.
Q So you actually did a new racial analysis of the 786,000, as you say, without sort of building it onto yours? You started from scratch and did a separate analysis?
A Yes. I effectively pretended I had a no match list of my own. I took his no match list and just did the analysis that I did in my reports on his no match list.
Q And what was the racial numbers compared to the numbers up here?
A Effectively the same.
Q So that was a check and you came up, again, with the same

numbers?
A Yeah, it's as if someone had shown up to me and said, "Throw out your data base, just use this no match list," and I said, "Okay," and then I just did the analysis on that, and then what I get is the exact numbers, I don't know off the top of my head. I remember that my white figure was exactly approximately 3.6 percent, which is what is in his ecological regression, you can see that here. And my Hispanic and black numbers were incredibly close.
Q One-tenth of a percent.
A I don't remember the exact numbers, but they were very close.
Q And this was based -- the analysis you did was based on your own Geocoding and your own programming?
A Yes. The -- just to be clear, I hired someone to Geocode since I'm not a geographer, I hired someone to Geocode the addresses for me. Yeah.
Q But that had nothing to do with -- you did not depend on Dr. Ansolahehere or anybody outside of you and your employee to do it?
A No, I depended on no one else.
Q Okay. And tell me the -- measure -- tell me what the percentage differences on these are? Let's take the whites, what's the percentage difference between you and Dr. Ansolahehere?

A So you want me to ask how much smaller is 3.6 percent and 3.7 percent?
Q Yes.
A Okay. So I'll just divide -- I'll take 3.6 minus 3.7, that's negative .01, and I'll divide this by 3.7, divide -- multiply by 100, that's around -- well, it's smaller, but 2.7 percent difference.
Q Okay. And are the percentage differences between the black figures and Hispanic figures also in that range, 2 or 3 percent?
A Sure. I could just divide -- for Hispanics, for example, I could just divide .3, that's 8.6 minus 8.3, divided by 8.3, multiply that by 100 and I get 3.61.
Q Okay.
A So, yes, they are all of that magnitude.
Q Okay. And that's the magnitude that you might expect for you and Dr. Ansolahehere did separate Geocoding, separate selection of block groups, et cetera, but using the same mathematical formula for the analysis?
A Yes. This doesn't -- these differences are in my experience completely consistent with the idea that there -- that we have two different people working on a common data set. There are some subtle changes in the way say I treated addresses and the way Dr. Ansolahehere treated addresses but in general we had the same idea and in situations like that in the

past when I have worked with other scholars and we've swapped data sets for example that you'll get differences of this magnitude.

Q And that confirms that you both really come to the same result?

A Well, it confirms -- yes. I mean it confirms that there's nothing idiosyncratic in what I did or what Dr. Ansolahehere here did that generated this.

Q You -- were you here yesterday for the examination of Dr. Ansolahehere?

A I was in the back row, yes, or second to back row.

Q And did you hear Mr. Scott suggest that there was some place where there's a 27 percent difference between your figure and Dr. Ansolahehere's figure?

A Yes, I heard that.

Q Okay. Did you look through your report and Dr. Ansolahehere's report last night to try to find some figure, figures that were comparable that were 27 percent off from each other?

A Yes. Well, what I did is I basically looked at the data that are behind this particular exhibit and I tried to check -- well, I did check whether any of these figures are 27 percent off and they're not but that's all I really checked.

Q Do you have any idea -- have you seen any place in your report that looks like figures are 27 percent away from a

comparable figure of Dr. Ansolahehere's?
A Well, the only -- strictly speaking, these are the comparable figures here and I don't in these figures observe 27 percent differences in the way you're describing them.
Q They're two or three percent, correct?
A Yes, much smaller.
Q Okay. So you've not seen anything that looks like a 27 percent difference so far?
A Not so far, not comparing these.
Q Maybe because (indiscernible). Okay.
I think we're almost done, your Honor.
THE COURT: Okay.
Q Professor Herron, did you do anything with your no match list to remove felons, dead wood, people who mean anything like that?
A No. After I generated my no match list I didn't remove people from it.
Q Now you were here yesterday when Mr. Scott asked Dr. Ansolahehere questions about convicted felons and the dramatic example of the woman in I guess it's Huntsville on death row.
How does that -- how could that potentially affect your analysis?
A Well, in principle it could affect my analysis if everyone on the no match list was a convicted felon.

