IN THE UNITED STATES DISTRICT COURT

FOR THE WESTERN DISTRICT OF WISCONSIN

WILLIAM WHITFORD, et al.,

Plaintiffs,

vs. Case No. 15-CV-421-bbc

GERALD NICHOL, et al.,

Defendants.

DEPOSITION OF

NICHOLAS GOEDERT

Milwaukee, Wisconsin

December 15, 2015

8:42 a.m. to 3:27 p.m.

Laura L. Kolnik, RPR/RMR/CRR

A P P E A R A N C E S

FOR THE PLAINTIFFS:
EARLE LAW OFFICE
MR. PETER GUYON EARLE
839 North Jefferson Street
Milwaukee, Wisconsin 53202
peter@earle-law.com
(414) 276-1076
CHICAGO LAWYERS' COMMITTEE FOR CIVIL RIGHTS UNDER THE LAW, INC.,
MS. RUTH GREENWOOD, Lead Attorney Voting Rights Project
MS. ANNABELLE HARLESS, Staff Attorney
100 North LaSalle Street, Suite 600
Chicago, Illinois 60602
rgreenwood@clccrul.org
aharless@clccrul.org
(312) 202-3469
NICHOLAS ODYSSEAS STEPHANOPOULOS
University of Chicago Law School
1111 East 60th Street
Chicago, Illinois 60637
nsteph@uchicago.edu
(781) 248-8145

FOR THE DEFENDANTS:
STATE OF WISCONSIN DEPARTMENT OF JUSTICE
MR. BRIAN P. KEENAN, Assistant Attorney General
17 West Main Street
P.O. Box 7857
Madison, Wisconsin 53707-7857
keenanbp@doj.state.wi.us
(608) 266-0020

I N D E X

(Original transcript supplied to Attorney Greenwood.)
(Original exhibits attached to original transcript.)

E X H I B I T S

P R O C E E D I N G S

NICHOLAS GOEDERT, called as a witness herein, having been first duly sworn, was examined and testified as follows:

EXAMINATION

BY MR. EARLE:
Q. Good morning.
A. Morning.
Q. You mind if I call you Nick?
A. Sure.
Q. Okay. Good. You can call me Peter.
A. Okay.
Q. All right. Nick, what did you do to prepare for today's deposition?
A. You mean after I filed the report?
Q. Yeah.
A. I -- sorry, I reread my report, I reread the various filings in the case from both the plaintiffs and the state. I looked over the other experts' reports on the -- on the plaintiffs' side. I briefly looked over some other articles that I thought related to the case. I discussed a little bit with Brian Keenan what he anticipated from the deposition. That's pretty much it.
Q. Could you list the articles you reviewed for me that

you -- that you thought -- whether -- how did you characterize them? Articles you thought were --
A. Related to the case.
Q. -- related to the case. Okay.
A. Yeah. I read Stephanopoulos and McGhee's article. I read an article in Studies Quarterly from 2014. I read -- I mean those are the two that I read most closely.
Q. Had you read those before?
A. Yes, I had. Yes.
Q. Okay. So did you read anything new?
A. Did I read anything new? I don't believe so.
Q. Okay.
A. I don't believe I read anything that was not cited in the -- in the report.
Q. Okay. We'll get into some more detail on this stuff. Let's start with some basic rules. I assume you've never been deposed before; is that correct?
A. I've never been deposed before.
Q. Okay. Now, I do know you're a lawyer. So you --
A. I do have a law degree.
Q. So you have an idea what a deposition is, correct?
A. Yes. Yes.
Q. Okay. Rules. We have a court reporter here, and she's taking down everything we say verbatim. And

the whole purpose of this is to get a transcript with a sequence of a question and an answer, and it's important we don't talk over each other for that reason.

Sometimes during a deposition people lapse into conversational mode and you kind of preempt my question by answering it because you know where I'm going, and you might be accurate. In a conversation that might be totally normal and comfortable and it means we've acquired a level of comfort with each other, but it's going to be very tough on the court reporter and it will make the transcript less readable, so I need you to not answer my question until I finish it. Is that okay?
A. Okay.
Q. Okay. If you don't understand my question and you answer it, anybody reading this transcript, including the court, will assume you understood the question and that your answer was intentional to the question as worded. So if you don't understand the question, you need to clarify that with me and ask me or tell me you don't understand the question, ask me to rephrase, and I will. Okay?
A. Okay.
Q. Um-hum, hu-ugh, all those kinds of noises that

people make to simulate a yes or a no, they don't work in a deposition. You have to actually articulate your answer. Okay?
A. Okay.
Q. You're under oath so you understand that it is a crime to not be completely truthful in this deposition?
A. Yes, I understand.
Q. Okay. Now, I always ask my deponents if they would be kind enough to be completely candid with me. Some people can artfully answer a question in a literal way and shave their answer so as to distort the context or meaning or perhaps not be fully forthcoming. Do you know what I'm talking about when I describe that?
A. I think so.
Q. Do you promise to avoid that with me in this deposition?
A. I will try.
Q. Okay. So you will be fully candid and answer all my questions with complete -- all relevant facts and background and so forth, right?
A. Yes.
Q. Okay. Good. Okay. Do you have any questions for me?

A. I don't think so.
Q. Okay. Let's start with the subpoena. Did you receive a subpoena?
A. I did receive it.
Q. Okay. We'll mark this as Exhibit --
A. I don't have a copy with me.
Q. I'm going to give you one.

(Exhibit No. 16 marked for identification.)

Q. Showing you what's been marked as Exhibit 16.
A. Okay.
Q. Have you seen this document before?
A. I have.
Q. Okay. And drawing your attention to the third page of Exhibit 16, documents to -- to be produced by Nicholas --
A. Goedert.
Q. -- Goedert, did you review the -- the 12 category -- the 17 categories of documents on pages 3 and 4 of Exhibit 16?
A. I did briefly.
Q. Okay. Is there anything that you did not produce that's listed amongst items 1 through 17?

MR. KEENAN: I'll assert an objection that we did make a written objection to the subpoena for producing books that are publicly available that

Page 10

would be burdensome for producing.
THE WITNESS: (Witness reading.)
BY MR. EARLE:
Q. Want the question reread?
A. Sorry?
Q. Do you want the question reread to you?
A. I'm just reviewing everything. I want to make sure. (Witness reading.) I believe I did with the exception of number 16. I didn't provide copies of the -- the Wonkblog or Monkey Cage blog posts.
Q. Why didn't you do that?
A. It was an oversight. I -- I did not rely on those in the -- in this case.
Q. This is a compulsory process. I asked you to produce them in a subpoena to a deposition. And the reason that you didn't do that is you -- it was an oversight?
A. Yes.
Q. Okay.
MR. KEENAN: Would you like to Google them so you can get them?
MR. EARLE: Well, I would ask that the -- so that I don't miss one, that the deponent during one of the breaks Google them and perhaps email them to me and I'll print them out.

Page 11

THE WITNESS: Okay.
BY MR. EARLE:
Q. Is that acceptable?
A. That's fine.
Q. Okay. Good. Anything else?
A. I don't believe so.
Q. Okay. Just out of -- as an aside, did you review any materials from the Baldus case?
A. No.
Q. Are you familiar with what the Baldus case is?
A. Not particularly familiar.
Q. Okay. Do you have any idea what the Baldus case is?
A. I recall it being referred to in the -- some of the filings for this case.
Q. Okay. And what do you recall about that?
A. Not very much.
Q. Okay. Are you familiar with whether there was prior litigation involving Act 43?
A. I am vaguely aware that there was litigation involving Latino representation in one or two particular districts.
Q. Okay. And anything else?
A. I --
Q. Well, I guess let's back up. I'll withdraw that question and rephrase.

Page 12

Did you -- do you have any knowledge of prior litigation involving Act 43?
A. Not beyond what was mentioned in the complaints that I read.
Q. Okay.
A. Or the other -- the other filings in this case.
Q. Did you ask to see any discovery from prior litigation relating to Act 43?
A. No.
Q. Is there a reason you did not ask to see discovery documents from prior litigation?
A. It didn't strike me as relevant to my report.
Q. Okay. Let's go to your -- to your resumT.
Before -- before we do that, let me ask you another couple questions. Who all did you speak to to prepare for this deposition other than counsel?
A. I didn't speak to anyone.
Q. You didn't speak to Nolan McCarty?
A. I did not.
Q. How about Joey Chen?
A. I did not. I will say that I mentioned to Brandice Canes-Wrone that I was considering serving as an expert witness in this case and asked her opinion on it. This was prior to my coming on as a -- as a witness in the first place.

Page 13

Q. Okay. And so who was this person you --
A. Brandice Canes-Wrone. She is a professor of politics at Princeton. She was my graduate school advisor.
Q. And would you spell her name for the court reporter?
A. B-R-A-N-D-I-C-E is her first name. Last name is Canes, C-A-N-E-S hyphen W-R-O-N-E, Canes-Wrone.
Q. Would you describe that conversation in more detail, please?
A. It was an email correspondence.
Q. Uh-huh.
A. I had just emailed her mentioning that an attorney for the State of Wisconsin had called me and asked me to -- if I was interested in serving as an expert witness. I mentioned a couple of the expert witnesses -- I mentioned both of the expert witnesses that were testifying on the plaintiffs' side, and I think I gave her a little one-sentence background on the case, and I asked her if she thought it was a good idea to serve as an expert witness in the case given that I had never served as an expert witness before.
She replied back the next day that she saw no problem with it and thought it was a perfectly fine idea. That's -- that is the only correspondence

Page 14

that I've had with anyone outside of friends who would not have any particular knowledge of -- friends and family who would not have any particular knowledge about the case.
Q. Okay.
A. I'm sorry, I guess I should say that I have mentioned this to other colleagues of my school or other colleagues, not in any way who would have any knowledge about the case, just to give background about myself and what I was doing with my time.
Q. Such as who?
A. Such as Bruce Murphy who is a professor at Lafayette. Such as Joshua Miller who is a professor at Lafayette. Again these are not any people who would have any information about the case or any insight into the case, just to mention sort of my professional responsibility to correspond with other people in my department that I am doing this work.
Q. And just so I'm clear and the record is complete, other than the people you've mentioned, you've not discussed your work in this case with anyone outside of counsel for the defendants in this case?
A. I have not -- I have not discussed my work in the case at all outside of telling people that I was serving as an expert witness on the case.

Page 15

Q. Okay. Do you know why the State of Wisconsin approached you to serve as an expert in this case?
A. I am not certain. It is my impression that Brian Keenan had read some of my articles that were available online related to redistricting, and that was probably where he got the background from.
I believe he also visited my academic website and looked up my background and some of the articles that I had written prior to contacting me, but I am not certain why the State of Wisconsin recruited me as an expert witness.
Q. Okay. Have you spoken with Sean Trende?
A. I have not.
Q. Do you know Sean Trende?
A. I do not know him personally. I have never met him in person. I am aware that he's a journalist who writes for Real Clear Politics, and I do read articles on their website. But outside of reading some of his work just casually, I do not know him.
Q. Okay. Did you review his report in this case?
A. I did not.
Q. So you've never read that report?
A. I have not read that report.
Q. So it's accurate to say that you have no knowledge as to what is in that report?

Page 16

A. Brian Keenan told me that he was doing some specific work related to partisan dispersion in Wisconsin in that report but only in the vaguest terms.
Q. Okay. Well, let's go to your resumT or your CV. I guess it's attached to your report. We'll mark that as Exhibit 17.
(Exhibit No. 17 marked for identification.)
Q. Okay. It's fair to say that you're fairly new in the field of academia, correct?
A. I suppose it depends on what you mean by "new." You can see --
Q. Post-graduate.
A. -- on my resumT --
Q. Post-graduate.
A. -- I received my Ph.D. three years ago.
Q. And you have three years of experience teaching?
A. Three-and-a-half. Yes.
Q. Three-and-a-half. Okay. Well, let's start from your current position as a visiting professor in the Department of Government and Law at Lafayette College, correct?
A. Yes.
Q. Okay. Is that a tenure track position?
A. It is not.
Q. Why don't you have a tenure track position?

Page 17

A. I have not received a tenure track position yet.
Q. Well, have you applied for any?
A. I have applied for many.
Q. How come you haven't been hired by anybody?
A. I don't have any knowledge of why a particular job would not hire me.
Q. You've not got any feedback as to why you weren't able to get a tenure track position at any college or university in the United States?
A. I don't know if you'd want to -- me to discuss the background of how the applying for jobs, applying for academic jobs work, but typically if you apply to a job and do not at least receive an interview, you would not get any feedback as to why you were not selected.
Q. Am I to -- does that imply -- are you intentionally trying to imply that you did not receive any interviews?
A. I am not trying to imply that. I have received --
Q. Okay.
A. -- a few interviews. Yes.
Q. How many interview did you receive?
A. Are you referring to campus interviews or are you referring to --
Q. I'm referring to any kind of interview --

Page 18

A. Okay.
Q. -- in the -- in the job application process.
A. I believe during the time that I have been applying for tenure track jobs I have received five interviews in some form or other for tenure track jobs.
Q. And how many applications have you placed with colleges and universities?
A. Over what time period?
Q. Over the entire time period you've been applying for tenure track positions.
A. I don't have a precise number. It would be over 100 and less than 200.
MR. EARLE: Can we take a quick break?
(Discussion held off the record.)
BY MR. EARLE:
Q. So we went off the record. You indicated that you applied for more than 100 positions, but less than 200?
A. Yes.
Q. And you got five interviews?
A. Let me just -- yes. I believe that's correct.
Q. And where were those five interviews?
A. One was at Bard College, one was at Lafayette College, one was at -- I am -- I'm slightly hesitant

Page 19

to talk about interviews that are ongoing in some sense if this is public record.
Q. Well, the problem, Nick, is that I'm trying to assess the -- the quality of your experience, knowledge, and qualifications, and we -- they're being presented to the court in the context of this case as a person who's an expert. And the court's going to have to evaluate the extent to which you're qualified to give opinions, and -- and in academia, being able to get hired by a university or college is important.
MR. KEENAN: I would just object to the speech as to the importance not necessarily to the court, but I think you should answer the questions.
THE WITNESS: Okay. Old Dominion University in Virginia.
BY MR. EARLE:
Q. Old?
A. Dominion. Virginia Tech and University of North Carolina-Wilmington, I believe.
Q. Did any of those five give you reasons as to why you were not hired?
A. In the case of some of those I am not sure that I have not been hired. They have not completed the process of deciding on who to hire yet.

Page 20

Q. Okay. So you have applications pending at this moment?
A. I have many applications pending. Yes.
Q. Okay. And you have -- but you identified the universities or colleges for which you have applications pending where you have been interviewed?
A. Old Dominion, Lafayette, Virginia Tech.
Q. Okay. Okay. And Lafayette is where you're currently visiting --
A. Yes. I'm sorry.
Q. -- as an assistant professor? Yeah, we have to take turns. See how easy it is to lapse into comfortable conversation?
Okay. Did any of these folks, and for the ones that did not hire you, indicate why?
A. You're speaking of the two where I had interviews in previous years?
Q. Uh-huh.
A. No.
Q. Do you have any perception yourself as to why you have not been successful in landing a tenure track position at this point in your career?
A. I don't have any specific knowledge.
Q. No, but I asked you if you had a perception.

Page 21

MR. KEENAN: I'll just object to the relevance of someone's perception. If you have one, you can answer.
THE WITNESS: I don't have a perception.
BY MR. EARLE:
Q. Are you confident that you're going to get a tenure track position in the near future?
A. Depends on what you mean by "near future."
Q. Well, in the next couple of years?
A. Yes.
Q. How long has your application for a tenure track position at Lafayette College been pending?
A. Six weeks.
Q. So you put that application in after you started working as a visiting citizen -- visiting assistant professor, correct?
A. Yes.
Q. All right. When you were a post-doctoral research associate at Washington University, who did you work for?
A. I was doing my own independent research. I suppose indirectly you could say I worked for Jim Spriggs, but I was not working on his research projects. I was working on my own research projects.
Q. What was his name?

Page 22

A. Jim Spriggs or James Spriggs.
Q. James Spriggs. Okay.
A. S-P-R-I-G-G-S.
Q. And he was your supervisor?
A. Only in the sense that he was the one who hired me. He did not directly supervise my research in any meaningful way.
Q. Okay. And what was the research you were working on in that position?
A. I was working on my research dealing with various aspects of legislative elections, including turning my dissertation into publishable articles and other articles related to legislative elections.
Q. Anything else?
A. I don't think so.
Q. Okay. Do you know Simon Jackman?
A. I have met him very briefly. It was several years ago while I was a graduate student at Princeton. I know like him by reputation.
Q. Okay. Would you describe that reputation for me, please? Or at least your perception of that --
A. My perception --
Q. Wait a minute. Hold it. We have -- would you please -- I'll withdraw that question.
Will you please describe your perception of the

Page 23

reputation of Simon Jackman?
A. My perception is that he has an excellent reputation overall in political science, particularly in dealing with quantitative methodology and developing statistical packages for use in political science.
Q. Do you consider him authoritative?
A. I think you'd have to be a little bit more specific.
Q. Well, do you consider his work to be authoritative in the field in which it's published?
A. I consider his work to be very good.
Q. Okay. Do you think that his peers in his profession consider him to be an authority in his field?
A. Yes, I think that's fair.
Q. And in fact, you have relied on him yourself in constructing your models, correct?
A. Yes.
Q. How about Professor Mayer? Wait. Let me withdraw that question.
On Professor Jackman, do you consider him to be experienced?
A. Yes.
Q. He's in your view qualified to render opinions on legislative redistricting matters; is that correct?
MR. KEENAN: Object to the question as vague.
THE WITNESS: I don't know his work on

Page 24

specifically legislative redistricting very well. I don't know that he has published recently specifically on legislative redistricting. I think that he is generally a very qualified political scientist.
BY MR. EARLE:
Q. Have you read Jackman's article with Richard Niemi, is it, on legislative redistricting?
A. Yes, I'm fairly certain that I have. If I recall correctly this is an article from at least 20 years ago. I don't know if I could specifically characterize anything in the article off the top of my head.
Q. All right. Let's turn to Professor Mayer. Are you familiar with Professor Ken Mayer's work?
A. Only vaguely. It is my -- prior to this -- reading his report in this case, it was my impression that most of his work dealt with institutions and especially executive institutions as opposed to legislative elections so I would say I was much less aware of his work than -- sorry.
Q. No. Go ahead. Finish. I did not mean to --
A. I would say than other scholars who deal more closely in the fields that I study.
Q. Okay. Do you consider Professor Mayer to be

Page 25

experienced in the political science field of elections?
A. Yes, only in the sense that I am aware that he has worked in this area for a very long -- for a relatively long time and published several articles related to elections.
Q. Do you consider him qualified?
A. Yes.
Q. So in your view qualified and experienced to render opinions in this case?
MR. KEENAN: Objection to the relevance and calling for a legal conclusion.
THE WITNESS: In a casual sense, yes.
BY MR. EARLE:
Q. Okay. Occasionally during the course of the deposition, counsel is going to interpose objections, and those are for the record. They have nothing to do with what's going on between you and me. I get to ask you questions, and you get to answer them, and he's making a record --
A. Okay.
Q. -- for subsequent use. And so it has no bearing on your answer to the question. You understand that?
A. So I should always answer the question even if there is an objection?

Page 26

Q. Unless he instructs you not to.
A. Okay.
Q. Okay. If he does instruct you not to, I'll ask him why.
MR. KEENAN: That deals with issues of attorney-client privilege and work product and things, but with just phrasing of questions, there won't be an instruction not to answer.
THE WITNESS: Okay.
BY MR. EARLE:
Q. Yeah. And so you understand that. Just so it's -- because you've never been in a deposition before, right?
A. Right.
Q. You've never taken a deposition?
A. No.
Q. Okay. Are you nervous?
A. Slightly.
Q. Okay. And why do you think you're nervous?
A. It's an unfamiliar situation.
Q. Uh-huh. Could it have anything to do with your lack of experience?
MR. KEENAN: Object as vague. Experience with what?
THE WITNESS: I think it would definitely have

Page 27

to do with my lack of experience in testifying in depositions, yes.
BY MR. EARLE:
Q. How about your lack of experience at being an expert?
A. As it would relate to my lack of experience in testifying at depositions, yes.
Q. Okay. Well, how about as your lack of experience as being an expert and rendering opinions for consideration by a court?
A. I don't think in general I'm uncomfortable at rendering opinions. I think -- I think that being in an official court circumstance when someone is inexperienced in that circumstance would likely make people nervous in general.
Q. I don't want to belabor your -- your CV too much, I mean at this point, but I guess I just want to be able to -- to have nailed down in this record here the extent of your experience. And as I look at your resumT and your background, it seems to me that you're -- you're kind of new. I think it would be fair to call it -- characterize you as -- as an inexperienced expert. Do you think that's right?
A. I have never served as an expert witness in a case so in that sense I am inexperienced.