I noticed two things in that presentation. One was that the number of convicted felons discussed was small; I believe it was one or two, and my no match list has approximately 800,000 people and -- so I was pleased that that number was small. Again, small is with respect to 800,000 and certainly the number discussed yesterday was small.

I was also pleased that I believe Ms. Baldwin at the Department of Justice brought some material to Dr. Ansolahehere's attention that explained that either the state or counties in Texas -- I can't remember which and I couldn't hear completely -- one of those agencies or agent -- one of those agencies or multiple daily attempt to remove felons, convicted felons excuse me, from the team database. And so, I was pleased to hear that that effort was done on a daily basis. Of course, it probably has some errors. It's a human process. But the fact that it's done daily as opposed to say every several years was reassuring to me.

Q So do you see any reason to think that the failure to remove convicted felons or find convicted felons on the no match list and remove them if they haven't already been removed by the process, would that likely have a significant effect on the numbers?

A I doubt it would have a significant effect in particular because that sort of error could be present on the match list as well. So I haven't seen any evidence that removing felons

would dramatically change my results in a statistically meaningful way.

Q Okay. There was also some discussion of people who voted in elections, actual elections, subsequent to the (indiscernible) deposition you were asked -- weren't you asked about the fact that some 32,000 people on your -- or 32,000 people from your no match list had voted in the November 2013 election?

Do you remember being asked about that in your deposition?

A Yes, I believe I was asked in my deposition if I knew that 32,000 people approximately on I think Dr. Ansolahehere's no match list --

Q Okay.

A -- but I'm not sure which no match list was the subject there -- had voted in I believe November 2013.

Q What does that tell you?

A That tells me a couple of things. One, it tells me that it's probably true that there are individuals on my no match list and Dr. Ansolahehere's no match list who probably have forms of identification. So those are errors and that's not surprising. When I began this project, meaning when I got the data, I noticed there were 13,400,000 people. It never occurred to me that I would or anyone would perfectly classify all of them.

So the fact that some people voted probably, you know, suggest to me that some number of them do have forms of identification. Now of course I can't say anything about precinct administration in Texas. So I don't believe that one can conclude that everyone who voted in that election had ID. I don't think that logically follows but I certainly think that probably some of them did.

I would also think that -- this also raises I think a related point which is that there are two types of errors that you can envision in this matching process. In one type of error or in one form of error, people who have ID are labeled as non-matches but people who don't have ID could also be labeled as matches. So errors like in most statistics work -- work because they work in two ways or both ways.

So when you mentioned it, well, perhaps 33,000 people might have IDs one thought that occurs to me is -- excuse me, 32,000 people who voted who I -- who someone considers not having IDs might actually have IDs. It occurs to me, well, it's also possible that approximately 33,000 other people who don't have IDs were labeled as having IDs. So that would be the other side of this error and as someone who does apply statistics, my interest is not just saying, "Okay, do we have an example of an error?" But rather, "Do we have an example of a process that systematically generates errors that will bias results?"

So if we have an example where some people who have IDs are labeled as not having one and some people who don't have IDs are labeled as having IDs, that would sort of lower all the numbers but leave percentages intact. So as you can see from this chart on the -- on the wall right here, I conclude that approximately 3.7 percent of whites don't have IDs and I believe this is equivalent to approximately 4 percent of all registered voters not having IDs. That 4 percent figure might be completely accurate if some people -- if errors work in both ways I've described. So that thought occurred to me when thinking about the possible 33,000 people.

Q One last question. I should have asked this before.

You referred -- you mentioned that after your report was done, the last report, you did that. You checked your numbers against Dr. Ansolahehere's 786,000 (indiscernible).

Did you explain that -- even though it's not in your report, did you explain that at your deposition?

A Yes, I discussed that in my deposition because it was, it was late.

Q And so in fact the defendants had an opportunity to examine you about that last little piece of information?

A Yes. I brought to -- or counsel brought to the deposition the bounds figure that I produced and corresponding to that analyses.

MR. DERFNER: Thank you very much, Professor Herron.

It's been a pleasure and now I think you belong to Mr. Scott.

THE WITNESS: Thank you.

MR. SCOTT: Well, let's not go that far. Hi, Dr. Herron.

THE WITNESS: Good morning.

CROSS EXAMINATION

BY MR. SCOTT:

Q Now you did a declaration in this case -- well, hold on a second. I'm going to leave that Elmo on for just a second.

(Pause)

Q You did a declaration in this case on -- signed on August 14th, 2014, correct?

A I believe that's the right date, yes, sir.

Q So in that -- this is page seven of that report, correct?

MR. DERFNER: Excuse me, your Honor, may I come over here?

THE COURT: Yes.

A Yes, I believe that's right.

Q And the summary of your main findings, number one, of the 13,564,410 registered voters in Texas conservatively speaking, 619,354 lack valid forms of voter -- of ID to vote, correct?

A Yes, that's from the -- right, yes, sir.

Q So you've got a calculator. You used it a second ago with Mr. Derfner. Would you pull it out for me?

A Yes, I have it.

Q Will you subtract or enter 786,727?
A Yes, I've done that.
Q Now hit the minus key.
A Done.
Q Put 619,354.
A Done.
Q Now hit equal.
A Done.
Q What number do you have?
A I have 167,373.
Q Would you divide that by 619,354 and what number do you get?
A Point 270 approximately.
Q Point 270 is I guess 27 percent? Is that right?
A If I multiply it by 100, yes.
Q Yes. And so your -- the difference between your number and 786,727 is 27 percent, right?
A When you say "your number," you're reading the number on line 9. There are other numbers in this report.
Q Fair point. We'll get to that in a little bit but when we're looking at your summary of main findings and number one of your summary of main findings you have 619,354, correct?
A Sir, there are several other main findings here but yes, if you ask me the difference between 619,000 and that, that difference is approximately 27 percent.

Q Sure. And I know last night you spent a lot of time trying to come up with a number that was 27 percent differential based upon what I had asked some questions yesterday, correct? I think that was the colloquy between you and Mr. Derfner.
A I looked at my regression results and compared them to his, yes.
Q Okay. So at least we've found that missing 27 percent number, right?
A Yes.
Q Let's talk if we could a little bit about ecological regression and homogeneous block analysis.

You did both methods in this case, correct?

A Block groups to be precise.
Q Block groups. And so you were talking up there and I started getting this mental image.

You ever been on Google Earth?

A I believe so.
Q You ever looked at your house from Google Earth?
A I've looked at my house on some mapping program. Whether it's Google Earth I'm not sure.
Q Well, on Google Earth you can put an address in there and you can go all the way down and just in your mind's eye you can see the outline of your house. You can see the outside of the house but you don't really get to see inside the house, right?

A I understand the principle, sure.
Q Sure. And so from an ecological regression method that's very similar. You're not -- you're not telling us what's actually inside any individual home, right?
A That's not the purpose of an ecological regression.
Q Sure, I understand that and I'm not here to debate whether it's the right way or wrong way to look at stuff. I guess what I'm asking you, sir, is you're not able to really go into the details of what's in that home. That's -- you're making an estimate. You're looking at it in a macro sense, correct?
A Yes. I think it's right to say I don't know what's going on in the individual house.
Q And so with regard to a number of the things and the two different analysis that you performed in this case, you're not able to tell this court with any accuracy whatsoever on any individual voter whether that individual possesses ID that complies with the terms of SB 14, correct?
A No.
Q That's correct. You're not here to testify about individual voters, correct?
A Well, I think there are two parts in this question. I'm not here to testify about individual voters. However, in a homogeneous block group I do know that if someone is registered to vote and so forth and is on my no match list for example, and if a homogeneous block group is 100 percent white then I'll

be able to say something about that person.
Q Well, in the homogeneous block groups you -- you testified about the homogenous block groups that you were able to derive in this case, correct?
A Yes, I did.
Q And I think that if my recollection's right you came up with about a total of 213 different block groups in this case, correct?
A Yes, that's correct.
Q And that represents almost 104,000 voters in the State of Texas, correct?
A I believe that's right, yes, sir.
Q And we have how many registered voters in the State of Texas?
A Around 13.4 million.
Q And so what is the percentage that 103,585 voters is of the 13.5 million?
A It's small.
Q Well, let's use your calculator. So would you enter 103,585 and would you divide that by 13,564,000?
A Excuse me, I made a mistake. Could you tell me the first number again?
Q Sure, 103,585 and for clarification in the record, that's the number of people you identified in your homogeneous block group, correct?