Page 28

Q. And you're relatively inexperienced as a scholar; isn't that true?
A. Relative to what?
Q. Relative to somebody like Simon Jackman.
A. Yes, Simon Jackman is a more experienced scholar than I am.
Q. Same is true with Ken Mayer, correct?
A. I suppose that would be accurate.
Q. Okay. We'll move off of your -- your resumT for now.
If you take the body of your work in political science related to elections, is it fair to say that you've mostly concentrated on congressional elections and not state legislative elections?
A. Yes.
Q. There's a different dynamic between the two, isn't there?
MR. KEENAN: Object.
THE WITNESS: That's rather vague.
BY MR. EARLE:
Q. You beat counsel to the -- to the objection. Your dissertation was on congressional redistricting, correct?
A. Yes.
Q. And your published work has all been focused on

Page 29

congressional redistricting, correct?
A. My published work related to redistricting has focused on congressional redistricting.
Q. Okay. That's a good example of a clarifying answer to -- to a question, a precise answer. That's good. All right.
And so we can also say that none of your published work has focused on legislative redistricting at a state level?
A. Certainly I think there would be applications to state legislative redistricting in -- in my work. To the extent that I have relied on empirical data in my work, it has all come from congressional data.
Q. All right. So one of the things I would like you to try to do is answer the questions I ask, and as opposed to advocating in a nuanced way in -- instead of answering the question I asked.
MR. EARLE: Can you repeat the question I asked to the deponent, please?
(Question read: And so we can also say that none of your published work has focused on legislative redistricting at a state level?)
THE WITNESS: I think it's fair to say that none of my published work has focused on legislative redistricting. I think that's a complete statement.

Page 30

BY MR. EARLE:
Q. So the answer is yes? It's fair to say that, right?
A. If by focus you mean was the primary subject matter of any of my published work legislative redistrict -- state legislative redistricting, the answer is yes.
Q. Okay. All right. So and just to understand some -- some of the concepts here, a state legislative redistricting plan has component parts, right? Individual districts, right?
A. Yes.
Q. And where -- while you're looking at congressional redistricting at a national level, there's no national congressional redistricting plan, is there?
A. No.
Q. So the two are not equivalent in that regard, correct?
MR. KEENAN: Object as vague.
THE WITNESS: When states draw congressional maps they also have districts.
BY MR. EARLE:
Q. Excuse me?
A. When states draw congressional maps, of course they also have districts just like you were characterizing state legislative maps.

Page 31

Q. But there are 50 of those, aren't there?
A. There are 50 states that draw congressional maps that all feed into the U.S. Congress, yes.
Q. There is not a single United States congressional redistricting plan?
A. True.
Q. There are 50 congressional redistricting plans?
A. Yes.
Q. And to be precise, we have to exclude those states that have a single congressman, correct?
A. Sure.
Q. Okay. Let's get some other basic definitions down as we go forward here. Because we're going to be talking about stuff, but I want to make sure that we're always on -- using the same language. All right?
You would agree that partisan gerrymandering exists, correct?
MR. KEENAN: Object as vague as to what "partisan gerrymandering" is.
THE WITNESS: I don't feel like I can answer the question unless you give a more precise definition of partisan gerrymandering.
BY MR. EARLE:
Q. You would agree that partisan gerrymandering is the drawing of legislative district lines to subordinate adherents of one political party and entrench the rival party in power, correct?
MR. KEENAN: Object to the extent it calls for a legal conclusion, but you can answer your understanding.
THE WITNESS: I believe that there are maps drawn with that intent.
BY MR. EARLE:
Q. Okay. Wisconsin is one of those maps that was drawn with that intent?
A. You're referring to state legislative map in Wisconsin?
Q. Yeah. Uh-huh.
A. My only knowledge of Wisconsin is what I had read in the complaint so my only knowledge of what the intent was would be as it was characterized by the plaintiffs in their complaint.
Q. Okay. All right. But just so let's just nail down this, the definition. Is it accurate to say that partisan gerrymandering is the drawing of legislative district lines to subordinate the adherents of one political party and to entrench the rival party in power?
A. That is not how I define partisan gerrymandering in

Page 32

Page 33

my own work. So I don't know that I would agree with that.
Q. So you think that the -- that the author of that definition is ill informed or wrong?
A. I think the term is vague. I think many people have different definitions of what they mean by the term so no, I wouldn't say that the particular definition that I use is more authoritative than what other people might use.
The way that I use it in my work is somewhat different and does not rely on intent. And it does not rely on empirical results of elections. I'm just looking at the process.
Q. Okay. Could you explain to me what is wrong with that definition?
MR. KEENAN: Which definition?
MR. EARLE: The definition I just provided to the deponent.
BY MR. EARLE:
Q. Partisan gerrymandering is the drawing of legislative district lines to subordinate adherents of one political party and entrench a rival party in power.
A. I think how that you define a term like that is going to depend on the context in which you're --

Page 34

you're using it. That term may be appropriate in a context that's different than the way that I am using it in my own work. So I would not characterize it as wrong so much as inappropriate for how I am analyzing gerrymandering in my own work.
Q. Do you consider that definition I just gave you to be irrelevant to this case as you understand this case?
A. As I understand this case, the plaintiffs are arguing -- as I understand this case, there -- the use of partisan gerrymandering in the context would essentially be a legal conclusion. I don't have any opinion on whether that definition is appropriate in this case.
Q. Do you know who the author of that opinion is, I mean that definition is that I just gave you?
A. I believe it comes from a Supreme Court opinion. Whether it is -- because the quote is familiar to me. Whether it comes from Bandemer or one of the later cases, I can't recall off the top of my head.
Q. It's Justice Ginsburg in the Arizona case.
A. Oh, okay.
Q. Okay. Now, Justice Ginsburg in that decision also said that partisan gerrymanders are incompatible

Page 35

with democratic principles. You agree with that statement, right?
A. Not in the way that I define partisan gerrymandering in my own work.
Q. Okay. So you don't think that partisan gerrymandering is incompatible with democratic principles?
A. The statement is very vague, both with respect to partisan gerrymandering and democratic principles.
Q. Well, we've just defined partisan gerrymandering, how Justice Ginsburg from the Arizona case. So what's ambiguous about democratic principles?
A. It sounds like you're asking for something which is very -- it sounds like you're asking for a personal opinion outside of the subject that I have been recruited to ask as an expert on.
Q. Okay. So can you identify what the democratic principles are that are injured by a successful partisan gerrymander?
MR. KEENAN: Object as vague.
THE WITNESS: Given the way that I think about partisan gerrymandering, I would not know what a successful partisan gerrymander was.
BY MR. EARLE:
Q. Okay. Why do you say that? Well, let me withdraw

Page 36

that question and rephrase my question.
Is it your opinion that there is no such thing as a partisan -- a successful partisan gerrymander?
A. In the way that I define partisan gerrymandering in my work, that would not be a meaningful statement because I define partisan gerrymandering as something related to the process of gerrymandering.
Now, I -- in a casual sense you do observe some partisan gerrymanders winning more seats for the gerrymandering party than others, so if you are relating partisan gerrymandering and the definition to the intent to -- again I don't remember the exact quote that you used. Some partisan gerrymanders are more successful than others, I suppose, but I'm using the term here very casually, and I don't -- in neither the way I would define it in my work nor the way I would expect a court to define it, even though I'm not -- not offering it as an opinion on how I would expect a court to define it.
Q. You're not offering an opinion as to how you would expect the court to define partisan gerrymandering?
A. Right.
Q. Okay. And you will not be doing that at trial?
A. I will be doing that at trial.
Q. Okay.

Page 37

A. I don't think that the way that I would characterize partisan gerrymandering would be compatible with the -- I'm sorry, I will be doing that at trial.
If you can go back to the previous question, you can refresh my memory as to what you're asking. I have forgotten it.
Q. Why don't we go back and refresh the deponent's recollection of the preceding question before that one. If you can read the question and then his answer, and then you can elaborate if you wish.
A. I'm sorry, the previous question was to how to expect a court to define partisan gerrymandering?
MR. KEENAN: She'll read it back.
BY MR. EARLE:
Q. Just so you're clear, I'm not trying to play got you with you, so I'm going to have the court reporter read the question -- the first question that you gave an answer to, and then my follow-up question that you struggled with answering, okay? So that --
(Question and answer read: Is it your opinion that there is no such thing as a partisan -- a successful partisan gerrymander?
Answer: In the way that I define partisan gerrymandering in my work, that would not be a meaningful statement because I define partisan

Page 38

gerrymandering as something related to the process of gerrymandering.)
MR. EARLE: It's the question before that.
MR. STEPHANOPOULOS: The one after.
THE COURT REPORTER: The one after is --
(Question and answer read: You're not offering an opinion as to how you would expect the court to define partisan gerrymandering?
Answer: Right.
Question: Okay. And you will not be doing that at trial?
Answer: I will be doing that at trial.)
(Discussion held off the record.)
MR. EARLE: He wants to amend his answer.
THE WITNESS: I am not offering an opinion on how I would expect a court to define partisan gerrymandering because I am not offering an opinion about what I think judges will do. I am offering an opinion on how the court should define partisan gerrymandering.
BY MR. EARLE:
Q. And what is your opinion -- is that opinion stated in your report?
A. I don't think it is directly stated in my report. And to the extent that it's -- okay. Sorry. To the

Page 39

extent that it is not stated in my report, I don't know that I expect to offer that particular opinion. I don't know exactly what I would be asked at a trial or something like that if that's what you're asking.
Q. So we're pretty much all over the map on this here because you've started by saying that you weren't going to do one thing, but then you were going to do that thing, and then you had a different version of that thing that applied to your work and that you're not sure how the court would do it so we've kind of like gone all over the place on this. So let's just go straight to the question.
A. Can I --
Q. Okay.
A. Okay.
Q. Exactly what is your definition of partisan gerrymandering?
A. The definition of partisan gerrymandering I use in my work is it would be a redistricting plan which is done under the complete control of one party. So typically where one party has control of the process of districting, and typically that would mean they have control over both houses of the state legislature and the governorship depending on how

Page 40

the process for gerrymandering works in a particular case or a particular state.
Q. And you define that as a partisan gerrymandering because one partisan party control the process?
A. Yes, in a formal way.
Q. Okay. Now, how would you define intentional partisan gerrymandering?
A. I would not define that term. I don't think it's a meaningful term in the context of my work.
Q. What about in the context of what happened in Wisconsin with Act 43?
A. Can you be more specific what you're asking?
Q. How would you define intentional partisan gerrymandering?
A. I would not define intentional partisan gerrymandering. I don't think it's a meaningful -- I -- in the context of my work.
Q. I'm asking you in the context of Act 43, how would you define intentional partisan gerrymandering?
MR. KEENAN: Just object as vague. He says it doesn't make any sense. He's asked and answered this like twice now.
MR. EARLE: Could you read the question to the witness?
(Question read: I'm asking you in the context

Page 41

of Act 43, how would you define intentional partisan gerrymandering?)
THE WITNESS: The question is not meaningful in a way that I can answer it.
BY MR. EARLE:
Q. Well, do you think it's a relevant question in the context of this case in which you've been hired to render opinions?
You're -- it looks like you're about ready to answer the question. Just so the record is clear, this is a transcript, and it's not time coded.
So --
A. That's fine.
Q. You've sat silently for quite some time, and you appear to be thinking, and I don't want to interfere with that. I just want the record to reflect that there has been the passage of time between the statement of the question and -- and the answer.
Take your time.
A. It sounds like the question is asking for a legal conclusion related to intent, which I don't think I am -- is related to what I have been recruited to act as an expert on.
Q. Okay. So is it correct to say that under your definition of Wisconsin's current -- under your

Page 42

definition, Wisconsin's current plan, Act 43, is a partisan gerrymander, correct?
A. **In the context of how I code partisan gerrymandering in my work, I would code it as a partisan gerrymander, yes.**
Q. That's because one party had complete control over the entire process, correct?
A. **As I understand the legislative control in Wisconsin, yes.**
Q. Is it correct that your definition does not take into account the electoral impact of a plan?
A. **My work studies the electoral impact of a plan. It studies the impact of partisan gerrymanders. It does not take into account their impact, whether I define them as partisan gerrymanders or not.**
Q. You do not connect the outcome of a plan to the intent of the plan, correct?
A. **I do not connect the outcome of the plan to whether I code it as a partisan gerrymander or not.**
Q. And you're not going to be rendering any opinion as to whether the impact of Act 43 was the intentional result of the design of Act 43, correct?
A. **I am not rendering an opinion on the specific intent of anyone who was crafting Act 43.**
MR. EARLE: Okay. Could you read the question

Page 43

to the deponent again?
(Question read: And you're not going to be rendering any opinion as to whether the impact of Act 43 was the intentional result of the design of Act 43, correct?)
THE WITNESS: Certainly I believe that the impact of a map is the result of intentional acts by the people who were drawing the map in addition to several other variables. I believe there is intent behind the drawing of legislative maps, and I'm sure that's true in this case as well.
BY MR. EARLE:
Q. I need you to answer the question I asked you, though.
A. **Okay.**
MR. EARLE: Read it again. And on the transcript each time could we have you re-print the question in parentheses?
(Question read: And you're not going to be rendering any opinion as to whether the impact of Act 43 was the intentional result of the design of Act 43, correct?)
THE WITNESS: I believe that the impact of any legislative map is in some way the result of the intent behind that map. The impact may be different

Page 44

than what the intent was, but it is still in some way the result of that intent, combined with other variables.
BY MR. EARLE:
Q. I'm not asking you about your beliefs. I'm asking whether you're going to be rendering an opinion very specifically and I'll ask the question be reread again. And listen very carefully to the question and answer the question that I asked. Okay?
(Question read: And you're not going to be rendering any opinion as to whether the impact of Act 43 was the intentional result of the design of Act 43, correct?)
THE WITNESS: I will not be rendering an opinion on the intent behind Act 43. I will be -- most of the opinions that I am giving in this case relate to the impact of adopting the standard for what would constitute unconstitutional partisan gerrymander as presented in the plaintiffs' complaint. That would also relate to Act 43 and the specific facts presented in this case.
BY MR. EARLE:
Q. We're going to move on.
Would you characterize your coding of partisan gerrymanders as idiosyncratic?

Page 45

A. **No, there are certainly cases in which there is a question as to how something could be coded and it might recall -- require a judgment call in certain specific cases.**
Q. Can you point to any legal or political science literature that codes plans in the same way that you did?
A. **That codes all of the plans in the exact same way that I did? There are -- there are -- is other literature that codes plans in a similar way that I would and for the most part, yes, relies on the same sort of standards and judgments that I use.**
Q. Can you identify those, please?
A. **There's an article by Michael McDonald in 2004. I don't know the title off the top of my head, but it certainly codes congressional plans in a similar way, and in part I have relied on that.**
There is an article by Squire from the early 1980s that codes plans from the 1970s I believe in a similar way. Again I am not recalling the titles off the top of my head. I could look them up if that's necessary.
Q. Okay. Are you familiar with Andrew Gelman and Gary King's measure of partisan symmetry?
A. **Yes.**

Q. Can you define it?
A. For a given share of the vote -- the definition would be that both parties win the same share of seats given a certain percentage of the vote if that party were to receive them. So if the democrats were to receive 55 percent of the vote, they would receive the same share of seats that the republicans would if the republicans received 55 percent of the vote.
Q. How do Gelman and King determine what the outcome of a hypothetical tied election would be?
A. My impression is that generally they -- I mean it's a little more subtle than this, but they would use a uniform swing across districts based on whatever underlying data they're using for -- so -- so they would take the deviation of the tied from whatever baseline they're using and use a uniform swing across districts to determine what the vote would be in those districts.
Q. How is that more subtle than this?
A. Well, Gelman and King's work has the potential to incorporate many other variables.
Q. Can you explain how their measure differs from the efficiency gap?
A. The efficiency gap defines a fair map under a

specific -- the efficiency gap is more specific in how it defines a fair map in that the efficiency gap prescribes a specific slope of responsiveness which partisan symmetry does not do. That is what I would say would be the most relevant difference between efficiency gap and partisan symmetry.
Q. Have you ever calculated their measure?
A. In what context?
Q. The context of your work, any legislative plan.
MR. KEENAN: Object to vague as "their measure."
THE WITNESS: Oh, sorry, which measure are you talking about?
BY MR. EARLE:
Q. I'll withdraw that question and rephrase it.
Are you familiar with the work of Roland Fryer and Richard Holden?
A. I don't believe so.
Q. Do you know what -- you're not familiar with their work on simulating districts plans?
A. Not in any specific sense.
Q. Are you familiar with the work of Adam Cox, John Friedman, and Richard Holden on how to construct optimal -- an optimal gerrymander?
A. Not in any specific way.

Q. Have you read Optimal Gerrymandering: Sometimes Pack but Never Crack?
A. Can you -- do you know who the authors of that are?
Q. Yeah. I said are you familiar with --
A. Oh, this is the -- sorry.
Q. -- Adam Cox, John Friedman, and Richard Holden?
A. I am not familiar with that.
MR. EARLE: Let's take a break.
(Break taken 9:47 to 9:55 a.m.)
BY MR. EARLE:
Q. Have you heard of global Moran's I?
A. No.
Q. How about local Moran's I?
A. No.
Q. How about the isolation index?
A. I don't think so.
Q. How about the index of dissimilarity?
A. I don't think so.
Q. Okay. Have you written anything about clustering analysis?
A. No.
Q. Have you ever produced simulated plans like Chen and Rodden? C-H-E-N, R-O-D-D-E-N.
A. No.
Q. I might have asked you this, but did you -- have you

read Simon Jackman's textbook?
A. I don't believe so. This is -- this is -- this is a methodology textbook?
Q. I think that's a fair description.
A. I'm not sure. I might -- I have not -- I'm not sure.
Q. Okay. Are you familiar with any of the authoritative textbooks on qualitative methodology?
A. I'm sorry, on qualitative methodology?
Q. Uh-huh.
A. No, I don't --
Q. Quantitative. I'm sorry, my eyes are -- quantitative methodology.
A. Am I aware of any authoritative textbooks on quantitative methodology? Certainly I have taken several classes in quantitative methodology in which we relied on textbooks. I don't know that I would say that one is particularly more authoritative than any others.
Q. Okay. Just a loose end here. Can you identify any -- any at all, any measures of the geographic clustering of different groups?
MR. KEENAN: I'm going to object as vague.
THE WITNESS: Not specifically.
BY MR. EARLE:

Page 50

Q. I just want to mark the -- that article by Fryer and Roland and Holden, Roland Gerhard and Holden that I asked you about earlier and give you a chance to take a look at it.
(Exhibit No. 18 marked for identification.)
BY MR. EARLE:
Q. Showing you what's been marked as Exhibit No. 18. Take a moment.
A. (Witness reading.)
MR. KEENAN: So you're saying you mentioned this article earlier?
MR. EARLE: Yeah.
THE WITNESS: Is there something specifically you want me to look at in this article?
BY MR. EARLE:
Q. No. Looking at the article, does this trigger any memory?
A. The article is not familiar to me.
Q. Okay. Are you familiar with any of the authors?
A. No.
Q. Okay. Now, you said you read the -- the complaint. We have a -- a copy of the complaint. Has it been marked yet?
(Exhibit No. 19 marked for identification.)
Q. Showing you what's been marked as Exhibit No. 19.