A There was 103,000 registered voters --
Q Yes, sir.
A -- or individuals? Yes. Those are the -- that's the number of individuals, 103,000, who live in these homogeneous block groups. That is correct.
Q Okay. So you punched in 103,585, correct?
A Yes, sir.
Q Now divide that by 13 million --
A One, three --
Q Five, six, four --
A Five, six, four.
Q -- Zero, zero, zero.
A Zero, zero, zero. Okay.
Q What do we get?
A I multiply by 100 again?
Q Well, what do you get right now?
A Well, I got .007.
Q So now times it times 100.
A And it's done.
Q And what number is that?
A Point seven six approximately.
Q And so it is seven tenths of one percent?
A Yes, if we were rounding I would say eight but --
Q Okay. Well, it says round. So less than eight percent of the state's population of registered voters goes into your

homogeneous block analysis, correct?
A Just to be clear, this is the completely homogeneous because --
Q Yes.
A -- there was some discussion of nearly homogeneous. That is correct.
Q And with regard to these groups that are 100 percent, where there's a high degree of accuracy from your standpoint, correct, that's the benefit of a homogeneous group, correct?
A Effectively so, yes.
Q So can you tell me the income makeup of those homogeneous groups?
A The income?
Q Sure, individual income.
A No, I -- I didn't look at any income figures in my report.
Q Can you tell me the voting history within any of those homogeneous block groups? Again, we're talking about these small groups, less than one percent, that you've been able to identify and tell us specifically here is 100 percent of a category, a race of people within this geographic area. Can you tell me what the voting history of that group is?
A Do you mean did I do so for the purposes of this?
Q Sure.
A No, I did not.
Q Were you asked to?

A No.
Q Can you tell me the rates of possession of birth certificates within that homogeneous group?
A Birth certificates?
Q Yes.
A No. I don't believe I could do that either.
Q Did you undertake to determine in these 100 percent homogeneous block groups that identified specific races the compliance -- I mean what percentage of those folks had the documents necessary to be able to obtain a photo ID that would comply with SB 14?
A No, I don't have data on that.
Q Did you provide the information about these groups that were 100 percent homogeneous and so there's no mistake that we're talking about a race specific within these groups, these block groups? Did you provide that information for any of the other experts to use in this case for any reason?
A That information is implicit in my data and I was asked to circulate my data. So I think the answer to that is yes.
Q And who did you provide it to?
A I sent four Fed Ex packages in the last -- on the Friday that we were talking about earlier to various counsels I believe.
Q Well, you got some material from the Department of Justice, about 183,000 names; is that correct?

A Are we talking about the August 8th -- excuse me, I don't even want to use that date.

Are we talking about the data that I received after my amended report was complete?

Q Yes. So we're talking about information that you received before August 14th, the report, the declaration that's filed in this case, correct?

A I have to look at my calendar. I'm confused on the dates here.

Q Sure. It's 183,000 people that you were notified were suspended -- had suspended licenses? I think that was your testimony.

A Right, those I received before my amended report was done, correct, yes.

MR. DERFNER: Object, that's not exactly -- not suspended. I thought the word was surrendered.

Q Oh. Well, tell me Doctor, what does a surrendered license in Texas entail?

A Counsel explained to me that "surrendered" means that an individual has to hand over a license.

Q Who told you that?

A Counsel.

Q Your counsel?

A Yes.

Q And did he tell you that was based on Texas law or did he

tell you that was based on a fact that actually happened with these 183,000 people?

A She, and I don't recall asking.

Q Was it important to you to understand if all 183,000 or any of the 183,000 people that you were identified and ultimately put on the no match list had actually surrendered their physical license to someone?

A I was told that these were individuals who had surrendered their licenses for the most part; again, subject to some of them possibly being dead.

Q So other than the people who are dead I think you already agreed that they shouldn't be on the list to begin with but those folks would be the only ones you excluded out of that 183,000, correct, off your no match number, your new no match number?

A I was given a list of 183,000 and I was told that for the most part these were individuals who had surrendered their licenses but it is possible -- this is what I was told by counsel -- that some of these individuals were dead and it is I believe possible that there might have been a very small number of other individuals although I'm not sure if it was -- the only exception was that. I cannot recall for sure but I didn't inquire as to the exact mechanism of handing in a license because I was told that they had surrendered it and that was sufficient for me.