Page 51

Take a moment to look at it. Is this the complaint that you're familiar with?
A. I believe so.
Q. How many times have you read this complaint?
A. Three or four.
Q. Fair to say you've studied it carefully?
A. I've studied parts of it carefully.
Q. What parts did you not study carefully?
A. Parts related to standing.
Q. When you say "standing," are you referring to the paragraphs describing the -- the parties?
A. Paragraphs describing where particular plaintiffs lived. I have also not -- not studied carefully the discussion of specific -- specific division of counties or areas in Wisconsin in the particular districts.
Q. I don't know what that meant, what you just said. You have not studied specifically what?
A. I have not reread carefully the parts of the complaint that deal with how particular counties or particular areas in Wisconsin were divided into specific districts. I have read them, but I have not reread them several times.
Q. Okay. After having -- well, just so I'm clear, based on what you said earlier in your testimony, is

Page 52

it fair to say that you have no opinion relevant to paragraphs 1 through 11?
A. I don't think that's fair to say.
Q. Okay. Which paragraphs between paragraphs 1 and 11 do you have an opinion that -- that is relevant to one of those paragraphs? That's poorly worded. Let me --
Which of the paragraphs do you have opinions about? Let me rephrase that.
Which of the paragraphs between paragraphs 1 and 11 would you offer an opinion contradicting the content of those paragraphs?
A. I would offer an opinion contrary certainly in paragraph 6.
Q. Let's read that paragraph into the record. "When the efficiency gap is relatively small and roughly equivalent to the efficiency gaps that have traditionally existed, the map should not be deemed unconstitutional."
Do you quibble with that, that first sentence?
A. As a stand-alone sentence?
Q. Yes. Yes.
A. Well, I believe that there are many reasons why a map might be declared unconstitutional which would be unrelated to an efficiency gap. So as a

Page 53

stand-alone sentence, I don't believe that that would be entirely accurate. I guess in the context of the rest of the report --
Q. This is a -- a complaint.
A. The complaint. Okay. I'm sorry. Yes. In the context of the rest of the complaint, I would not quibble with that for the -- that it should not be deemed unconstitutional for the reasons that the complaint is citing.
Q. So -- so that means you're okay with the first sentence of paragraph number 6 in the context it's offered?
A. Yes.
Q. Okay. What about the second sentence of paragraph 6. "In such cases there may be no intent to treat voters unequally; in any event, the effects of any gerrymandering are likely to be redressable through the political process."
Do you have the same reaction to that sentence?
A. In the first clause it's very general. I have no objection to the first clause. Let me see, "the effects of any gerrymandering are likely to be redressable through the political process." I don't know that that's particularly true just -- I don't know that a small efficiency gap as related to the

first -- as it relates to the first sentence would necessarily be related to whether a gerrymander is redressable to the political process.
Q. Rather than go through each of the remaining sentences, identify the sentences in the rest of this paragraph that you would take a contrary opinion to.
A. I would take a contrary opinion to the next sentence.
Q. Would you read that sentence?
A. "But where the efficiency gap is large and much greater than the historical norm, there should be a presumption of unconstitutionality."
Q. Okay. Anything else in the paragraph you take exception to?
A. The next sentence.
Q. Read that one.
A. "In such a case, an intent to systematically disadvantage voters based on political beliefs can be inferred from the severity of the gerrymander alone."
Q. Okay. Is there anything else in the paragraph that you disagree with?
A. I think I would probably disagree with the next sentence.

Q. Read that one into --
A. "And because such severe gerrymanders are likely to be extremely durable as well, it is unlikely that the disadvantaged party's adherents will be able to protect themselves through the political process."
Q. Okay. What about that sentence do you disagree with?
A. That particular sentence?
Q. Yeah. Uh-huh.
A. I would disagree with the notion that if you are defining a severe gerrymander as having a large efficiency gap in a particular instance, that that gerrymander is likely to be extremely durable.
Q. Do you agree with the -- with the statement that any severe gerrymander that is, in fact, extremely durable makes it unlikely that the disadvantaged party's adherents will be able to protect themselves through the political process?
A. I don't think I would agree with that. No, I would not agree with that.
Q. So just so I am clear here, if you have a severe gerrymander that skews the electoral districts in a way that substantively disadvantages the adherents of one party, and the gerrymander is as a matter of fact durable enough to last the entire decennial

cycle with its disadvantaging effects -- all right, those circumstances are givens that I want you to assume the existence of. All right. You would agree that under such a circumstance, recourse to the political process becomes unavailable to the adherents of the disadvantaged party?
MR. KEENAN: Object as vague.
THE WITNESS: No, I wouldn't say I agree with that. I think there are other possible recourses.
BY MR. EARLE:
Q. Other than filing a lawsuit like this one, what recourses do those adherents have?
A. Well, assuming that the gerrymander was done through the normal process of typical legislation in the state, the adherents could, for instance, elect a governor of their party, and that governor in subsequent redistricting -- redistricting cycles could have some power over how the lines are drawn. That would be one recourse.
Q. That -- by definition that's a recourse that would only exist in the last election of the decennial cycle; isn't that true?
A. I'm sorry, so you're asking that if a gerrymander is extremely durable --
Q. Right. These are -- these are the -- the givens

that I want you to assume for this hypothetical. Okay? The gerrymander is severe so that all the adherents of one party are either packed into a few heavily populated districts or cracked and spread out amongst the remaining districts such that they cannot obtain elect -- win an election. All right? And that structure is durable enough to last the entire decennial cycle. I want you to assume those givens.
A. If you are stipulating that an election system is set up such that it is impossible for a party to win representation, then I agree it is impossible for the party to win representation.
Q. Okay.
A. So I think that would depend on your definition of durable. It sounds like you're defining durable to say that it is impossible for the -- the out party to win representation, in which case your question I think is tautological.
Q. Well, you just made the question tautological. I don't think I intended it that way. I think that if you have a -- a skew that is intentionally imposed on the adherents of one party that's adverse to them, and that skew is substantive through cracking and packing, and it is severe enough that it is

Page 58

durable for the entire decennial cycle, the disadvantaged adherents would have no political recourse, correct?
MR. KEENAN: Objection. Asked and answered.
THE WITNESS: You're stipulating that it is durable enough that the opposing party adherents could not elect members of their party. With that stipulation, I think your question is tautological in that the opposing party's adherents could not win representation. I'm not sure what else you're asking.
BY MR. EARLE:
Q. Any other paragraphs between -- how about -- that you disagree with?
A. (Witness reading.) I would disagree with paragraph 7.
Q. Okay. How about 8? Do you disagree with paragraph 8?
A. Given that my -- so given that my only knowledge of the specific process for the -- an enactment of the current plan was what I read in this complaint, I don't know that I have enough information to agree or disagree with paragraph 8.
Q. Okay. Now, did you ask to see the documents related to paragraph 8?

Page 59

A. I did not.
Q. How about paragraph 9? Just before we go on to 9, you're not going to be offering any testimony that's contrary to paragraph 8, correct?
A. I will not be offering testimony contrary to paragraph 8. Paragraph 9 is a little bit vague.
Q. Before you go further with paragraph 9, on paragraph 8, you would agree that the content of that paragraph means that the -- that Act 43 is a partisan gerrymander under your definition, right?
A. I don't think the content of paragraph 8 informs my opinion about whether the Act 43 was a partisan gerrymander. I -- for instance, whether it was the result of the ordinary political process or not does not inform my opinion about whether it was a partisan gerrymander under the ordinary political process of a normal bill drafted and enacted by a republican-controlled legislature and signed by a republican governor, I believe that would be a partisan gerrymander as well in the way that I code partisan gerrymanders in my work.
Q. So the content of paragraph 8 means that as far as you're concerned, Act 43 was a partisan gerrymander? It says here in paragraph 8 it was a -- it was drawn up in secret by the republican leadership without

Page 60

consultation of the democrats with the purpose and intent of altering a map that was already favorable to them, and the proposal was passed through, ran through the assembly without any opportunity for real debate. Correct? That means it satisfies all your criteria, right?
MR. KEENAN: Object.
BY MR. EARLE:
Q. For coding purposes.
A. My criteria for what I would code as a partisan gerrymander is that the process -- the normal political process was controlled by one party.
Q. Let's go on to 9.
A. Okay.
Q. We have a quibble with 9?
A. Because this is the introductory part of the complaint, I assume that many of these terms are further defined in the body of the complaint. So I don't -- whether I would have a quibble with 9 would depend on how they are defined.
Q. Which terms?
A. Well, outlier.
Q. What does an outlier mean?
A. Well, an outlier would be a data point which is -- which is very far to one extreme of the rest of the

Page 61

data set that would compile the data that data point is a part of. It would be -- it would be a data point that is not part of the -- of the rest of whatever the distribution of that data.
Q. And you know what partisan symmetry is, right?
A. I believe there can be a number of definitions for partisan symmetry. You have asked me about one, right?
Q. Uh-huh.
A. So I am assuming that that is the definition that they are using. It is not clear to me from this paragraph of the complaint that that is the definition that they are using.
Q. Okay. And what would you assume that definition to be?
A. I would assume that that definition was that each party will win an equal number of seats given a particular share of the vote, and that both parties will win the same number of seats if they receive the same -- that particular share of the vote.
Q. Anything else on paragraph 9?
A. Well, I do see that it is -- sorry, the second sentence of paragraph 9 is defining partisan bias as only the share of seats that each party would win if they tied statewide each receiving 50 percent of the

vote.
Subsequent to that, it's giving I presume an estimate for what percent of the vote each party would win in a 50/50 election. Again it doesn't specify it in this paragraph. I assume that's using a uniform swing across districts. If that is true, there is nothing specific in the data here that I would object to.
Q. Okay. Let's go to paragraph 10.
A. **(Witness reading.) I would certainly disagree with the last sentence as implied by the previous sentences.**
Q. Okay. Read what -- what you're referring to.
A. **The last sentence is, "Thus, defendants cannot salvage the current plan on the theory that adherence to redistricting criteria or the state's underlying political geography made an unfair plan unavoidable."**
Q. What's your quibble with that sentence?
A. **My quibble with that sentence is that the fact that a single plan can be drawn that would display different characteristics under measures like partisan bias or efficiency gap under a particular election result, the fact that a single plan can be drawn that would display those characteristics would**

imply that the state's underlying political geography would not contribute to how -- I mean this is not exactly what the sentence is stating, but that a state's underlying political geography would not contribute to how a typical plan would be drawn up or how one might expect a plan to be drawn up, even absent specific partisan control.
Q. Let me see if we can -- well, let's finish paragraph 11 then, and then we'll go back on some of this stuff here. Do you quibble with paragraph 11?
A. **I don't know what a neutral plan would be. I mean this relates to the plaintiffs' intent -- sorry.**
Q. All right. Let -- let's get -- let's nail down what your understanding of the proposed test that the plaintiffs have in this case is.
A. **Sure.**
Q. What is it?
A. **My understanding from the complaint of the plaintiffs' test is that they would propose that you would measure the efficiency gap in an election result in the first election following a redistricting cycle. If that cleared a certain threshold, and I believe that the complaint suggests that threshold should be seven percent, that the plan should be presumptively unconstitutional.**

Once a plan is deemed presumptively unconstitutional, that the defendants could offer evidence that some other factor should redeem it and make it constitutional instead.
Q. What is the first step of the plaintiffs' proposed test?
A. **My impression is that the first step of the plaintiffs' proposed test differs in the complaint from other documents that I have read that the plaintiffs have filed in the case. So it is unclear to me what the plaintiffs' first step in the proposed test is.**
Q. Okay. You don't understand the first step of the test to be a showing that the plan was adopted with the express intent to subordinate the opposing party --
A. **That --**
Q. -- through a process of cracking and packing?
A. **That first step is not clear to me from the complaint.**
Q. Okay. Is it clear to you from subsequent filings in the case that that is the first step of the plaintiffs' test?
A. **I believe that subsequent filings from the plaintiffs claim that they would use as a first step**

some sort of subjective measure of partisan intent or evaluation of partisan intent. Again that's not clear from the complaint so I -- given that the -- the various documents are contradictory, it is not clear what the plaintiffs' test to me is.
Q. Okay. Now, you used the word "contradictory." Contradictory means the documents take non-reconcilable positions, right? What -- where is anything in this complaint contradictory to any other document you've seen filed in this case?
A. **I believe so let me find it. So paragraph 84 of the complaint, "The same two-part approach should be applied to partisan gerrymandering claims, only with the efficiency gap substituted for total population deviation. The first step in the analysis is whether a plan's efficiency gap exceeds a numerical threshold."**
Q. Why don't you read paragraph 89.
A. **"Finally, there is no doubt that the current plan was specifically intended and indeed designed to benefit republican candidates, and to disadvantage democratic candidates, to the greatest possible extent. Thus, the current plan had both the purpose and effect of subordinating the adherents of one political party and entrenching a rival party in**

Page 66

power, in violation of their right to equal protection under the law."
Q. You understand and -- would you read paragraph 31?
A. I should mention I think there is another part of the complaint that I would want to highlight, but I am having a little bit of trouble finding it.
Paragraph 31. "The current plan was drafted and enacted with the specific intent to maximize the electoral advantage of republicans and harm democrats to the great possible extent, by packing and cracking democratic voters and thus wasting as many democratic votes as possible. Indeed, after a trial in prior litigation, a three-judge court characterized claims by the current plan's drafters that they had not been influenced by partisan factors as 'almost laughable' and concluded that 'partisan motivation' clearly lay behind Act 43."
Q. Now, did you go to that citation?
A. The citation to Baldus?
Q. Yeah.
A. I did not go to that citation.
Q. Do you question the content of paragraph 31?
A. I don't question the content. I just don't understand how it relates to what you were asking me previously.

Page 67

Q. Well, okay. I want you to assume that the plaintiffs' test has three parts: First, a showing of intent to discriminate on the basis of partisanship. All right? Second, a showing of effects as measured by the efficiency gap. And third, an opportunity for the defendants to make a showing that the plan was the result of legitimate public purpose or public policy or geography. All right? Is that familiar to you?
A. Previously you asked me whether anything in the complaint was contradictory to anything in the later filings.
Q. Uh-huh.
A. The test as expressed by the plaintiffs in the complaint is contradictory to what you just said.
Q. How is that so?
A. The test as explicitly laid out in the complaint has two steps. It does not include the first step.
Q. How would you characterize paragraphs 31 through 41? And 43. I'm sorry.
A. I would characterize them as providing factual background. I would not certainly -- I would certainly not characterize them as in any way expressing a legal test that would be integrated into part of the express two-part approach as it's

Page 68

explained in the complaint.
Q. Has anybody instructed you to assume a two-part test as opposed to a three-part test?
A. No one has instructed me to assume that.
Q. You arrived at this conclusion yourself by reading the complaint?
A. This is what the complaint states. It repeatedly states a two-part test.
Q. Okay.
A. I also think that -- am I allowed to refer to my own notes with respect to this complaint?
Q. If you show them to me. You have notes? Were they produced in response to the subpoena?
A. I have some handwritten things that are highlights that I put on the complaint.
Q. Okay. And you want to use them now? Okay. Let me look at them.
MR. KEENAN: Do you think you need to use them?
THE WITNESS: Okay. As long as I have --
BY MR. EARLE:
Q. If it would make your testimony more efficient, you can.
A. I can reread the complaint.
Q. Well --
A. Find --

Page 69

Q. -- not on my clock you can't. I have seven hours with you, and if you have notes that are going to make it faster, you can go ahead and look at those notes.
MR. KEENAN: You don't have to do it.
BY MR. EARLE:
Q. You can show them to me first before you -- before you use them.
A. I believe there is a statement in the plaintiffs' filings that over a certain threshold of efficiency gap partisan intent can be assumed.
Q. Okay.
A. That's what I'm trying to look for.
MR. KEENAN: Which we just saw in paragraph --
MR. EARLE: Well -- okay. Let's just move on.
MR. KEENAN: Paragraph 6.
THE WITNESS: It's paragraph 6?
BY MR. EARLE:
Q. I want to go into another -- another area now of questioning.
A. It's -- sorry. That's not actually the part that I'm referring to. But I --
Q. So you want to --
A. I think it -- go ahead.
Q. Okay. All right. What is the commonly accepted

Page 70

error rate in social sciences?
A. Error rate?
Q. Yeah.
A. Can you define "error rate"?
Q. You don't understand what error rate means?
A. If you're referring to a standard of statistical significance?
Q. Okay. Yeah.
A. Yes.
Q. Yeah.
A. I would say the most common threshold would be five percent.
Q. Okay. All right. I want to draw your attention to your quote on page 5 of your report.
MR. KEENAN: Exhibit 17.
BY MR. EARLE:
Q. It's the quote is at the bottom second to last sentence of the first full paragraph. "I concur that this shortcut is an appropriate and useful summary measure of efficiency gap and also use it in subsequent examples in this report."
Do you see that there?
A. Yes.
Q. Okay. You're referring to Jackman's report, correct?

Page 71

A. Yes.
Q. Okay. And you're referring to the methodology used by Jackman in calculating the efficiency gap?
A. I'm referring to a part of his methodology. Yes.
Q. Let's nail that down. Can we look at page 16 -- is Jackman's report in -- it's already been marked as an exhibit.
MS. GREENWOOD: It's 11.
Q. Okay. So we -- let me -- I'm going to show you what has been marked as Exhibit 11 in this case. On this exhibit it's marked Exhibit 3 because it's Exhibit 3 to the complaint. Okay. So assuming the reader of this transcript figures that out, we --
MR. KEENAN: There was an issue with the copy that has the exhibit sticker on it having all the pages. So that's why we're using this one.
MR. EARLE: Oh, there is?
MR. KEENAN: The court reporter I think scanned the one with the 11 sticker wrong so there's some missing pages. So that's why I said we could just use the one that's attached to the complaint because it's an identical document, just doesn't have the exhibit sticker on it.
MR. EARLE: He cites a correct version of Exhibit 11 in the record of these depositions?

Page 72

MR. KEENAN: Yes.
MR. EARLE: And we're going to -- I want to go to equation number one.
MR. KEENAN: Can I get a copy of --
MR. EARLE: Oh, sure.
BY MR. EARLE:
Q. I want to draw your attention to page 16?
MS. GREENWOOD: Page 16. Yep.
Q. Equation one. I'm sorry. Equation one, in paragraph 6.1, the efficiency gap when districts are of equal size. And the first sentence reads: Under the assumption of equally sized districts, McGhee, parens 2014 comma 80 re-expresses the efficiency gap as, and then there's a formula -- formula number one?
A. Yes.
Q. That's what you're referring to in that sentence from page 5 of your report, correct?
A. Yes.
Q. Okay. And you have yourself repeatedly calculated plans' biases by comparing the parties' actual seats to their expected seat shares given a responsiveness of two, correct?
A. Yes.
Q. And that is essentially identical to the efficiency gap, correct?

Page 73

A. Yes.
Q. Okay. Now, let's move over to your article. Gerrymandering or Geography -- well, two articles. Gerrymandering or Geography or Disappearing Biases?
A. Yes.
Q. You're familiar with those, right?
A. Yes.
(Exhibits Nos. 20 and 21 marked for identification.)
BY MR. EARLE:
Q. Now -- so -- okay. Do you think your models in these two articles are reliable?
A. You know, you haven't given me a copy of my articles.
Q. Oh, I'm sorry, your lawyer has them.
A. Okay. Can you repeat the question?
Q. Okay. Are your -- are your models reliable?
A. What do you mean by "reliable"?
Q. How would you -- I mean the term reliable has substantive meaning in your profession, doesn't it?
A. Yes, I believe that -- so what I am seeking to do in these articles is to characterize from an historical perspective how many seats a party would expect to win given a vote -- particular vote share.

Given that, I think that I have used a very simple model which could be made I believe slightly more accurate by increasing the complexity, but for the simplicity of the model that I am using, which I think is appropriate given the venue that I'm publishing, I believe that the model is reliable.
Q. Given the venue that you're publishing. What does that mean?
A. So the -- the journal Research & Politics is an open-access journal which I believe is trying to target, in addition to academics, other people who are interested in empirical political science research. Does that make sense?
Q. Yeah.
A. Okay.
Q. Do these models reflect modern political science techniques?
A. Yes.
Q. And you would trust their predictions for 2012 and 2014?
MR. KEENAN: Object as vague as to predictions.
THE WITNESS: I do not believe that these models are providing predictions.
BY MR. EARLE:
Q. Okay. Go to table 3 in each of those articles.

Yeah, the table 3, the regression results.
A. Yes.
Q. In the Disappearing Bias article.
A. I'm sorry, this is the --
MR. KEENAN: Which number?
THE WITNESS: Is this 21?
MR. STEPHANOPOULOS: Exhibit 21.
BY MR. EARLE:
Q. Twenty-one. Page 13.
A. Okay.
Q. You trust the predictions here?
A. Can I ask as an aside, do you know where you acquired this from?
Q. Website.
A. You acquired this from my website. Okay. So this is the current version. Because this a forthcoming article which I have very recently made some edits to before it's being --
Q. Why don't we do this. Let's take a very quick break, look at the article, make sure it's the latest version and make sure we're not operating off of a previously edited version.
(Break taken 10:39 a.m. to 10:43 a.m.)
THE WITNESS: So it appears that all of the data in this version of the article is identical to

what will be in the published version. I think there are some copy edits that I made to the text which wouldn't substantively alter anything in the article.
BY MR. EARLE:
Q. Good. So let's proceed. Okay. The design of -- of this regression exercise on table 3, it enables us to differentiate between the effects of the redistricting institution on bias and the effect of other demographic and political information, correct?
A. Right.
Q. Okay. This design also lets us make predictions about what a state's bias would be under hypothetical conditions, correct?
A. Well, I don't know if it would enable you to do that.
Q. Well, for example, we could predict what a state's bias would be if its map was a democratic gerrymander or a republican gerrymander or a partisan or court-drawn plan, correct?
A. It would give a prediction about the average impact of republican control of the process given that the electoral conditions are identical to the electoral conditions in a particular election. Right. So it

shows that the impact of gerrymandering is, for instance, different depending on the electoral conditions as they differed between 2012 and 2014.
Q. So predictions for 2012, 2014 are covered by the model, right?
A. Yes. That is what covered by the model.
Q. All right. So you present models -- okay. So what is the dependent variable in your model?
A. The dependent variable is the deviation in democratic seats won from historical expectation given a certain vote share.
Q. Okay. And the -- and this dependent variable is essentially identical to the efficiency gap, right?
A. No. It uses a slightly different functional form than efficiency gap does.
Q. Okay. Explain that.
A. I'm using a probit functional form that I think is better adapted to extreme -- extreme election results on one side or other. So it ends up -- when -- the model that I use ends up I think rather coincidentally being very close to efficiency gap when one party wins say between 40 and 60 percent of the vote. They deviate fairly strongly when one party wins an overwhelming percentage of the vote.
Q. Okay. So other than that, would you expect there to

be any material differences between your dependent variable calculations and the efficiency gap?
A. Again my calculations would lead different results in cases where states deviate strongly from -- from parody. For instance, in Massachusetts, right, if one party won more than 75 percent of the vote in Massachusetts, efficiency gap would predict that they would win more than 100 percent of the seats in Massachusetts while my operationalization would not.
Q. Most states have democratic statewide vote shares in the 40 to 60 range -- percent range, correct?
A. I think that that is a fair characterization more often than not that most states in most years will have democratic vote shares between 40 and 60 percent. I don't think that is universally true. I don't think it is universally true of all states or in all election cycles.
Q. So okay, so in these circumstances the efficiency gap is about equal to your dependent variable, correct?
A. In cases where the parties are fairly close to parody, my dependent variable will be fairly close to efficiency gap. Yes.
Q. Gotcha. Okay. And the independent variables include both which institution was responsible for

redistricting and other demographic and political information from the state, correct?
A. Yes.
Q. Okay. Do you trust the model's predictions for 2012, 2014?
MR. KEENAN: Object as vague as to the word "predictions."
THE WITNESS: I would not characterize them as predictions.
BY MR. EARLE:
Q. What would you characterize them as?
A. I would -- I would -- I would say they are assessing the effects of the dependent -- sorry, of the independent variables on deviation from historical seat expectation.
Q. Would you characterize the models as reliable for 2012 and 2014?
A. I think this is a very simple model. It is intentionally very simple.
Q. Is it reliable?
A. Yes.
Q. Okay. Goes a lot faster when you answer the question I asked. Okay.
The R-squared is 0.83 for 2012, 2012 model, correct?

A. Yes.
Q. And it's 0.57 for the 2014 model, correct?
A. Yes.
Q. You would characterize these scores as high by political standard -- by political science standards?
A. Certainly the first one is higher than the second one. I would say by political science standards R-squared values tend to be fairly low in political science so I think -- again it really depends -- in many cases it depends very arbitrarily, the R-squared values, on how you define your model, how you define your data set. So I would say in general there are many contexts in which I would not give a lot of weight to R-squared values. Right?
There's research that I've done that has very high r squared values, there's research that I've done, results I have produced has very very low R-squared values. I don't -- I wouldn't say that the lower R-squared values are necessarily implying that the model is less reliable, just that the variables that I am testing are accounting for less of the differentiation in the independent variable -- sorry, in the dependent variable than in a model that has, you know, a greater R-squared

value.
Q. Okay. Given the party in charge of redistricting in a state, the state's black and Hispanic population shares, the state's urbanization, the state's democratic vote share, and the state's number of seats, you would agree that your own model is a reliable way to assess the relative impact of geography and partisan control?
MR. KEENAN: Object as compound.
THE WITNESS: Sorry, what's the objection?
MR. KEENAN: It's a compound objection.
THE WITNESS: I would say there is always a trade-off.
BY MR. EARLE:
Q. Let me -- let me just modify. I'll change the question in light of the objection.
Answer the question first as to relative impact of geography. No, I'm going to fix it. Let's have the question reread and we'll stand with it.
No, I'll rephrase the question.
Given the independent variables, all right, your model is a reliable way to assess the relative impact of geography and partisan control, correct?
A. This particular regression model tests the relative impact of urbanization, the percentage of the state

which the census deems as urbanized. I don't know if I would conclude that that is a measure of geography as a whole.
Q. Okay.
A. **But I think it is a -- it is a test of the impact of urbanization, and that is a facet of geography, and also partisan control of redistricting. Yes.**
Q. Okay. So I'd like to go through an exercise here. Okay. And what I'm going to ask you to do is to plug in some values for Wisconsin into your model and see what we find. Okay?
A. **Okay.**
Q. Okay. So you've got a pen and paper? I think what we should do, the easiest way to do this is on Exhibit No. --
MS. GREENWOOD: I can give you an Excel if you want to use Excel.
BY MR. EARLE:
Q. But on Exhibit 21, what we're going to do is I'm going to give you some Wisconsin values, and then we can offer you -- you can write those down in red on Exhibit No. 21, and then what we're going to do is provide you with a -- a Excel worksheet where you can do your math and put your answers down on Exhibit 21. That will become part of the record.

Okay?
A. **Okay.**
Q. Ready?
A. **Sure.**
Q. Okay. Wisconsin is 6.6 percent black.
A. **Okay.**
Q. Okay. It's 6.5 percent Hispanic.
A. **Okay.**
Q. 70.2 percent urbanized.
A. **Okay.**
MR. KEENAN: 72 point what?
MR. EARLE: This is based on the 2010 --
THE WITNESS: 70.2?
BY MR. EARLE:
Q. 70.2 urbanized. Okay. That's based on the 2010 census. And its democratic vote share was 50.8 percent.
MR. KEENAN: Democratic vote share of what?
MR. EARLE: In 2012.
MR. KEENAN: Of what election?
BY MR. EARLE:
Q. Congressional elections. And 47.2 percent in 2014. It has eight congressional seats. Okay. So I'm going to give you the -- the Excel here, and the question that you're going to answer for me here is

what -- what bias would your model predict in 2012 and 2014 if Wisconsin had a bipartisan or court-drawn plan?
A. **Bipartisan.**
Q. Okay.
MR. KEENAN: For congressional districts?
MR. EARLE: Okay.
MS. GREENWOOD: Are you okay with Mac? Do you want a PC?
THE WITNESS: That's fine.
(Discussion held off the record.)
MR. EARLE: Back on the record.
BY MR. EARLE:
Q. All right. So your findings for what your model predict for 2012 and 2014 if Wisconsin had a bipartisan or court-drawn plan?
A. **Oh, I didn't do 2014 yet. I'm sorry.**
Q. Oh, you didn't do 2014?
A. **I didn't do 2014.**
(Discussion held off the record.)
MR. EARLE: Back on the record.
BY MR. EARLE:
Q. Okay. So the question -- okay. So your answer to the question which is what bias would your model predict in 2012 and 2014 if Wisconsin had a

bipartisan or court-drawn plan?
A. **This model would predict that Wisconsin would have a -- in both years, I mean the number is rounded to the same percentage, the same both years would be four percent in favor of the democratics in both years.**
Q. Do you want to check your 2012 calculation?
A. **My 2012 calculation is 5 point -- sorry, 3.58.**
Q. Is it 3.58?
A. **Sorry, what is the -- let me just make sure I have all the -- oh, you're right. You're right. I did -- Sorry.**
Q. So you have to make another adjustment here?
A. **Yeah, I just typed in one of the numerals wrong. Sorry, I'm getting 1.85.**
Q. 1.85?
A. **Yes.**
Q. And 4.392 for '14?
A. **I'm getting 4.2. Are you -- are you adjusting for -- oh, 47.2. 4.392. Yes.**
Q. So the record is going to be a little jumbled there in terms of clear questions and clear answers so what I'd like you to do at this point now is to write down your findings in red at the bottom of table 3 on Exhibit 21.

Page 86

And so you agree that your models -- as you have said, your models predict that if Wisconsin had a bipartisan or court-drawn map, it would have a modest pro democratic bias in both 2012 and 2014, correct?
A. I don't know that I would be able to say with any confidence that it had a pro democratic bias considering like a two percent bias in favor of the democratics would be a small fraction of a seat, right? It would be like 1/10 of a seat.
Q. Okay. But it's still a bias in favor of the democratics, right, given the state's actual urbanization, its racial demographics, and the political environment, correct?
A. Yes. I mean again I wouldn't characterize the confidence that I would -- of the bias.
Q. There's no republican bias?
A. I certainly could not confidently say that there is a republican bias generated from the model. Yes.
Q. All right. Let's do one more exercise, okay? I'll give you some more numbers. We're going to do -- now let's plug in the values for a state that looks like America as a whole, the United States as a whole.
A. Okay.

Page 87

Q. Okay. So according to the 2010 census -- now why don't you write this -- write this one in blue ink.
A. I should -- so I should mention the --
Q. Let's -- you don't have a question so unless you're --
A. I do --
Q. -- modifying a prior answer.
A. No, I'm not.
Q. You're not modifying a prior answer. Okay.
A. Well, I would -- so I would like to modify a prior answer.
Q. All right. Which answer -- which question are you modifying the answer to?
A. When you asked me whether these are reliable estimates for bias, whether the model -- I believe the model generates reliable estimates for bias.
Q. Right.
A. This model is only predicting for states -- for medium or large states that have greater than six congressional districts. I would not say that I am trying to provide an estimate of bias for smaller states than that. So if you're giving me data that's drawn from smaller states than that, I would not necessarily say that this model provides a reliable --

Page 88

Q. Anticipating the next -- Okay. That's fine. All right. As far as the exercise we just went through --
A. Right.
Q. -- we had enough congressional seats that we don't have that problem?
A. Yes.
Q. Okay. Good. All right. So now, a model -- we're going to apply your model in blue ink, and you've written down your results on table 3 in red ink for Wisconsin given the demographic independent variables we just gave you, right?
A. Uh-huh.
Q. Now we're going to do the state -- a state that looks like the United States as a whole. Okay? According to the 2010 census, the United States was 13.2 percent black, 17.4 percent Hispanic, 80.7 percent urbanized.
A. Uh-huh.
Q. And according to your papers, the democratic share of the two-party congressional vote was 51 percent in 2012 and 47 percent in 2014. And the average state had nine congressional districts. Okay? Using these variables, what would be the predicted bias if the average state had a bipartisan or

Page 89

court-drawn map in 2012 and 2014? Okay. Plug those numbers in. And we'll go off the record while you do the math.

(Discussion held off the record.)

MR. EARLE: Back on the record.

MR. KEENAN: I'm just going to interpose an objection that this hypothetical has no basis in fact, but you can answer.

MR. EARLE: Did you get the objection?

THE COURT REPORTER: Yeah.

MR. EARLE: You're objecting to the hypothetical?

MR. KEENAN: Yeah, I mean you can ask hypotheticals, but you have to have a basis and evidence in fact, and I'm saying it's not.

MR. EARLE: I think it's the United States census 2010.

MR. KEENAN: Well, you're asking to assume that every state has the same --

MR. EARLE: It is a hypothetical with a state with an average number of congressional districts matches a proportion of the United States census demographics, and it's plugged into his model to see what kind of result it gives.

BY MR. EARLE:

Page 90

Q. So why don't you give us your values.
A. Again I just want to mention I think the model is only for larger states.
Q. I okay. I understand. This is one of the things about depositions and it's also true about trials. You've got to answer the questions that are in front of you, and it's not an opportunity to speak openly and this is not a dialectic here. Okay. It's a question and answer. I get to ask the questions, you get to answer them. Okay. So you got the question. What's the answer?
A. I'm getting a less than one percent bias in favor of the democratics in both cases. 0.6 percent in 2012. 0.2 percent in 2014.
Q. Okay. I think you want to look at 2014. I think it should come out to 1.6.
A. Let's see. Oh, you're right. I got the wrong -- I'm having a problem here. 1.6. You're right.
Q. So there's a slight pro democratic bias. So given again -- so again, giving these -- given these values that we gave to you just now, your models again show a slight pro democratic bias, correct?
MR. KEENAN: Object as vague.
THE WITNESS: So --
BY MR. EARLE:

Page 91

Q. That's the question. Answer the question I just gave you. And I'll have the court reporter read the question again. We have an objection to the form of the question, and I'm going to have the court reporter read the question to you. Answer that question. If you want to give other testimony later under other circumstances about stuff, that's fine. But right now you have one question in front of you.
MR. EARLE: If you could read the question.
(Question read: So there's a slight pro democratic bias. So given again -- so again, giving these -- given these values that we gave to you just now, your models again show a slight pro democratic bias, correct?)
THE WITNESS: Yes, given the caveat that there are a lot of people who would object to characterizing a coefficient that I don't have a statistical confidence level on.
BY MR. EARLE:
Q. You don't have a what?
A. A level of statistical confidence on. The coefficient value is very small.
Q. Your models don't show any pro republican bias; isn't that true?
A. Given the values for the independent variables that you gave me, the model does not show a pro republican bias.

Page 92

Q. And you acknowledge that the values we gave you come from the 2010 United States census, correct?
MR. KEENAN: Objection as vague.
THE WITNESS: Yes, although that includes states the model is not meant to apply to.
BY MR. EARLE:
Q. Okay. All right. So you used the term hyper-responsive, right?
A. Yes.
Q. And hyper -- hyper-proportionate?
A. Yes.
Q. You used those terms interchangeably?
A. I use those terms casually to refer to the same concept.
Q. So you use those terms casually to refer to the same concept in your report?
A. Yes.
Q. Okay. And I don't understand what you mean by "casually."
A. Do you want your laptop back?
MS. GREENWOOD: Yes.
MR. KEENAN: Can I make a request that we save this document as an Excel file?

Page 93

MS. GREENWOOD: And I'll send it to you.
MR. EARLE: Let's do this. Let's actually -- I'll one up you. Brian, let's print it up and mark it and attach it to the -- and make it -- we'll do that right now.
MS. GREENWOOD: Okay. Yeah. Did you have them as separate calculations or did you put them back into the same cells?
MR. KEENAN: I had them as separate calculations. They're unlabeled. I'll label them.
MR. EARLE: Why don't you label them, we'll take a break, we'll print them out and mark them for the record.
(Discussion held off the record.)
(Exhibit No. 22 marked for identification.)
BY MR. EARLE:
Q. All right. So we're showing you what's been marked as Exhibit 22. This is a printout of the calculations we've just gone through for both the United States average state and the State of Wisconsin in 2012 and 2014 for each?
A. Yeah.
Q. Okay. All right. So that's now part of the record in your deposition. Okay.
I want to draw your attention to this article

that you wrote entitled Redistricting, Risk, and Representation: How Five State Gerrymanders Weathered the Tides of the 2000s. We'll mark that as Exhibit 23.
(Exhibit No. 23 marked for identification.)
BY MR. EARLE:
Q. This is an article you authored, right?
A. Yes.
Q. It's a peer-reviewed article?
A. Yes.
Q. Okay. And drawing your attention to page 8, I guess, of the article, the section that's Section 2, Dimensions of Representation, and under that subsection A, Bias and Responsiveness. And if you look at the second column, right above the reference to table 1 in the middle of the page, there's a quote there that I have in mind, which is -- begins -- the words begin, "The relationship between seats..."
A. Yes.
Q. Would you read that quote for the -- from that sentence through the end of the paragraph?
A. The relationship between seats and votes under one regime could be considered unresponsive if it displays a higher -- if it displays -- excuse me --

if it displays a responsiveness slope much below 2, and hyper-responsive if this slope is substantially greater than 2.
Q. Okay. And okay. And given your article, do you think it's fair to characterize a responsiveness -- well, that's an accurate statement, right? I mean you stand by that statement in your article?
A. Yes. Given -- defining responsiveness as how much a change in votes change the number of seats a party won compared to an historical average, yes, that is accurate.
Q. So it would have to be substantially off the slope, right, greater than 2 for it to be hyper-responsive?
A. Right.
Q. Correct?
A. Yes.
Q. Okay. And how would you characterize a responsiveness of 2, exactly 2?
A. I would characterize that as average responsiveness compared to historical trends or historical averages, historical observations.
Q. Okay. I lost my spot here. Hold on a second. And based on your definition of hyper-responsiveness in this article, that would not qualify as hyper-responsiveness? A slope of 2 would not -- a

slope --
A. Yes.
Q. A slope of 2 would not qualify as hyper-responsiveness, correct?
A. Given how I'm defining hyper-responsiveness in this article, yes.
Q. Okay. All right. Now, and that's contradictory to what you wrote in your report on page 5, correct? I'm sorry, page 6. You're right. Page 6.
A. I don't think I define hyper-responsiveness in my report, and I think I stated earlier that I used the term casually.
Q. That's what you meant by using the term casually, you were stretching that in -- in your report from what you indicated substantively in your article?
A. What I indicated substantively in the article was that I was defining hyper-responsiveness for the purpose of the article as a deviation from historical average. I don't think that is necessarily how a lay person would define responsiveness, particularly in the context of comparing it to proportionate representation, which is what I'm doing in the report.
Q. Okay. So in your report at the bottom of page 5, you say -- and I'm reading from your report, the

bottom page 5. You tell me if I read this incorrectly. And the court has additionally been wary of adopting a standard for partisan gerrymanders that would amount to proportional representation, yet the efficiency gap test would codify a very specific translation of seats to votes that is essentially -- essentially, quote, hyper-proportional, close quote, representation.
Did I read that correctly?
A. Yes.
Q. All right. So you're using that term as -- and the efficiency gap does not deviate from the slope of 2, does it?
A. The -- the term that you're using in your quote is hyper-proportional, not hyper-responsive. It's a different term.
Q. Five minutes ago you said the terms were equivalent.
A. I said in the report I casually used the term hyper-responsive to be equivalent to hyper-proportional. It is clearly defined in the article that you gave me as being a certain definition, which I think would not be the same as hyper-proportional.
Q. Okay. In your report, subsection B on page 6, you say that an efficiency gap may discourage drawing

Page 98

competitive districts. Do you read that there?
A. **I'm sorry, can you point out it again?**
Q. It's page 6.
A. **Oh, yes.**
Q. Okay. You say, "An efficiency gap standard may discourage the drawing of competitive districts"?
A. **Yes.**
Q. And you say this is an example of a normative value. What are normative values in your mind?
A. **Values that a person who is designing a political system may wish to imbue their system with in order to represent some idea of good government.**
Q. Okay. Do you know whether competitive districts is a -- a value defined in Wisconsin law for purposes of redistricting?
A. **I do not know. No, I don't know.**
Q. Okay. Do you know what the values are that are defined for purposes of redistricting in Wisconsin?
A. **Can you be more specific?**
Q. Well, do you know what they are? I don't want to answer the question I just gave you. I want your answer to the question.
A. **I don't know that Wisconsin law states that the drawing of maps requires the consideration of certain values.**

Page 99

Q. You don't know that?
A. **Specific to Wisconsin law I don't know.**
Q. Okay. What are Wisconsin's legal requirements for redistricting plans? Do you know?
A. **I don't know specifically Wisconsin's legal requirements beyond the standard federal requirements.**
Q. Okay.
MR. EARLE: This is a good time to take a break. Let's take a break for lunch.
(Lunch break taken 11:31 a.m. to 12:25 p.m.)
BY MR. EARLE:
Q. Nick, do you know Keith Gaddie?
A. **Are you asking if I know him personally?**
Q. Yeah.
A. **No.**
Q. Have you read any of his work?
A. **I don't recall specifically. I feel like I have, but it's -- nothing is -- I -- yeah, I don't recall specifically.**
Q. Okay. All right. Now, in your report, you have basically five opinions?
A. **Okay.**
Q. Correct? And each of those opinions is -- is expanded upon in the body of the report with your

Page 100

reasoning and the support that you have for those opinions, correct?
A. **Okay. Yes.**
Q. Yes. Okay. And I just want to go through and see if we can just kind of corral those a little bit more precisely.
A. **Okay.**
Q. You understand that under Rule 26, you have to state all of your opinions in your report, and as worded by the rule itself, that your report must contain a complete statement of all opinions that the witness will express and the bases and reasons for them, and all the facts and data that you considered in forming your opinions and any exhibits that will be used to summarize or support those facts or data. All right? You understand that?
A. **Yes. It sounds like you were saying that what I will be expressing -- yes, I understand that.**
Q. Okay. So I'm looking at your report, and in the context of Rule 26 requirements, and you have five opinions that you will express, and you expand upon the bases and reasons for those opinions in the body of the report that corresponds to each of the five opinions, correct?
A. **Yes.**

Page 101

Q. Okay. And the first opinion you have -- and I'm going to state it to you as I understand it, and you tell me if I'm right or wrong, all right?
A. **Okay.**
Q. Your first opinion is that a high efficiency gap doesn't mean an unbalanced map, rather a high efficiency gap implies a deviation from a pre-determined seat/vote curve that discourages normatively desirable objectives such as maximizing competition and proportionality. Correct?
A. **Okay. I -- the last clause I think you'd have to -- I'm not saying that a high efficiency gap itself discourages those particular -- use of those particular normative standards. I'm saying that adopting -- adopting a legal standard where a high efficiency gap would imply presumptive unconstitutionality of a map, adopting that standard could potentially discourage the use of those normative values in the drawing of districts. Does that make sense?**
Q. No, it doesn't make sense because as I would understand this those normative values would be responses that would legitimize the map in the face of the inference. Isn't that correct?
A. **Could you repeat the question?**

Page 102

Q. I'll have the court reporter read it to you.
(Question read: No. It doesn't make sense because as I would understand this those normative values would be responses that would legitimize the map in the face of the inference. Isn't that correct?)
MR. KEENAN: I'm going to object as vague to the extent that it's asking Mr. Goedert to apply the burden shifting frame of plaintiffs' test.
MR. EARLE: In other words, you're objecting to the form of the question?
MR. KEENAN: Yep.
MR. EARLE: As opposed to trying to answer the question yourself?
MR. KEENAN: Well --
THE WITNESS: I think that adopting a test that would make something presumptively unconstitutional would discourage the drawing of maps that would be presumptively unconstitutional even if they could be rebutted by some other standard.
BY MR. EARLE:
Q. Okay. All right. So I guess what I'm trying to figure out here is what is the meat of that first opinion?
A. Okay.