THE COURT: Was that number matched to see if they had other SB 14 ID or what did we do with that number?

MR. SCOTT: That's what -- I was going to --

THE COURT: Okay, okay.

MR. SCOTT: I was going to try -- I agree with you, your Honor.

THE COURT: Sorry, go ahead.

Q And so was there a process that was undertaken on another no match process that was done with the Department of Public Safety to confirm that those individuals were (a) without their licenses so that they'd actually physically surrendered those, those materials, before you included them in your no match list?

A Yes, sure, I can tell you what I was told which is that these were individuals who had surrendered their licenses and they didn't have other forms of SB 14 qualifying identification.

Q Okay.

A Obviously without that last bit if for example all of them had passports then this would be a different world.

Q So you would agree that none of those individuals to the extent they have other ID, if somebody could show that, you would agree that they would not be properly included in the no match list, correct? Again, you were just given data and told to believe this is true, correct?

A Yes, if you could show -- I mean if you could show me that some of the -- some ethogeneric to be true, if there's someone on my no match list that had ID then that person shouldn't be on the no match list.
Q Which of the two ways that you performed your analysis, the ecological regression analysis or the homogeneous block --
A Group.
Q -- group analysis, do you have more faith in? Which one's better?
A You know in my experience as a I guess a scientist the hallmark of good science is multiple methods and so the reason I used two methods here was not because I didn't have faith in one of them it's because every method has strengths and weaknesses as I explained earlier. And so, I think the goal of any endeavor in this instance is to try to find methods that complement each other and then to ask if they show different things or the same things. In my case, they show the same thing effectively. So I'm not comfortable saying that I trust one more than the other because they're compliments.
Q And again, when we're talking about either of those two groups we're not really talking about individuals in a Section Two case who we are trying to determine if they have -- if they don't have the proper ID in order to actually cast a ballot, correct?
A I'm not sure.

Q Well, I mean, again, we're talking -- you're not able -- you didn't do a database match, correct?
A No, as I explained I didn't carry that process out.
Q I mean is voter turnout important from the standpoint of your analysis in this case, the effect on voter turnout?
A No, you know, I was asked what I was asked to do and I was asked to study the database and access non-match rates. I wasn't asked to look at turnout.
Q Well, is it important to know whether any individual you're identifying in your group has not had the ability to vote as a result of SB 14?
A If the question is who has ID then the answer's no.
Q Well, the question is do you believe in your opinion that any individual within your group has been deprived of the ability to vote as a result of SB 14? Do you have an opinion on that?
A I wasn't asked to study who's been deprived of rights to vote. I was asked to study who has IDs.
Q Do you know how many cities in Texas have -- well, strike that.

You used to teach with Dr. Ansolahehere?

A I don't believe I ever taught with him.
Q Oh, I thought you were at Harvard at the same time.
A I visited at Harvard for a semester. I'm not sure if he was on the faculty then.

Q Okay. Have you read any of his articles that he's published?
A Over the course of my career I'm confident I've read some of them.
MR. SCOTT: Brian, will you put the first article up from Dr. Ansolahehere?
Q Have you read Effects of Identification Requirements on Voting, Evidence From Experiences of Voters on Election Day that was written by Dr. Stephen Ansolahehere?
A I'm looking for the date on this, I'm sorry. How can I make this a bit bigger?
Q Well, you and I can't.
A Oh, thank you.
At some point I believe I've read this but I cannot testify now as to all of its conclusions.
Q Well, there's one conclusion I'd like to show you.
A Thank you.
MR. SCOTT: Let's see if I can't find it here.
(Pause)
MR. SCOTT: There, it's highlighted.
Q So some of the denials -- this is occasionally -- this is an exceptionally low rate of (indiscernible) access to the vote. Some of these denials may have been legitimate and some of them may have been erroneous but the actual analysis of the vote in these two surveys suggests that photo ID laws may

prevent almost no one from voting.

Do you agree or disagree with Dr. Ansolahehere?

A I didn't write this paper.

Q I know, that's why I asked you do you agree or disagree.

A I mean if you want me to give a scholarly opinion on this you would have to give me his data and ask me to analyze it which I could do but until I've done that I'm hesitant to sort of say "Yes, this is right" or "No, this is wrong."

MR. SCOTT: Okay. Brian, can you bring up another article by Dr. Ansolahehere? Oh, I'm sorry, it's the second page.