Page 103

Q. And as I understand it, that -- that a -- first, a high efficiency gap in your view does not mean an unbalanced map. Is that a part of your opinion?
A. Yes, an observation of a high efficiency gap does not imply that a map is unbalanced.
Q. Okay. That's -- that's the basic opinion in the first opinion that you have, right?
A. Yes.
Q. Okay. And then you go on and elaborate that the -- a large efficiency gap implies deviation from a pre-determined vote/seat curve representing hyper-proportionate or hyper-responsive representation?
A. Yes.
Q. All right? But a high efficiency gap is based on what kind of a -- of a -- of seats-to-votes curve?
A. It's based on a curve that would -- that a one percent increase in the number of votes that a party receives should correspond with a two percent increase in the number of seats.
Q. So your use of the term hyper-proportionate and hyper-responsive representation being a dev -- a deviation, right, is wrong, right?
A. I'm sorry, I would think that hyper-proportionate -- so proportionate would mean a one percent -- a one

Page 104

percent increase in votes would correspond with a one percent increase in seats.
Q. That's what you think?
A. Hyper-proportionate -- that's how I would describe proportionate or proportional representation. Yes.
Q. Okay. And where does -- you would agree that the United States has exhibited a hyper-proportionate seats/vote curve over the history -- over history?
A. Yes. The historical average responsiveness of -- and I have in particular studied congressional elections -- of the congressional elections that I've studied does display a hyper-proportionate response to changes in vote share, on average.
Q. Okay. So this is your first opinion. We'll -- we'll elaborate on that in a little bit. I just want to nail down that's the first opinion you've got?
A. Yes.
Q. Okay. Second opinion, that an EG threshold of seven percent is a highly -- highly unstable metric and doesn't inform future efficiency gaps or durability. Is that your second opinion?
A. Are you quoting from something here?
Q. No, I'm reading -- I'm just characterizing.
A. I don't know if I would say it doesn't inform at

Page 105

all. I would say it is a very weakly informative signal of future efficiency gaps.
Q. And the basis for that is?
A. The basis for that is prior research on efficiency gaps as well as the expert report of Jackman.
Q. Okay. What is the prior research?
A. Stephanopoulos and McGhee.
Q. So your interpretation of Stephanopoulos --
A. My interpretation of Stephanopoulos and McGhee --
Q. We can't talk over each other.
A. Sorry.
Q. So your interpretation of the Stephanopoulos and McGhee article is that it supports your view that the efficiency gap threshold of seven percent is highly unstable and is a weak informer of future efficiency gaps and durability, correct?
A. Yes.
MR. EARLE: Okay. Mark this exhibit.
(Exhibit No. 24 marked for identification.)
Q. Showing you what's been marked as Exhibit 24. Please show me where in this report you draw that.
A. (Witness reading.) Page -- page 26, second full paragraph.
Q. Okay.
A. The second most specifically -- I would say the

Page 106

paragraph as a whole, most specifically the second to last sentence beginning specifically. "Specifically, a plan's efficiency gap in one election is a relatively weak predictor of its gap in the next election, coefficient equals 0.23, in a model that also includes a variety of other factors. Many partisan gerrymanders, therefore" --
(Court reporter interrupted.)
THE WITNESS: I'll stop there.
Q. So when the efficiency gap is small, it's not a good predictor is what you're saying?
A. This is overall measures of efficiency gap are a relatively weak predictor as I interpret the statement.
Q. Okay.
A. I could find other instances where you have -- I also think that if you look at the graphs that they show -- and my reference to that is a little bit awkward considering one of the authors is in the room.
MR. STEPHANOPOULOS: Criticize all you want.
BY MR. EARLE:
Q. What page?
A. So this is on page 38 and 39. You can see that many of the maps which exceed their threshold, which I

Page 107

believe is -- so on page 39, eight percent in one direction or another, many of the maps that exceed that threshold when observed throughout the decade observe a wide range of efficiency gaps in other years in the decade, including efficiency gaps that cross over to the other side of bias.
Q. All right. And then you say the other basis that you have is the Jackman report?
A. Yes.
Q. Okay. You've got that exhibit in front of you?
A. Oh, it's right there.
Q. Let's hold off on that. We'll come back to this. I just want to try to get these opinions reduced to -- to a clear couple of sentences. Okay. So how would you state your second opinion then in a couple sentences?
A. Let me just look at my report to make sure I'm referring to the right --
Q. Yeah, you summarize them on the -- on page 2.
A. Yeah.
Q. State it in two sentences.
A. I would say that the plaintiffs' alleged threshold for unconstitutionality of seven percent in a single election is not a strong or particularly informative signal of what an efficiency gap will be in future

Page 108

elections.
Q. Okay. Let's go to the third one. Would you state your third opinion in one or two sentences?
A. The third opinion is differentiating between the standard as expressed in the complaint and the standard as expressed in the other academic research or suggested in the other academic research as to how efficiency gap should be applied to determine constitutionality. The other academic research, specifically the Stephanopoulos and McGhee article that I was referring to, also requires that a sensitivity step -- a sensitivity test be applied to measure I suppose the hypothetical durability of a map sufficiency gap, and that this is not stated in the complaint, that this is not part of the -- the test as stated in the complaint, and that this is really -- something along these lines is essential to determine durability of an efficiency gap. And also I am not certain that the test as expressed in even in the Stephanopoulos and McGhee is sufficient to establish the durability of the efficiency gap in a map.
Q. That was a lot more than two sentences.
A. I'm sorry.
Q. Give me two sentences. What is the opinion?

Page 109

A. The opinion is that the complaint does not sufficiently establish the durability -- the test suggested in the complaint does not sufficiently establish the durability of a efficiency gap.
Q. That's your third opinion?
A. Yes.
Q. Okay. And what is your fourth opinion? Is it accurate to say that your fourth opinion is that Ken Mayer's demonstration map is hindsight based on 2012 results not available at the time of drawing?
A. Yes.
Q. That's it?
A. There are some other I think less important quibbles that I would have with the -- the way that Mayer is -- is drawing up his demonstration plan, but that is the most important point that I am making in the --
Q. And your fifth opinion is that any judgment about partisan bias must account for the political geography that favors republicans supposedly?
A. Yes.
Q. And that's an accurate statement of your fifth opinion?
A. I think I say should account for bias, but --
Q. Okay. So those are -- those are the five opinions

that you're going -- that you've rendered in your report, and the rest of the report represents your reasoning basis for each of those opinions, correct?
A. Yes.
Q. Okay. All right. So you -- you assert that -- that increasing competitiveness or achieving proportional representation are legitimate goals that might result in a large efficiency gap, correct?
A. Yes.
Q. Okay. And but you have no reason to think that Act 43 was intended to increase competitiveness or achieve proportional representation, do you?
A. I have no specific knowledge that would suggest that was a goal.
Q. Okay.
A. But I don't have any specific knowledge related to much about the intent behind that act.
Q. Okay. How many states other than Arizona include competitiveness as a legal criteria for district lines?
A. It is certainly a minority difference between congressional and state legislative maps. I don't know the number off the top of my head.
Q. It's zero, isn't it?
A. I don't --

Q. Now, you paused. So I'll --
A. If you're talking about stated in the law, that might be true.
Q. Okay. You can't name any other state other than Arizona as you sit here in this deposition; isn't that true?
A. Sure.
Q. Okay. And how many states include the achievement of proportional representation as a legal criteria for districting plans?
A. I'm fairly sure that none do.
Q. Okay. Are you familiar with -- you use -- so it's zero, right?
A. I believe -- I'm not aware of any.
Q. Okay. So the answer is that it's zero. Zero states require that, correct?
A. I believe that's true.
Q. Okay. You use examples from Arizona and California, right?
A. Yes.
Q. In your report. And these are both congressional examples, not state legislative examples?
A. Yes, they're both congressional examples.
Q. Now, Chen and Rodden -- I'm sorry. In his expert report Jackman quotes Stephanopoulos and McGhee as

follows on page 24. You want to open the report?
A. This is the Jackman report?
Q. Yeah. Uh-huh.
A. Okay.
Q. All right. And we're going to look at page 24. And so Professor Jackman quotes Stephanopoulos and McGhee as, quote, we strongly discourage analysts from either dropping uncontested races from the computation or treating them as if they produced unanimous support for a party. The former approach eliminates important information about a plan, while the latter assumes that coerced votes accurately reflect political support, period, close quote. I concur with this advice, close quote. All right.
Do you agree with Jackman, Stephanopoulos, and McGhee that uncontested races should neither be dropped nor treated as if they produced 100-percent-to-zero outcomes?
A. I think it depends on context.
Q. Okay. And yeah, what's the context that matters to you in answering that question?
A. I think there are a variety of perfectly acceptable things that could be done to -- in the treatment of uncontested races, and Stephanopoulos and McGhee adopt one method and Jackman adopts two different

methods, and of course the Mayer report adopts a totally different method. I don't have any specific objection to any of those.
Q. Okay. So in your opinion -- in your opinion you don't object to any of those?
A. Not for the purpose of measuring -- estimating efficiency gap in a particular election.
Q. Okay. Let's go -- let's go on to what you did then because when you calculated the efficiency gap for Arizona's congressional map from 2002 to 2012 on pages 7 and 8 of your report, and the California congressional map in 2008 at page 10 of your report, didn't you treat uncontested races as if they produced 100-percent-to-zero-percent outcomes?
A. I did not do any imputation for uncontested races. That's true.
Q. Right. So you --
A. Yes.
Q. You treated them as -- as producing a 100-percent-to-zero-percent outcome?
A. I believe that's accurate.
Q. Isn't it correct that you don't know what the plan's efficiency gaps would have been if you hadn't treated uncontested races that way?
A. Well, because I have not done any particular

imputations for uncontested races, that's true, I do not know what the results would have been if uncontested races had been imputed with some sort of -- under one of the methodologies of the report or the scholarship. That's true.

Q. In their article Stephanopoulos and McGhee state, quote, we -- we report the efficiency gap in seats for congressional plans and in seat shares for house -- state house plans. What matters in congressional plans is their impact on the total number of seats held by each party at the national level. Conversely, state houses are self-contained bodies of varying sizes for which seat shares reveal the scale of parties' advantages and enable temporal and spatial compatibility, close quote. That's at page 868 -- 869 of the Stephanopoulos and McGhee article, the final version.

MR. STEPHANOPOULOS: This is -- this is the same text but without the final page numbers.

MR. EARLE: Okay. So what page is that on?

MR. STEPHANOPOULOS: I don't know. I'll have to find that.

MR. EARLE: We'll take a quick break and get that for you because there's a variation between the exhibit and the -- my notes here.

THE WITNESS: If you know what section it is, you could probably find it more easily.

MR. EARLE: We've got it right here. It's coming up.

MR. STEPHANOPOULOS: Page 29.

MR. EARLE: Page 29.

MR. STEPHANOPOULOS: The bottom of page 29 to page 30.

MR. EARLE: To page 30.

THE WITNESS: Okay. I see where you're talking about.

BY MR. EARLE:

Q. Yeah. All right. So you see the quote. And you heard the quote that I read?

A. Can you tell me where you're getting the quote again?

Q. We report the efficiency gap in seats for congressional plans and in seat shares for state house plans. What matters in congressional plans is their impact on the total number of seats held by each party at the national level. Conversely, state houses are self-contained bodies of varying sizes for which seat shares reveal the scale of the parties' advantage and enable temporal or spatial comparability. All right?

A. Okay.

Q. And you agree that that approach is reasonable, correct?

A. A reasonable way of reporting the data?

Q. Yeah.

A. I -- yes, I think it's reasonable.

Q. Okay. And so why do you report your efficiency gaps for California and Arizona in percentages rather than seats since those are congressional?

A. Because I think my approach is reasonable as well.

Q. Okay. Do you know what the efficiency gaps would be in seats rather than percentages?

A. I could refer to my report and figure out what the -- very quickly off the top of my head. I mean Arizona had eight congressional seats in 2002 to 2010 and nine in 2012. So the greatest efficiency gap it looks like would be a little over one seat in 2002 or a little over one seat in 2012. In the case of Arizona -- did you just ask me about Arizona or are you asking about California as well?

Q. Arizona. Arizona is fine.

A. So slightly over one seat would be the greatest deviation.

Q. And what would be -- what would it be under Stephanopoulos and McGhee's? And that -- oh, I'm

sorry. And that would be -- I'm trying to -- and that would be under Stephanopoulos and McGhee's proposed two-seat threshold at all times, correct?

A. Well, you asked me if I thought it was a reasonable way to report the data. You didn't ask me about the reasonableness of the threshold.

Q. Okay. Okay. Let's go to the next section. Okay.

On page 11 of your report, I draw your attention to the quote: "Yet both the academic research and data presented by the plaintiffs' expert show that such intent cannot be inferred." It's the last sentence on the first paragraph of page 11.

A. Yes.

Q. Do you have any objection to efficiency gap scores when they're being used to establish effect rather than intent?

MR. KEENAN: Just object as vague.

THE WITNESS: I do not object to them being used as a summary measure for deviation from a pre-determined seats/votes curve.

BY MR. EARLE:

Q. So it's --

A. So I wouldn't necessarily say that a particular efficiency gap implies that some particular factor

Page 118

has an effect. Does that make sense?
Q. No. No, it doesn't.
A. I don't think that a particular efficiency gap measure implies that some -- implies any particular thing has an effect.
Q. Well, it -- isn't it a measure of the bias effect?
A. It is not a measure of the bias. It is a measure of the deviation from a pre-determined seats/votes curve.
Q. Okay.
A. And as I think I demonstrate in the first section of the report, there can be a variety of ways in which an unbiased map can show a high deviation from that pre-determined seats/votes curve.
Q. Okay. All right. So then you go on and you -- also on page 11, you say, quote, past results demonstrate enormous instability even within a given decade and sensitivity to very realistic partisan ties, close quote. Right?
A. Okay.
Q. Okay. And you can't know if a -- so the implication here is that you can't know if a plan will go on to advantage one party over another just from the first election? Is that what your implication is?
A. I think that you can't know -- I would say you

Page 119

cannot be particularly confident about that one -- that one efficiency gap measure in one particular election implies that the efficiency gap will be in the same direction in a subsequent election.
Q. Okay. What do you understand Jackman's methodology in calculating the plus/minus seven percent threshold?
A. I'm sorry, what do you mean -- do you mean what is his methodology for calculating the efficiency gap?
Q. No, plus/minus seven percent threshold.
A. Calculating the threshold? I'm sorry, the threshold is pre-determined.
Q. Setting it.
A. Oh, I don't -- other than the fact that this is the threshold that's stated in the complaint, I am not sure why he set the threshold here, although I do note that at the end of his report he has some data with respect to this threshold and some data with respect to his confidence that -- that an efficiency gap observation in the first decade after redistricting will imply efficiency gap in the same direction in subsequent years. I don't know if he set the threshold because that was the threshold that the plaintiffs wanted him to set or whether he set it for another reason. That I'm not sure of.

Page 120

Q. Why don't you go to figure 32 in the Jackman report. It's on page 67.
A. Sixty-seven. Okay. I see it.
Q. Looks remarkably like our art museum here at the lakefront. Would you explain to us or describe for us what Jackman is displaying in figure 32?
A. Give me a moment. He has a lot of figures. (Witness reading.) I believe this represents Jackman's confidence that an efficiency gap observed in a first election of a certain number would be durable to a certain rate of confidence over the rest of the decade.
Q. Okay. And can you identify any flaws in how Jackman assembled figure 32?
A. Flaws in how he assembled it?
Q. Yeah. That's the question.
A. Do you mean do I think it is a correct representation of the data that he says he is representing?
Q. That's -- that's --
A. Yes, I think it is a correct representation of the data that he is representing. I don't have any reason to believe it's not.
Q. And what's the confidence rate associated with an EG of below minus seven percent?

Page 121

A. I mean it looks like it's just over .95.
Q. And that's above the 95 percent confidence commonly used in the social sciences, isn't it?
A. Yes, but I would disagree that his method actually represents a useful measure of what will happen in the future.
Q. Well, the answer to my question is yes?
A. Yes.
Q. Okay. Thank you. All right. And --
MS. GREENWOOD: Can we take a break?
MR. EARLE: Yeah. We can take a break.
(Break taken from 1:00 p.m. to 1:10 p.m.)
MR. EARLE: Back on the record.
BY MR. EARLE:
Q. Okay. On page 13 of your report, subsection A, you've got a -- we're going back to this -- and this is I guess opinion number three. Okay. "The plaintiffs' complaint does not include a crucial second part to the empirical test for presumptive unconstitutionality," which you define as sens -- "sensitivity testing for future results." Correct?
A. Yes.
Q. All right. And how would you recommend that that sensitivity testing be carried out?
A. I would have no particular recommendation as to how

Page 122

the sensitivity testing be carried out since I'm not the one proposing the test.
Q. Okay. But if you were asked to carry out sensitivity testing for that work, for those calculations, how would you do it as a matter of methodology?
A. So if you're asking if I were asked what is the likelihood that an efficiency gap would -- observed in a particular election would --
Q. No, observed in the first election after the map -- the decennial cycle, the first election after the map is drawn.
A. That an efficiency gap observed in the first election after a particular cycle -- if I were being asked what the likelihood that that efficiency gap would persist throughout potential future elections in the decade, with the caveat that I am not suggesting that this should be the test for constitutionality, just if I was asked to do that from an academic perspective, I think what I would want to do is develop some sort of measure for the plausibility of future overall electoral environments in some way --
Q. I'm sorry, future what?
A. The plausibility of future electoral environments,

Page 123

and by that I mean the overall statewide vote share for a particular party. All right. So there could be an electoral environment where there is a democratic wave where the democrats get 58 percent, and that might occur with some probability. You know, you might have a 50/50 election with some probability. You might have a 60/40 republican election with low probability, but some probability. So you'd probably want to develop some sort of methodology to think about what the range of possible electoral environments would be.
Q. So is that a long-worded way of saying that you would recommend using a uniform swing assumption?
A. Right. And then applying that range and situating the current -- the immediately previous or the observed election results within that range, I would probably do something like applying a uniform swing.
Q. Okay. So --
A. I think -- I think that's not -- off the top of my head that's the sort of a reasonable way to do that.
Q. Okay. So you would base that -- that -- that uniform swing assumption based on future electoral environments based on past electoral data, election data, correct?
A. Yes. Based on past election data. Yes.

Page 124

Q. And now, in their article, Stephanopoulos and McGhee make this determination by looking at the variation that has historically occurred in state legislative elections using the entire range from the 10th to the 19th -- 90th percentile of this historical variation. You -- you agree this is a reasonable approach?
A. Can you point to where in the article you're finding that?
Q. Sure. Coming right up.
MR. STEPHANOPOULOS: Look on page --
MR. EARLE: Look on page --
MR. STEPHANOPOULOS: -- 35.
MR. EARLE: 35.
MR. STEPHANOPOULOS: So it's the beginning of section 3B.
THE WITNESS: This is footnote 153 that you're drawing from?
MR. STEPHANOPOULOS: Yeah.
MR. EARLE: Yeah.
THE WITNESS: And is your question whether I would object with the --
BY MR. EARLE:
Q. No, the question is that this is a reasonable approach, isn't it?

Page 125

A. What I think is not particularly reasonable about this specific approach is that you are taking the result of a particular election and swinging that result, rather than trying to situate that particular election result within the context of possible election results and altering your swings based on where that particular wave that you're observing happened within the range of possible election results.
Q. I'm not sure I got what you just said. Can you -- could you restate that? Because the question is the methodology or approach exemplified by the Stephanopoulos and McGhee in the footnote 53, that's a reasonable approach?
A. Can I give an example of where I think it might not be reasonable that might eliminate this?
Q. Well, first from a methodological point of view is it a reasonable approach?
A. I think there are aspects of it that are not reasonable.
Q. Okay. What would you recommend instead?
A. Again stating that this is not what I would recommend as a test of constitutionality, but specifically if I were asked the empirical question of what is the likelihood that an efficiency gap

Page 126

will endure to be the same sign in future election
results given the result of a particular -- one
particular election, presumably --
Q. The first --
A. -- the first --
Q. -- election.
A. -- right, what would I do? What you -- what I think
you would want to do is you would want to figure out
the range of possible election -- possible overall
election results, say statewide election results,
all right?
Q. Uh-huh.
A. You would want to situate the actual result that was
observed within that range.
Q. Uh-huh.
A. And that might cause you to want to deviate more in
one direction than in another, right? So, for
instance, if we took a result like 2008 where
overall the democrats say won the overall vote by 11
percent, this is nationwide congressional popular
vote, right? If we were to deviate that 11 percent,
say what is the 7.5 percent in either direction, all
right, that wouldn't give you the range of possible
election results because it wouldn't give you the
possible election result that happened two years

Page 127

later which was that the republicans won the
national result by I think seven or eight percent.
So you'd want to come up with some sense that
that was a -- that particular election that was
observed was one that lie -- that lied on the
extreme of the range of possible election results.
And so when you were doing the sensitivity testing,
you would want to test for more sensitivity in the
republican direction than in the democratic
direction.
Q. Okay. So if you did that, you would think the
results of the sensitivity testing are reliable
then, right?
A. I would think that they would give you a fairly
accurate estimate for the likelihood that an
efficiency gap would persist throughout the decade.
Q. Okay. Thank you. So now, when Jackman calculates
the odds that a plan with a certain efficiency gap
in this first election will flip signs over its
lifetime, those are figures 29 and 30 --
A. Uh-huh.
Q. -- he takes into account all recorded plans, doesn't
he?
A. I believe he excludes a number of plans for various
different reasons.

Page 128

Q. Well, all reported plans for which data is
available?
A. Yes, I believe he includes all the plans which meet
certain qualifications. Right. Like he excludes
multimember districts and various things like that.
Q. Yeah. That's what you meant. I understand.
A. Yes.
Q. So when Jackman calculates the confidence rate
associated with different efficiency gap thresholds,
he again takes into account all recorded plans for
which the data is available, right? Figures -- and
I'm referring here figures 32 and 33. Doesn't he?
A. Okay.
Q. Is that correct?
A. Those are his data points.
Q. Right. Don't all recorded plans exhibit a greater
total variation than sensitivity testing would
capture?
A. So my understanding of this graph is that he is only
looking at the first election result following
redistricting.
Q. Right.
A. Right?
Q. That's -- that's the -- the --
A. Yes.

Page 129

Q. -- the threshold EG that is used --
A. Yes.
Q. So you asked me a question. That is the -- the
first EG that he's using to determine durability.
A. Yes. So he's only looking at a very narrow range of
actual electoral environments, those being the
specific environments that occurred in 1972, 1982,
1992, 2002, I believe.
Q. That's all election environments that historically
preexisted the EG; isn't that right?
A. He's only looking at the specific electoral
environments that occurred in those four specific
years. Those do not encompass the range of possible
electoral environments that might happen in the
future. And in fact, I think they don't include
what we would normally consider wave elections or
more extreme election results that might generate
less durability.
Q. Isn't he looking at all -- after -- isn't he looking
at all elections that occurred after the 1972, 1982,
1992, 2002 elections?
A. What he is observing here is what is the likelihood
that given the election results, say for instance,
in 1992, what is the likelihood that it will deviate
in future elections in the 1990s? All right?

Page 130

However, what he is not looking at is given the wave election that occurred in 1994 -- 1992 was a roughly evenly balanced national election, 1994 was not. He doesn't include any elections in his baseline here that would be considered wave elections at a national level.

And there's no reason why if we're applying this in the future we wouldn't observe a wave election at a national or statewide level during the first election following redistricting.

Q. Well, in order to develop or determine that likelihood, he considers all election results --

A. That's not --

Q. -- not just the 1972, 1982, et cetera; isn't that so?

A. No, he is not -- this graph has nothing to do with the likelihood that a wave election will occur. He's observing the likelihood that an efficiency gap will be observed given a wave election.

Q. Right.

A. Not -- not the likelihood that a wave election will occur. And what I'm saying here is that there is a completely reasonable likelihood that a wave election could occur in the first election cycle following redistricting, which would generate

Page 131

completely different results with respect to the durability of the efficiency gap during the election subsequent in that decade than what Jackman observes in his graph.

Q. Do you agree that a uniform swing assumption is not entirely reliable?

A. I agree that it's not entirely reliable.

Q. Do you agree that Jackman's approach avoids the reliance on the uniform swing assumption?

A. Yes, Jack -- well, Jackman's approach does not perhaps have some of the problems that a uniform swing assumption might have, but I think his methodology is much more problematic in other ways.

Q. All right. If you -- if you weren't going to carry out sensitivity testing, how would you recommend setting the efficiency gap threshold using past electoral data?

A. I'm not recommending setting a threshold for constitutionality of an efficiency gap.

Q. But if you were asked to do that, how would you do it?

MR. KEENAN: Objection. Calls for a legal conclusion.

MR. EARLE: No, I'm asking how he would do it.

MR. KEENAN: Yeah, how he would set a legal

Page 132

standard. So like --

MR. EARLE: How you would -- I'm asking -- well, you can call it a legal standard. We're asking him to set the threshold.

THE WITNESS: I would not set a threshold because I don't believe efficiency gap is an appropriate measure of the constitutionality of a gerrymander.

BY MR. EARLE:

Q. You told us how you would do it with sensitivity. Now tell us how you would do it with past election results.

A. You were asking an empirical question with respect to sensitivity testing. All right? How would I determine the likelihood that X will happen given Y? Now you're asking me to make a judgment about the constitutionality.

Q. No, I'm asking how you -- same question. I'm not -- read -- read the question. Listen to it carefully. And I understand you want -- you're anxious to advocate your -- you know, your position and your opposition to using the efficiency gap, but that's not what I'm asking you.

A. Okay. Read the question.

(Question read: You told us how you would do

Page 133

it with sensitivity. Now tell us how you would do it with past election results.)

THE WITNESS: Can you define "it"?

BY MR. EARLE:

Q. How you would develop a re -- a robustness check, if you will, a reliability for testing durability, a reliable way of testing durability of the first --

A. I think I answered that with respect to the sensitivity testing.

Q. Right. Now if you were limited to using past election results, how would you -- how would you check that robustness, if you will, of the durability measure?

A. I see. So you're asking if I was not allowed to use a hypothetical -- use a uniform swing or develop hypothetical --

Q. Right.

A. Now I understand it better. I think at a minimum you would want to look at all election results, say given not just the first election after redistricting. You'd want to look at given any election, what is the probability of a deviation in the sign of the -- of the efficiency gap at some other point during the decade or at some other point given some time period that you're interested in,

but you wouldn't want to only highlight those particular -- arbitrarily highlight only those particular elections that occurred in the first decade after redistricting, which I think rather coincidentally don't include any wave elections where I think the least durability is going to be observed.
Q. Okay. Look at now figures 27 and 28. Got them? Those take into account all elections, don't they? Not just the first election after redistricting?
A. So the point estimates here are the proportion of elections that display an efficiency gap at least that large, including all elections in Jackman's data set.
Q. Right.
A. Yes. Sorry, I -- okay. This is the blue dots. Right?
Q. Right.
A. Yes.
Q. That's what it says.
A. The blue dots. That's what I'm defining. Is there another question? I'm sorry.
Q. Read the -- read the -- could you read the question, please?
A. And the red dots show the proportion among all

plans --
Q. Exceeding --
A. -- that have -- exceeding threshold to have an EG with opposite sign.
Q. Right.
A. And this is at some other point during the decade? Again he has a lot of figures here so I forget exactly which figures are showing what.
Q. Well, you're looking at figure 27 and you're looking at figure 28. Do you understand those?
A. I do understand them. Yes.
Q. Okay. So now I'll have the question read to you.
(Question read: Okay. Look at now figures 27 and 28. Got them? Those take into account all elections, don't they? Not just the first election after redistricting?)
THE WITNESS: Yes.
BY MR. EARLE:
Q. That's the question you have before you. What's the answer?
A. Yes.
Q. Okay. And so they're the kind of analysis you would want to conduct, right, that you just described a few moments ago?
A. This is closer to the analysis that I want to

conduct, yes.
Q. How different would you say figures 27 and 28 are from figures 29 and 30?
A. I think they're very different.
Q. How?
A. They show a much greater number of plans having an efficiency gap the opposite sign given a particular threshold.
Q. What exactly are those differences? Can you quantify them?
A. Well, I believe if you look at say a efficiency gap of negative .7, right, it's showing that like 35 percent of plans at that threshold, negative .7, have an efficiency gap of the -- show an efficiency gap of the opposite sign at some point during the decade. I believe I am interpreting this correctly. Again the figure's a little bit complicated so --
Q. Okay. Compare the EG of minus seven percent in figure 27 and figure 29. What are the corresponding blue and red dots?
MR. KEENAN: Object as vague. I don't understand it. If you do, you can answer.
THE WITNESS: So this is only showing the first election? 29?
BY MR. EARLE:

Q. Yeah. 27 is all elections. 29 is the first election.
A. Well, in this case it does look like he's showing that an efficiency gap of negative .7 will have an opposite sign efficiency gap at 25 percent at negative .7.
Q. Okay.
A. Which is a little bit -- seems a little bit inconsistent with his confidence estimates in the later graphs so I'm not completely sure why he's getting those confidence estimates.
Q. What's the figure for 29? What's the figure -- what's it for 29?
A. For figure 29? I think at negative .7 it was showing something like 25 percent. Again I'm just looking at the graph and eyeballing.
Q. Compare 27 and 29 then.
A. So 27 was 35 percent and 36 percent, something like that. 29 was 24 percent, 25 percent, something like that.
Q. Okay. So that's the entirety of the gap that comes from considering just the first election, right?
A. It looks like figure 9 --
Q. Versus all elections?
A. First election or all elections?

Q. No, first election.
A. Twenty-nine shows the first election.
Q. Uh-huh.
A. And it's showing that at a point -- an efficiency gap observed in the first election of negative .7 there is about a 24 percent chance that it will change in sign at some point during the next decade. Again I believe I'm interpreting this correctly.
Q. Okay. All right. Buried the theme a little bit here.
The second sentence of the first paragraph on page 16 of your report, when, quote -- you write, when measuring the bias in a map from an academic standpoint, imputing vote share in unopposed races seems entirely appropriate as do the specific methods used in both reports to make these imputations, close quote. All right.
Did I read that accurately?
A. Yes.
Q. And that's your position, right, that Jackman imputed vote share in unopposed races in an entirely appropriate method?
A. Yes, I do not have any objection to the imputation decisions in the Jackman report.
Q. Do you know what Sean Trende said about that?

A. No, I don't.
Q. Okay. Haven't you used presidential election results in your work to measure districts' underlying partisanship?
A. Yes.
Q. And don't these assume all districts are contested and there's no incumbent when you do it?
A. Yes, but I am not doing it to predict future election results.
Q. Okay. Didn't Wisconsin's own redistricting advisor, Keith Gaddie, assume no incumbents in coming up with his predictions for Act 43?
A. The only knowledge I have of this is what was written in the plaintiffs' filings.
Q. Well, do you find it curious that you have been hired by the State of Wisconsin, the GAB, to defend the map and criticize the EG, and they didn't provide you with this information?
A. No.
Q. If the state in formulating the map came up with an analysis, wouldn't you want to see it, that predicted partisan performance?
A. I don't think it's particularly relevant to the opinions that I'm offering.
Q. You think it's irrelevant that the person drawing

the map for the state, the expert hired by the state to help them draw the map came up with a model to predict partisan performance? You think that's irrelevant?
A. Irrelevant to what?
Q. To determining whether or not there's been an intentional gerrymander here.
MR. KEENAN: Object to the extent it calls for a legal conclusion.
THE WITNESS: I don't believe that I have been hired as an expert to determine intent.
MR. EARLE: Okay.
(Exhibit No. 25 marked for identification.)
BY MR. EARLE:
Q. Showing you what's been marked as Exhibit 25, and I will represent to you that this was a memo written by Keith Gaddie dated April 17, 2011 and was produced by him as part of his reliance material, and he provided testimony on behalf of the GAB in the Baldus case. Take a look at it.
A. Okay.
Q. Take a moment to read it.
A. (Witness reading.)
MR. KEENAN: So you're representing that?
MR. EARLE: Yes.

MR. KEENAN: We don't have the actual document?
MR. EARLE: You have it. You guys produced it to us. The State of Wisconsin attorney general representing the GAB produced it to us in -- at the Keith Gaddie deposition when they produced his thumb drive of his --
MR. KEENAN: Is this the actual document that was produced?
MR. EARLE: It's a print of his thumb drive that was produced.
MR. KEENAN: But I'm asking did you copy --
MR. EARLE: This is a printout of what was on --
MR. KEENAN: Of what was on -- like the exact thing?
MR. EARLE: The metadata would show that this was drafted by Keith Gaddie, and his testimony would say that this was drafted by him on April 17, 2011.
MR. KEENAN: Okay. I just wanted to know whether it was an actual document or just like a --
MR. EARLE: No, this was -- and it was the subject of deposition testimony as well.
THE WITNESS: Okay.
BY MR. EARLE:
Q. Okay. So you read the document?

Page 142

A. Yes.
Q. Okay. Now, don't you think it would be useful for you to have the data that the State of Wisconsin used, the authors of the Act 43 used to calculate and predict partisan performance in the remap process as they redistricted?
A. Not necessarily. I don't see how this particular data would be particularly informative to my report.
Q. Okay. Do you think it would be significant to compare the predicted partisanship performance to the actual partisan performance after the passage of the act?
A. Can you repeat the question?
(Question read: Do you think it would be significant to compare the predicted partisanship performance to the actual partisan performance after the passage of the act?)
THE WITNESS: Depends what you mean by "significant." I am not surprised that the election results in 2012 would conform well to predicted partisan performance based on an even baseline of partisan balance. Because that was the actual result in 2012, that you had a result that was very even on the basis of partisan ballots.
So it would certainly not surprise me that the

Page 143

results were very close in that particular election. I don't think that's necessarily informative for future elections.
BY MR. EARLE:
Q. Okay. Do you know -- do you know if Ken Mayer had a variable for incumbency in his report?
A. Well, he does have a variable for whether an incumbent was running in a particular seat.
Q. Okay. You criticized Mayer for assuming that all districts didn't have incumbents, right?
A. In the demonstration plan, the performance in the demonstration plan, I believe I am accurate in his assumption that there are no incumbents -- or there is no incumbency effect in any district.
Q. How do you recommend that Mayer take into account which districts have incumbents?
A. I am not recommending that Mayer necessarily take that into account, but I think the fact that that is not taken into account reflects the possibility -- reflects the plausibility of the expectation that a legislature could draw such a map or would draw such a map in a hypothetical circumstance.
Q. Okay. Well, but if you were to take it into account, how would you recommend he do it in order to meet with your satisfaction?

Page 144

A. If I were asked to --
Q. Your confidence. I'm sorry.
We have to take turns.
A. I'm sorry.
Q. If -- how would you recommend that Ken Mayer take incumbency into account in order to satisfy your confidence in his work?
A. I would not -- I would not recommend anything to take incumbency into account. I would just discount the effectiveness of this particular methodology in general in rebutting a presumption of constitutionality if we're accepting the plaintiffs' test in the first place.
Q. You've criticized him for not taking it into account, right? Am I understanding that properly?
A. I am not --
Q. Am I understanding that properly that you criticize Ken Mayer for not taking incumbency into account?
MR. KEENAN: I object as it calls for speculation as to what Mr. Earle understands.
THE WITNESS: Okay. I object to making the -- drawing the conclusion that this is a plausible map that could have been drawn because there are so many other factors that are different from the time when the map actually had to be drawn. The amount of

Page 145

knowledge that he both uses and does not use is so much different from that which the legislature knew and didn't know at the time when they had to draw the map.
BY MR. EARLE:
Q. So it would be more realistic to take incumbency into account then, right?
A. It would be realistic if we were predicting what type of map a legislature will draw.
Q. Answer the question I asked. I'll have the court reporter read it to you.
(Question read: So it would be more realistic to take incumbency into account then, right?)
THE WITNESS: I think you would generate more real -- you would probably generate more realistic results for the particular election that you are generating counter factual results for if you did take incumbency into account.
BY MR. EARLE:
Q. How would you do it?
A. Well, I suppose that you know where the incumbents live and you know which incumbents actually did run in the election so you could apply the incumbency advantage in those elections that are in districts where the incumbents live.

Page 146

Q. Okay. Do you criticize -- you criticize Mayer for using 2012 election results to calculate his plan's efficiency gap, right?
A. I -- I don't -- I criticize using that as a conclusion for what the legislature would have expected the efficiency gap in 2012 to be.
Q. Okay. So how would you recommend that Mayer use pre-2012 election results to calculate a plan's efficiency gap?
A. To calculate what the efficiency gap in 2012 would have been?
Q. Right.
A. Again I don't have an objection to that, but from the perspective of the legislature prior to knowing what the election result in 2012 would have been, if he's trying to simulate what they would have guessed the efficiency gap of a plan would be, they would not know the 2012 election result.
Q. If they had to make the prediction, what data would you use?
A. I suppose you would use a range of possible election results judging from the historical range -- drawn from the historical range of observed historical election results.
Q. If you're -- if these approaches that you've now

Page 147

stated had been used by Mayer, would you still have an objection to his choice of data?
A. I'm trying to picture how that data would -- how that would actually be incorporated into his methodology. Certainly I think it was unrealistic to expect a legislature to actually use that sort of methodology. It is -- it is entirely unclear to me if you were to ask a legislature to draw a map that will have a low efficiency gap in the next election, without knowing what the next election would be, I would have no idea how to instruct the legislature to do that.
Q. Okay. So to sum up, if you were given the assignment before 2012, okay, of estimating the efficiency gap that the demonstration plan would exhibit in 2012, exactly how would you do it? I want you to take what you've testified --
A. So before -- so before an election has actually happened --
Q. Right.
A. -- how would I estimate what would -- what would be the efficiency gap in this plan without knowing what the overall election result in a particular year will be?
Q. Yeah.