Q There's another page, I'm sorry, and it goes on to say "Voter ID does not appear to present a significant barrier to voting." Skip down a little. "Although the debate over this issue is often draped in the language of civil and voting rights movements, voter ID appears to be -- to present no real barrier to access."

Do you agree with Dr. Ansolahehere?

A I have the same answer. I mean this is a statement about analysis of some data that I haven't done.

MR. SCOTT: Pass the witness.

MS. BALDWIN: I have a couple of quick questions. Anna Baldwin for the United States.

THE WITNESS: Good morning.

MS. BALDWIN: Thank you, Dr. Herron.

CROSS EXAMINATION

BY MS. BALDWIN:

Q The last article from Dr. Ansolahehere that was up did you see very briefly on that first page -- could we take a look at that again?

(Pause)

Q And could we -- that really small date at the bottom, could we zoom in on that?

(No audible response)

Q In January 2009, do you know the number of states that Dr. Ansolahehere was looking at in this article that had strict photo ID laws akin to the law that we're talking about here today in Texas?

A I don't know the number off the top of my head but I'm confident that it's very small.

Q Do you know if there were any strict photo ID laws akin to Texas' that were in place prior to January 2009?

A I would venture to say that there is nothing as strict as Texas'. So the number is probably zero.

Q I just want to pull up one other exhibit briefly, Plaintiff's 944, if we can switch --

(Pause)

Q Could you read the title of this document that's underlined if you can?

A Could you make it a bit bigger?

Q Yes, can we get that a little --
A Should I read?
Q Please.
A DPS Responses to Written Deposition Questions Pursuant to Rule 31 Regarding Specified Topics From the United States' Notice of Rule 30(b)(6) Deposition of the Texas Department of Public Safety.
Q Great. Could we go to the second page of this document and on question number two and three -- question number two lists the number of card statuses and then if we could see the answer. Can you scroll down a little farther?
(Pause)
Q The answer here says if a card has a status of voluntary surrender, voluntary surrender CSO, or voluntary surrender insurance, or voluntary surrender medical, it means the customer has voluntarily surrendered their card.
Do you agree that such people should be placed on the no match list if that's the only form of ID that they have based on this answer from DPS?
A I'm going to read the complete answer.
Q Sure.
(Pause)
A Yes.
MS. BALDWIN: Thank you, Dr. Herron.
MR. DERFNER: I've got -- I don't think

(indiscernible).

THE COURT: Yes. We've -- someone else has redirected. Maybe another plaintiff we can call as a cross by another party, however we want to designate that. The state gets to question again.

MR. SCOTT: So could we pull that exhibit back down?

(Voices off the record)

CROSS EXAMINATION

BY MR. SCOTT:

Q Do you know as we sit here today when a person physically -- out of that 183,000 who surrenders a license by turning that license in, do you know how many of those people get a different -- a different ID card back?

A Do you mean in response to handing in the license?

Q Yes, sir.

A Do you mean an SB 14 qualifying ID?

Q Yes, why not?

A So you're asking me if an individual hands in the license and is handed a what?

Q It sounded like we all agreed an ID that qualifies under the terms of Senate Bill 14.

A Well, I know that -- I know the list of IDs that qualify and I don't believe -- I have no reason to believe that anyone handing in a driver's license would be handed another form of ID in exchange.

MR. SCOTT: Perfect, thank you. Pass the witness.

(Pause)

MR. DERFNER: Hold on a second.

(Pause)

REDIRECT EXAMINATION

BY MR. DERFNER:

Q Do you know that any of those people, any of these 183,000 people, had another form of qualifying ID?

A I was told that the reason they were on that list is because they didn't have another form of qualifying ID.

MR. DERFNER: Thank you. We'll be -- you'll be hearing a lot more about this in later witnesses, your Honor.

THE COURT: All right. Anything else for this witness?

MR. SCOTT: Nothing from the defense.

THE COURT: You can step down, sir, and we'll go ahead and take our 15 minute morning break.

COURTROOM DEPUTY: All rise.

(A recess was taken from 10:08 a.m. to 10:24 a.m.; parties present)

MS. VAN DALEN: The Plaintiffs call Eulalio Mendoza -- or Mendez.

(Pause / Voices heard off the record)

THE COURT: Good morning, sir. Okay. Hold on one second. I'm going to have him sworn in so just tell him to