Page 148

A. Right? I suppose the best estimate for something like that would be to look at the range of historical election results, perhaps overweighting by recent election results, and estimate what the probability of an overall say statewide result would be. And then you would want to, for each probability, weighted probability, you would want to generate the efficiency gap under that election result, and then you would compile the -- right, essentially you would be integrating across that whole range of election results.
Q. Okay. And that would satisfy your concerns?
A. It would satisfy my concerns about what the expected efficiency gap of this particular plan would be in a hypothetical election where you didn't know the result.
Q. Right. It would satisfy your concerns about knowability?
A. Correct.
Q. How close is what Gaddie did to your preferred approach?
A. I don't know what Gaddie did.
MS. GREENWOOD: Exhibit 25.
Q. Exhibit 25.
A. As far as I know, the Wisconsin legislature did not

Page 149

attempt to estimate an efficiency gap so I couldn't tell you.
Q. Well, you have a description of what he did here, right, in Exhibit 25?
A. Yes, he provided data to the legislature, but I don't know how they used this data.
Q. He described what he -- how he organized the data, correct? He says he -- he created a measure of partisanship?
A. Yes. So he is creating -- sorry. Go ahead.
Q. He -- he went through the electoral data for state office and built a partisan score for the assembly districts that was based on a regression analysis of the assembly vote from 2006, 2008, and 2010, and it was based on prior election indicators of future election performance, right?
A. Right.
Q. Okay. Now, how similar is that to what -- what your approach is?
A. It's not very similar.
Q. How -- how is that different?
A. It sounds like what Gaddie is doing is he's determining a single partisanship score for each sub unit, whatever the sub units are. This is the district, for each hypothetical district, a single

Page 150

partisanship score, but he doesn't associate that single partisanship score with a probability of being -- with a probability of actually observing a specific election outcome in a future election.
So he's looking at -- he's looking at districts relative to each other. He's not looking at the possibility of variation within a district across time.
Q. Okay.
(Break taken 1:51 p.m. to 1:58 p.m.)
BY MR. EARLE:
Q. So you say that Ken Mayer's way is not the way to evaluate the propensity of a state's underlying geography to generate bias, right?
A. Well, showing that you could design one hypothetical plan that would show a particular efficiency gap does not demonstrate the underlying propensity of a state to show an efficiency gap or whatever -- however you want to measure the bias.
Q. And in your report you suggest three ways that that might be approached?
A. Okay.
Q. Right?
A. Yes.
Q. Yes?

Page 151

A. Yes.
Q. Okay. And one of those is comparing the bias observed in Wisconsin to other comparable states during the same period, correct? That's at page 18 of your --
A. Yes.
Q. -- report? We have to -- you have to wait until I finish before you say yes. Okay. So the answer is yes?
A. Yes.
Q. Okay. What other comparable states during that same period are you referring to?
A. Well, no particular -- no specific states in particular, but I would say states that are similar to Wisconsin hypothetically on a range of possible factors that, you know, I look at in some of the research that I've done, like urbanization or like underlying partisan propensity or state size, right? So you wouldn't want to necessarily compare it to a very small or very large state.
Q. So you didn't have any states in mind when you wrote that in your report on page 18?
A. Well, I think I probably would say I had states in mind that were of similar size to Wisconsin, but again it was not referring specifically to any particular states.

Page 152

Q. What states of similar size did you have in mind?
A. Hypothetically perhaps something like Michigan or Missouri.
Q. Did you carry out any quantitative comparison between Wisconsin and any other states?
A. For this report, no.
Q. Okay. Do you have a methodology in mind as to how that comparison would be executed?
A. Well, in the article that you referred to that I had published earlier in the deposition, I do compare the bias that it's generated in different states. Again for congressional maps of course I'm only looking at a couple of election cycles and with a very simplified model. So that's the idea that I have in mind.
Q. Is Pennsylvania one of those?
A. Yes, I think it's fair to say Pennsylvania would probably be a fairly comparable state.
Q. You think that the dynamic, the geo political dynamic of the Philadelphia metropolitan area is similar to that of the Milwaukee metropolitan area? Let me rephrase that.
Do you think that the geographic clustering of partisans in the Philadelphia metropolitan area is

Page 153

analogous to the geographic clustering of partisans in the Milwaukee metropolitan area?
A. This is an empirical question which I have not done any specific measurements for. I would tend to think that there are probably similarities based on my background knowledge.
Q. Is this -- would you be applying an eyeball test to that?
A. Eyeball test. I suppose that's fair.
Q. What?
A. I suppose that's a fair characterization.
Q. Would you ever rely on an eyeball test?
A. Occasionally it's probably sufficient.
Q. You think it's sufficient for providing an opinion to a court?
A. Well, if I see someone shoot another person, I'm seeing that with my eyeballs and I would testify to that in court so that would be an eyeball test.
Q. But you're not here as a witness to a murder. You're here as an expert providing the court or trying to provide the court with expert opinion.
A. Yes.
Q. Presumably grounded in standards that govern your profession, correct? An eyeball test meet those standards?

Page 154

A. Well, probably not, but this test has not been -- I mean it's not like there's a better test that's been done by any of the plaintiffs' experts on this point.
Q. The third way to evaluate the propensity of a state's underlying geography to generate bias that you mention in your report on page 18 is simulating non-partisan districts?
A. Yes.
Q. You have that in mind?
A. Yes.
Q. What do you mean by "non-partisan districts"?
A. I am -- I think there are a number of hypothetical ways that someone could simulate hypothetical districts. The one that I know most prominently and which I mention in the report is the Chen and Rodden method of randomly selecting, randomly generating districts for a variety of states based on certain standards of compactness and continuity.
Q. Can you identify any flaws with the Chen and Rodden methodology?
A. Well, for instance, I know they don't take into account like Voting Rights Act considerations.
Q. Okay. That's one. Do they comply with respect for political subdivisions?

Page 155

A. Well, in the way that I believe you're thinking, no.
Q. Okay. And do they respect communities of interests?
A. Not deliberately so.
Q. And I would represent to you that the Wisconsin constitutional -- Constitution Article IV, section 4 says that redistricting districts are to be bound by county, precinct, town, or ward lines to consist of contiguous territory and to be in as compact a form as practicable. Does that approach comply with that?
A. I think for the most part it does. They do -- they do respect precinct lines, or I guess they would be ward lines. In Wisconsin -- they don't actually simulate Wisconsin specifically in their article. They also do try -- they certainly try to respect contiguity, and I believe they try to respect compactness. Again they have a number of ways that they program this, which might vary the amount of weight that they give that sort of standard, but they do definitely consider those factors.
Q. But not political subdivisions, correct?
A. Beyond the precinct, or I suppose in Wisconsin ward level, I believe not.
Q. Okay. And Chen and Rodden simulated plans aren't a random sample of the whole universe of possible

Page 156

plans, are they?
A. They are a sample of the possible plans given their -- the methodology that they have programmed, which of course does respect contiguity and compactness and precinct lines.
Q. Haven't Chen and Rodden been criticized for not coming up with a random sample even of that universe?
A. Are you referring to something specific? I am not sure. I am not aware of the criticism that you're referring to.
Q. How about any work by any political -- Princeton political scientists?
A. Are you referring to the Does Gerrymandering Cause Polarization work that does other random sampling?
Q. No. All right. Chen and Rodden haven't simulated any maps for Wisconsin, have they?
A. Certainly not in any published work.
Q. Okay. And Chen and Rodden haven't simulated any maps using state legislative election results, have they?
A. Well, Chen and Rodden do not use -- Chen and Rodden simulate maps over a variety of number of possible districts. All right? So they're not just specifically simulating this is the maps that will

Page 157

be generated by -- in Florida given that Florida has 25 or 27 congressional districts. They look at what would be the bias observed if we allot 100 districts to Florida or 50 districts to Florida or 200 districts to Florida. So they measure at each number of districts what the average bias would be. So I would say they do set out to simulate both the congressional and the state legislative electoral environment across a variety of a number of possible seats.
Q. But they only use presidential election results, don't they?
A. You mean in terms of determining the baseline partisanship?
Q. Yes.
A. I believe that's true. Yes.
Q. Okay. Have you tried to quantify what percent of the overall pro republican trend you talk about in the EG is due to greater republican control of the redistricting process versus the partisan political geography of the state?
A. I don't believe that I specifically have tried to quantify that.
Q. You have not?
A. No.

Q. Okay. Are there any studies to your knowledge that do that?
A. That use efficiency gap in particular? No, not as far as I know.
Q. Are there any studies that try to tease out partisan control of the legislative process as opposed to partisan geographic clustering as the basis or the contribution to bias?
A. Well, of course my article does that in a fairly simplistic way with respect to the congressional election results in 2012, 2014.
Q. You have no idea what the relative contribution to republican bias in Wisconsin is as a result of political geography, do you?
A. I have not generated a specific estimate of that.
Q. Okay. And you're not going to be providing an opinion about that in the course of this case; isn't that right?
A. I have not been asked to provide an opinion about that specifically.
Q. It's not in your report, right?
A. It is not in my report. No.
Q. And, therefore, it will not enter this case, correct?
A. I don't know if I'm -- at some future point I'm

allowed to provide another report or something like that.
Q. It's true that republicans controlled many more state legislatures in 2010 and 2000 cycles than in previous cycles; isn't that true?
A. Yes. Well, than in at least immediately previous cycles. Yes.
Q. Are you aware of any studies on the trends in partisan clustering over time?
A. Studies on the trends in partisan clustering over time. There are certainly studies on the bias over time in election results.
Q. Uh-huh. The question was partisan clustering.
A. There are studies about the way that people are increasingly identify -- or increasingly correlating where they live and what their partisanship is.
Q. What are the studies?
A. It was a book by Levendusky on partisan sorting. I mean I'd have to get back to you off the top of my head.
Q. Are you aware of Glazer's work in finding --
A. Oh, well, this isn't -- can you -- I believe I am somewhat aware of this, yes, but if you can be more specific.
Q. Well, I asked you if you can identify any studies

that found an increase in partisan clustering over time, right? And you didn't identify any studies, right?
A. Okay.
Q. Is that correct?
A. Yes.
Q. Okay. And I asked then you if you're aware of Glazer's work finding that there has not been an increase in partisan clustering?
A. Okay, I'm not --
MR. KEENAN: You did not ask him that.
BY MR. EARLE:
Q. I'm asking you that now then.
A. I'm aware of work by Glazer. I'm not particularly aware of that study.
Q. In Wisconsin what evidence do you have that there's a concentration of democratic voters in compact urban areas in Wisconsin?
A. Well, in my report I did do the analysis at the ward level that shows that there are a lot more wards that are concentrated with democrats than heavily concentrated with republicans.
Q. Do you translate that into districts anywhere?
A. Do I do an analysis of districts?
Q. No. Do you translate that into comparative

concentration in districts?
A. Well, what districts would you be referring to?
Q. Wisconsin legislative districts.
A. Do I look at the concentration of voters in the Wisconsin legislative districts? This is what you're litigating over. That's the result of intentional districting as opposed -- that's not the result of like -- I'm a little bit confused by your question.
Q. I'll rephrase. I'll come at it a different way. Go to figure 1.
A. Yes.
Q. Have you got it in front of you?
A. Yes, I do.
Q. Okay. Let me open up to it. Tell us what it shows.
A. It shows the number of wards in Wisconsin and the share of the population that's -- or the share of the voting population that those wards consist of as reflected in the 2012 voting data that would have a particular baseline partisanship or what I would predict would be a particular share of the vote in a 50/50 statewide election.
Q. The number of wards data doesn't take into account the population of each ward, does it?
A. The number of wards data does not. The share of

Page 162

population data does.
Q. Okay. Let's compare the -- the heights of the bars here in the 40 to 50 percent range.
A. Okay.
Q. The democrat and the 60 -- the 50 to 60 democratic range. Do you see those two there?
A. Yes. I do see them.
Q. Okay. And first of all, what are those heights? Compare them. Tell me what they would represent.
A. The fact that the bars in the 40 to 50 percent range are higher than the bars in the 50 to 60 percent range suggest that there are more wards that marginally lean republican than there are wards that marginally lean democratic.
Q. Okay. And what are the heights? What's the difference?
A. In which bars? The blue bars or the --
Q. The red bars.
A. In the red bars. Well, it looks like it's about 27 percent in the case of these lean republican districts and about 22 percent in the case of the lean democratic.
Q. Is that a significant difference?
A. I think it is a -- I think it is a substantively significant difference. I would also include the

Page 163

bars that are in other places on the graph showing a substantively significant difference.
Q. That's not the question.
MR. EARLE: Read the question back.
(Question read: Is that a significant difference?)
THE WITNESS: If you're speaking of statistical significance, that doesn't have a particular meaning in this case because I'm not drawing from the sample. This is the entire universe of Wisconsin wards. So it is a difference.
BY MR. EARLE:
Q. Do you think that it's a substantively large difference?
A. Yes, I think it's a substantively large difference.
Q. Three percent is a substantively large difference?
A. I think it's a little more like four or five percent, but --
Q. Why don't you give us a precise difference.
A. Well, okay. I don't have the figures. It's probably about four percent. Right?
Q. Okay.
A. I think that is -- I think that is a fairly large difference if you're talking about bias.
Q. What is the height of just the 40 or 50 dem column?

Page 164

A. This is the red bar?
Q. Yeah.
A. It looks like it's about 26 percent.
Q. That's a pretty high number. It shows a republican skew, right?
A. Yes.
Q. Uh-huh. This has already been marked as Exhibit No. 1 in the deposition, the Ken Mayer report. Page 41. Compare the height of the Mayer bar at 50 to 60 republican.
A. I should mention this includes an annex that I did not receive.
Q. It's a part of the report itself on page 41.
A. Okay.
Q. It's not an annex.
A. Okay.
Q. You did see this, right?
A. Yes.
Q. Okay. Can you compare that? How many districts?
A. Compare it in what way?
Q. Well, I'm going to ask you a question. How many districts are in the 50 to 60 percent range in the -- in the chart --
A. This is --
Q. -- on figure 12?

Page 165

A. So this is -- the 55 percent bar?
Q. Fifty all the way to 60. So --
A. The 55 percent bar is 17, the 60 percent bar is 25. Is that what you're asking?
Q. Yeah. What's the sum of that? How many districts is that?
A. Forty-two.
Q. And that's --
A. Forty-two percent of districts. Is that -- so I guess out of 99 so it's pretty close, right?
Q. Uh-huh. That's a significant difference, right?
A. That is a large number of districts I suppose.
Q. So 42 percent of the districts. How does that percentage compare to the share of wards in the 50 to 60 percent range on your chart?
A. Well, it's somewhat larger.
Q. In the 50 -- 50 to 60 percent republican bar?
A. The number of districts in that range is somewhat larger than the number of wards in that range as a percentage of population in Wisconsin. Yes.
Q. Right. It's fair to say that the district distribution under Act 43 does not look like the ward distribution on your chart; isn't that right?
A. I think there are -- there are differences. I think that's fair to say.

Page 166

Q. And those are significant differences in terms of the -- of this case; isn't that right?
MR. KEENAN: Object as vague.
THE WITNESS: I think there are noticeable differences.
BY MR. EARLE:
Q. Well, it's fair to say that -- that the district distribution is substantially more skewed in the republican direction; isn't that correct?
A. It would depend on what you mean by substantially, but yes, it is -- there are a greater percentage of districts in that bin than there are wards. That's true.
Q. Okay. Well, how -- by how much does it skew in that direction?
A. Well, it looks like it's about 15 percent.
Q. Okay. What's that calculation based on?
A. The percentage of districts in that bin minus the percentage of wards in that -- minus the share of population that lives in wards in that bin.
Q. That's much more significant than the three or four percent difference you talked about earlier based on your chart; isn't that true?
A. Well, I think it's imprecise to say significant in this context. It's larger.

Page 167

Q. It's a lot larger, isn't it?
A. I mean a 15 seat difference is a substantial difference.
Q. Okay. How many districts are in the 40 to 50 percent republican bucket?
A. Isn't that what you just asked? Oh, 40 to 50? 17.
Q. How many?
A. Seventeen.
Q. And what's the difference within the 50 to 60 percent republican bucket?
A. Difference from what?
Q. From 40 to 50 percent to 50 to 60 percent.
A. This is in the Mayer graph?
Q. Yeah. Figure 12.
A. Well, it would be about 25.
Q. Twenty-five. 25 percent, correct?
A. Twenty-five districts or I guess that's about 25 percent.
Q. And how much larger than the four percent in your chart is that?
A. Well, it's about 20 percent larger.
Q. Okay. So your chart dramatically distorts the practical political reality represented by figure 12 in Ken Mayer's report, correct?
A. My chart shows the underlying geography of

Page 168

Wisconsin. It's not a representation of the Wisconsin district map. I'm not sure what you mean by "distorts." This does not purport to be a representation of the Wisconsin district map.
Q. Well, is your chart supposedly a representation of partisan geography that supports your thesis?
A. I think the chart is -- supports the thesis that there is republican bias in the partisan geography of Wisconsin.
MR. EARLE: Can you -- I'm sorry, can you repeat that?
(Answer read: I think the chart is -- supports the thesis that there is republican bias in the partisan geography of Wisconsin.)
BY MR. EARLE:
Q. To the extent that that thesis is predicated on your chart is inconsistent with what is shown in Ken Mayer's chart, figure number 12 on page 41 of the report; isn't that correct. Let me withdraw that question. I'll phrase it this way.
Wisconsin's underlying geography is not accurately reflected in the current districts of Act 43; isn't that true?
MR. KEENAN: Object as vague.
THE WITNESS: It's --

Page 169

BY MR. EARLE:
Q. I mean at least --
A. The districts -- okay, the distribution of partisanship in the districts in Wisconsin is not identical to the distribution of partisanship of the wards. I agree with that. Yes.
Q. Okay. And -- and using -- the district distribution is much more skewed in the republican direction than the ward distribution; isn't that true?
A. It is more skewed in the republican direction.
Q. Significantly more skewed; isn't that true?
A. Depends on what you mean by "significantly." I --
Q. Well, what is the percentage difference?
A. It depends on how you define it, but it is noticeably more skewed.
MS. GREENWOOD: Can we take a break?
MR. EARLE: Yeah. Take a break.
(Break taken 2:24 p.m. to 2:38 p.m.)
MR. EARLE: On the record. Mark this as the next exhibit.
(Exhibit No. 26 marked for identification.)
BY MR. EARLE:
Q. Showing you what's been marked as Exhibit 26.
A. This is my website.
Q. Exhibit 26. This is a printout that we made off of

your academic home page, and it's the section that's captioned "Media"?
A. **Okay.**
Q. Do you recognize it?
A. **I do.**
Q. Now, you listed these items on there because you consider them to be relevant representations of your work in the area of political science and elections, correct?
A. **Yes.**
Q. In particular gerrymandering, correct?
A. **Yes.**
Q. And on there you have a citation to The Monkey Cage blog, correct?
A. **Yes.**
Q. And didn't -- you also have posted on another blog. What's it called?
A. **Oh, okay. So I think Wonkblog. I think what the Wonkblog -- I think Wonkblog just posted linking to one of my Monkey Cage articles. It's part of the Washington Post website. It's just two sections.**
Q. Okay. So off -- we had an off-the-record discussion relative to the subpoena duces tecum in which you indicated that you would later produce to us printouts of the -- all the Monkey Cage material

that you've offered?
A. **Yes.**
Q. Authored, correct?
A. **Yes.**
Q. And you'll provide that to counsel and counsel will provide it to us at your convenience --
A. **Sure.**
Q. -- after this deposition.
A. **Yes.**
Q. Okay. Now, did you post anything on the Wonkblog yourself, any -- any entries, any commentary?
A. **No, I believe that the Wonkblog entry I'm referring to is simply a link to one of my Monkey Cage posts.**
Q. Okay. The -- okay. Your Monkey Cage posts, are there comments that you've placed on the Monkey Cage website in response to other posts by other people? In other words, have you participated in discussion on the Monkey Cage web page regarding the postings of other authors regarding redistricting?
A. **Oh, have I made comments about other articles on the Monkey Cage website?**
Q. Right.
A. **I think I might have at the old Monkey Cage website before it was associated with the Washington Post.**
Q. Okay. Would you include those in your production to

counsel? Because I believe they are -- it is my opinion that they are responsive to the subpoena.
A. **Okay. I will try to find them. I'm not completely certain that they would be saved, but I don't see why they wouldn't be. I can look for them.**
Q. Okay. And we'll mark -- and we'll ask the court reporter to mark this section as a document request to -- as a -- I guess a deferred compliance with the subpoena that we'll get later. Okay?
A. **Okay.**
Q. Can you do that in the next week?
A. **Yes.**
Q. Okay. Okay. Then we'd like to actually ideally get it before our rebuttal report is due.
MR. KEENAN: Yeah, can you just look this week then?
THE WITNESS: Yeah.
MR. EARLE: Okay. Good.
THE WITNESS: That's fine.
MR. EARLE: All right. Great.
BY MR. EARLE:
Q. Now, as I look at Exhibit 26, I'm assuming that what you list here is material that you consider to be credible, right?
A. **Yes.**

Q. Okay. And -- and your -- you list amongst these things a caption that's called What is Gerrymandering? Discussion -- and you cite it as "discussion of my research in gerrymandering primer on Vox.com" dated April of 2014. Are you familiar with that?
A. **I am familiar with it. I don't exactly recall the details of what the -- the whole article was.**
Q. Okay. Well, you -- you're placing it as an example of your work, right?
A. **Yes.**
Q. On --
A. **Yes. If I recall correctly the way that Vox.com formats this sort of article is it's a series of like cards. It's almost like a slide show, and my research is discussed on one slide in the slide show. I don't remember the content of what all of the slides in this -- what they call a card stack referred to in the gerrymandering primer.**
Q. Well, do you consider the card stack to be an accurate description of the substance that's within it?
A. **Well, I would imagine the card stack includes opinions from many -- from both journalists and politicians and various political scientists with**

various opinions on subjects related to gerrymandering, some of which I would consider reliable and some of which I perhaps would not.
Q. Okay. Well, we'll figure that out then. Mark this as Exhibit 27.
(Exhibit No. 27 marked for identification.)
THE WITNESS: Can I get a copy of this?
MR. EARLE: Yes, I'm sorry. Here you go. Now, I just selected various pages.
MS. GREENWOOD: This is just card one.
MR. EARLE: This is just card one. Okay.
MS. GREENWOOD: I marked them separately.
BY MR. EARLE:
Q. This is card one, and it's captioned What is Gerrymandering? And it reads, in the U.S., every state elects a certain number of people to the House of Representatives, a number that's based on the U -- on the census count of the state's population, Pennsylvania, for instance, elects 18 House members so Pennsylvania has to be divided into 18 congressional districts with roughly equal populations. In most U.S. states this process is controlled by the majority party in the state legislature.
Did I read that correctly?

A. Yes.
Q. Partisan gerrymandering occurs when this map-drawing process is intentionally used to benefit a political party -- a particular political party to help that party win more seats in the legislature or more easily protect the ones it has. The goal is to create many districts that will elect members of one party and only a few that will elect members of the opposite party. You can see Pennsylvania's congressional district map below. And there's a portrayal of the map. Correct?
Do you have any substantive disagreement with the two paragraphs that I just read under the caption What is Gerrymandering?
A. Well, I assume you're referring to their ostensible definition of partisan gerrymandering which I think I've already clarified is not the definition that I use in coding partisan gerrymanders in my research.
Q. But do you think that from the perspective of a political scientist studying the process of gerrymandering, that that's an inaccurate definition?
A. I don't think there is a uniform definition among political scientists of what they would call partisan gerrymandering or how they would code that

in a data set or use that in casual conversation.
Q. Okay. Would you like to distance yourself from the definition that's used here in What is Gerrymandering, Exhibit 27?
A. I wouldn't say I would want to explicitly distance myself in that here it's being used in a very casual way. People use -- people use terms to refer to many different things.
Q. Okay. At the back of the -- further on the other side of the maps on the last page of Exhibit 27, it reads, "Gerrymandering can affect any legislative body that has to have districts drawn, which includes both the U.S. House of Representatives and every state legislature. And since political power is at stake, fights over redistricting are often quite intense."
Do you disagree with anything that I just read?
A. I don't see anything that I would disagree with there.
Q. Okay.
MS. GREENWOOD: Next? This is 28.
(Exhibit No. 28 marked for identification.)
BY MR. EARLE:
Q. Showing you what's been marked as Exhibit 28. This is another page.

A. Okay.
Q. Okay. And it's captioned How Does Gerrymandering Work? Okay. Would you read into the record the -- the first paragraph?
A. "The idea behind gerrymandering is pretty simple. You pack your opponent's supporters together into very few districts. Then you make other districts relatively more balanced, but you place enough of your supporters in most of them to give you an advantage. The hoped-for result is that your party loses a few districts hugely, yet wins a majority of districts comfortably. All partisan gerrymanders boil down to that basic concept, Eric McGhee of the Public Policy Institute of California told me in 2014."
Q. Do you have any argument with Eric McGhee's quote on this Exhibit 28?
A. I would agree that in general partisan bodies who are drawing political maps tend to use the technique of packing opponent supporters together into very few districts.
Q. And they also use the technique of cracking?
A. And making the other districts relatively more balanced but place enough supporters in most -- I think that is a fair description of the way that

Page 178

most partisan gerrymanders operate. Yes.
Q. So you would agree that the efficiency gap is a tally of all the cracking and all the packing that goes on in a given plan?
A. No.
Q. What is the basis of your disagreement with that?
A. I believe that efficiency gap -- that gap -- the efficiency gaps can be generated from many sources, of which packing and cracking could potentially be one, but there are many other sources I think as I observe in my report where you could observe a large efficiency gap that would not result from packing and cracking, nor would packing and cracking necessarily generate a high efficiency gap depending on the overall electoral environment.
Q. List the other factors.
A. Well, I think as I demonstrated, a desire to create competitive elections in a balanced way would not be evidence of packing and cracking, yet in certain electoral environments that would display a high efficiency gap.
My example of proportional representation, right, in certain electoral environments, for instance, favoring strongly one party would display an efficiency gap against that party because

Page 179

proportionate representation -- it would be proportionate rather than the hyper-proportionate representation that would be described by the neutral efficiency gap. And also you could, for instance, have an electoral environment in which the districts that are relatively more balanced in a packing and cracking scenario, you'd have an electoral environment in which the districts that are relatively more balanced, but slightly advantageous towards the gerrymandering party in a 50/50 baseline scenario would instead all be won by the out party, the non-gerrymandering party in a wave election favoring them. That would generate an efficiency gap in favor of the party that did not gerrymander the map.
Q. Did you -- are you aware of any facts in Wisconsin that would indicate that what you just described was a part of the gerrymandering process here in Wisconsin?
A. I am not aware of any facts that considered, for instance, competitiveness or proportional representation.
Q. Answer the question I asked you, please.
MR. EARLE: Can you read it -- read the question to him, please?

Page 180

(Question read: Did you -- are you aware of any facts in Wisconsin that would indicate that what you just described was a part of the gerrymandering process here in Wisconsin?)
THE WITNESS: I think it's possible that districts drawn in Wisconsin, you could observe a reverse efficiency gap if the electoral environment was strongly favoring the democrats enough. That would not be evidence of packing and cracking on the democratic side. I think again this is a hypothetical.
BY MR. EARLE:
Q. Okay. Let's go to the next one. If we define -- one more question. If we define packing as districts that one party wins by a large margin and cracking as districts that the one party loses -- that party loses by a relatively smaller margin in a particular election, can anything other than packing or cracking result in a large efficiency gap?
A. I think I listed various other factors that could result in a larger efficiency gap in the previous answer.
Q. Wouldn't those other factors simply result in packing and cracking?
A. What I'm suggesting is if you were to draw a map in

Page 181

which every district was competitive, you could observe a large efficiency gap in the case of a relatively mild wave election in which one party won a super majority of seats. That would not be the result of what I would consider packing and cracking.
Q. The fact that it's a small margin despite a wave election means it's the result of cracking, isn't it?
A. No, no, I'm sorry. Let's say you have democrats -- there was a wave election in which democrats won 55 percent of the statewide vote. If all of the seats were drawn to be roughly 50/50 or 51/49 or 52/48 --
Q. Some of the seats are cracked or I mean some of the seats are packed.
A. You're saying factors outside of packing and cracking?
Q. Right.
A. Your question was factors outside of packing and cracking. I'm suggesting factors outside of packing and cracking that could generate a large efficiency gap. Now, if you're adding on seats that are intentionally packed or cracked, then that seems to be -- I'm not understanding the premise of your question now.

Q. Yeah. Yeah. Okay. We'll -- we'll end our quibble over the definition of packing and cracking.
A. Okay.
Q. Okay. Let's go to Exhibit 29.
(Exhibit No. 29 marked for identification.)
Q. Showing you what's been marked as Exhibit 29.
A. Yes.
Q. This is the section or the card in the Vox.com web page that contains your quotes. Okay?
A. Okay.
Q. I draw your attention to category number 1, Geography as a GOP Bias.
A. Okay.
Q. Okay. You're quoted as saying -- and I'll read the quote in the -- it says, "And Nicholas Goedert" --
A. It's go Goedert.
Q. -- "Goedert, a post-doc fellow at Washington University in St. Louis, wrote a paper that found geography was more important in explaining the 2012 House results than gerrymandering was."
Did I read that correctly?
A. Yes.
Q. Okay. And a little bit further down, it says, quote, but the more you account for incumbency, the less the intentional partisan gerrymandering is

going to matter, close quote, Goedert told me in 2014.
Is that an accurate quote?
A. I'm sure it's an accurate quote. Yes.
Q. Now, the most important one, the one -- principal one I wanted to read to you is the first sentence of section 3. And it says Partisan Gerrymandering.
And it says, "Finally, all of the analysts quoted above," and that includes you, "agree that there truly were some egregious partisan gerrymanders that affected 2012 results. For instance, Ohio, Pennsylvania, North Carolina, and Virginia. Republican candidates won 49 percent and 53 percent of the House vote in each state, yet each state's congressional delegation ended up about 70 percent republican. States such as Michigan and Florida on the GOP side and Illinois and Maryland on the" republican -- "on the democrat side are also frequently pointed to as being gerrymandered. But any analysts blaming the democrats' failure to take the House solely on gerrymandering is probably too simplistic."
Did I read that accurately?
A. Yes.
Q. Is there anything I just read that you disagree

with?
A. I don't think I ever explicitly characterized any map as an egregious partisan gerrymander. If they are referring to me as among all of analysts, they're probably just referring to the data that was provided in my article. I certainly don't think I gave a quote suggesting that, you know, there are truly some egregious partisan gerrymanders that affected 2012 results. I certainly would not have characterized a partisan gerrymander as egregious.
Q. You don't think there were any egregious gerrymanders affecting 2012 results; is that what your testimony is?
A. That's not what my testimony is. My testimony is that I was -- that it would not be correct to assume that I was an analyst quoted agreeing that there were egregious partisan gerrymanders.
Q. I gotcha. A follow-up question?
A. All right.
Q. Is it your position that there were no egregious partisan gerrymandering affecting the 2012 congressional election results?
A. Well, I would -- I think the term egregious is asking for a personal opinion rather than an expert opinion drawn from political science. I think that

I impulsively believe that there were some republican gerrymanders that were egregious in the sense that I am personally a democrat and I would like to see democrats elected, and the fact that the democrats were not able to be elected in the 2012 results in some of these elections, was -- you know, made me unhappy.
Q. Now, in your view all political discourse that occurs about the existence or non-existence -- well, about the existence of supposedly egregious gerrymanders is they're all wrong, they're -- it's just political geography?
A. No, no, no, no. I don't -- I don't think that -- it is certainly not my position that the bias generated in the maps in 2012 was entirely the -- was entirely the effect of political geography. I certainly think there was an intentional gerrymander on these maps, yes.
Q. I'm going into a slightly different subject here.
When examining partisan trends within a state, do you agree that the -- that the optimal geographic unit is one that has roughly the same population?
A. Same population as -- you mean --
Q. Each other.
A. -- across time you should be analyzing population

Page 186

units that have the same pop -- you should be examining units that have the same population.
Q. Or similar populations?
A. I think that's probably a fair characterization.
Q. Do you agree that counties vary dramatically in their populations?
A. Yes.
Q. Okay. And do you know how large the difference between Wisconsin's most populous and least populous counties are?
A. Not off the top of my head, no.
Q. Do you think that a county-level version of your figure 1 would be useful as your ward level version?
A. Would be as useful as my ward level version?
Q. Yeah.
A. I think it would probably not be as useful because many counties are going to be very very large, and thus, if you're looking at -- you might not -- because I assume the largest counties in Wisconsin are going to be urban counties, that you might not entirely characterize -- accurately characterize the clustering of population that might occur within those counties, within different areas of those very large counties that might be captured by analysis of ward level data.

Page 187

Q. So you would recommend against carrying out an analysis of partisanship using Wisconsin's counties as the unit for analysis?
A. I wouldn't necessarily recommend against it.
Q. So you think --
A. I -- I think all analysis can be helpful. Some can be more helpful than others.
Q. So notwithstanding the variance in population?
A. I think it would still be informative. I think it might not be quite as informative as the analysis at the ward level.
Q. Would it be as reliable empirically?
A. I don't know what you mean by "reliable."
Q. Accurate, to use another --
A. You mean would it be an accurate characterization of trends in partisanship?
Q. Right.
A. I think it would be slightly less accurate than the analysis at the ward level.
Q. Would you trust the conclusions from the county level analysis as at the same level you would trust an analysis based on ward level?
A. I would trust them slightly less.
Q. Slightly less?
A. It would depend on the exact form of the analysis.

Page 188

Q. There are various ways to measure the geographic clustering of democratic and republican voters, right?
A. Yes, I would imagine so.
Q. If we divide wards into democratic and republican leaning groups and then calculate the average margin of victory of the top-of-the-ticket candidate for each group, what does that tell us, if anything, about the extent of geographic clustering?
A. Well, it would -- I think it would pretty much tell you the same thing that my figure 1 told you. It would tell you -- it would tell you the distribution of partisanship among the wards.
Q. Okay. Can you identify any peer-reviewed literature that has studied geographic clustering the way I just described my prior question?
A. Not off the top of my head.
Q. Would this kind of analysis incorporate any data about wards -- the actual geographic location of wards?
A. The hypothetical analysis that you've just told me?
Q. Yeah.
A. It doesn't sound like it would.
Q. Okay. If we take wards of a certain partisan composition and then calculate how close they are on

Page 189

average to other wards of the same partisan composition, what does that tell us, if anything, about the extent of geographic clustering?
A. Well, it would be informative as to how easy it is to link those wards together in a compact district.
Q. Does this approach tell us which wards are adjacent to a ward of a certain partisan composition?
A. The distance of one ward from another?
Q. I'm sorry, adjacent. I'll rephrase the question.
Does this approach tell us which wards are adjacent to a ward of certain partisan composition?
A. The previous hypothetical analysis that you told me --
Q. Uh-huh.
A. -- would not tell you which wards are adjacent to other wards.
Q. Okay. Could this approach be influenced by the geographic size of the wards?
A. Well, it wouldn't be influenced by the geographic size of the wards. It would be -- the -- I mean it's possible that geographic size of ward could correlate with partisanship in some way, but the analysis itself would not be influenced by the geographic size of the wards.
Q. Would you recommend using the mean instead of the

Page 190

median difference between wards of the same partisan composition?
A. **Difference of what?**
Q. Huh?
A. **Difference between what?**
Q. Between the wards.
A. **The partisan composition of the wards?**
Q. Yeah. Uh-huh.
A. **I don't understand what you mean by the mean or the median in terms of partisan composition.**
Q. Of the distance.
A. **I haven't advocated for using any measure of distance.**
Q. Okay. Yeah, hypothetically I'm saying.
A. **I would advocate for using the mean or the median distance between wards? You mean --**
MR. KEENAN: I'm just going to object as incomplete hypothetical, but --
THE WITNESS: Okay. So you mean some measure of the cen -- like the centroid of the ward compared to the centroid of the other ward as opposed to the distance of all points in a ward compared to all points in another ward that -- that mean distance?
BY MR. EARLE:
Q. Yeah.

Page 191

A. **Yes?**
Q. Yeah. Uh-huh.
A. **I think those would be relatively equivalent. I would think that using the -- the distance between centroids of a ward would probably be sufficient to satisfy any minor differences that they would -- like using the distance between the centroids of the wards seems like a reasonable -- a reasonable method. Using the mean, that's a very -- that's very complicated.**
Q. Well, we're comparing the mean distance between wards of same composition and the median distance between such wards. Which is better, mean or median?
A. **I see. So you're saying -- okay. Now I understand what you're saying.**
Q. Which is better? The question is which is better?
MR. KEENAN: Object again as incomplete hypothetical.
THE WITNESS: Well, the advantage of median in general is that -- is that it -- it doesn't distort your data for outliers, right? So in that sense I can certainly see there being outliers in that sort of data that you wouldn't want to -- again you're talking about a hypothetical that I have not

Page 192

actually given a lot of detailed thought to.
I think that using a median would not be inappropriate there, but again, I'd have to think through this a lot more.
BY MR. EARLE:
Q. Okay. Do you agree that it would be easy to draw a district around wards of the same partisan composition that are geographically distant but adjacent to one another?
A. **Yes, but in general I think it's fairly easy to draw districts in many ways.**
Q. Do you agree that it would be difficult to draw a district around wards of the same partisan composition that are geographically close but not adjacent to each other?
A. **That are geographically close but not adjacent to each other?**
Q. Right.
A. **I believe there are examples of districts that include those sort of wards in many cases so I don't necessarily think it would be difficult.**
Q. Can you identify any peer-reviewed literature that has studied geographic clustering this way?
A. **Not off the top of my head.**
Q. Does this method strike you as an accurate and

Page 193

reliable way to study the geographic clustering of democratic and republican voters in Wisconsin?
A. **I don't know, I'd have to think about it more.**
Q. Well, how would you approach it while you're thinking about it? What would be the criteria that you would contemplate?
A. **Well, I think -- I actually think that the Rodden and Chen methodology is a fairly good one in that it doesn't -- it -- it doesn't really prescribe any particular method for drawing districts other than sort of adjustable parameters for contiguity and compactness.**
So to the extent that they would describe a district as easy to draw if it's easy to randomly generate, maybe you could measure something along those lines. If you're -- like it sounds like the study that you're suggesting is something like is the av -- like are districts -- are districts or wards or counties or something like that of similar political persuasion close to each other? If you're asking like throughout the state, well, I don't know that's a good measure, right?
Philadelphia and Pittsburgh are very far away from each other within the context of Pennsylvania and those might have similar political persuasions.

Page 194

So the fact that Philadelphia is far from Pittsburgh doesn't say anything about the actual concentration of voters within those particular cities. And depending on the size of the district, you could very well draw many districts that would just include packed democrats in Pittsburgh or just include packed democrats in Pennsylvania.

Q. Okay. I'm going to read something to you. All right?

A. **Okay.**

Q. Next, the distance to the nearest neighbor for each ward was calculated for each subset of partisan indices. To visualize this, imagine creating a grid with all of the D plus 1 wards listed both horizontally and vertically, parens, if you prefer an IXJ matrix where both dimensions are defined as including the number of wards, close parens. The distance from the first ward to every other ward is calculated filling in the first row of our grid. The smallest value is noted, which represents the distance from ward 1 to the nearest other ward of similar partisan index. The process then repeats for ward 2, 3, and so forth. At the end, the median of the smallest distances is calculated, which gives us an idea how close the D plus 1 wards are to each

Page 195

other.

I utilized the median rather than the mean here because outlying wards such as Menomonee County exert an undue amount of leverage on averages. Okay. The process is then repeated for D plus 2, D plus 3 and so forth. Okay?

Does that seem like -- are you familiar with any literature that supports that approach?

A. **Not off the top of my head.**

Q. What problems can you think of with this approach?

A. **Off the top of my head I don't see any problems, but again it's a little bit out of context for what it's trying to determine.**

Q. How so?

A. **You haven't determined what it's test -- you haven't told me what it's testing. I will say --**

Q. Does this strike you as a good way to -- to measure the clustering of partisanship?

A. **It does not strike me off the top of my head as an inappropriate way to -- an inappropriate methodology given what you've just told me, but again it's still out of context. Like, for instance, the use of the median as opposed to a mean there sounds totally fine to me. Right? You'd have to let me like inspect it a little more closely and give me more**

Page 196

context to actually come up with a reasonably informed critique.

Q. So if you ran this analysis, what would it tell you?

A. **I've -- you're --**

Q. In terms of your type of clustering.

A. **You're reading me -- without even showing me the like -- what would it tell me?**

Q. Yeah. About geographic clustering.

A. **It sounds like it would tell you for a particular partisan -- for a unit with a particular partisan makeup, it would tell you on average how close the nearest district was that had the same partisan makeup. That sounds to me like what it's telling me.**

Q. Okay. Would it tell you what wards are adjacent to each other for purposes of remapping?

A. **I don't think that in itself would tell me what wards were adjacent to each other.**

Q. Let's go to Exhibit 20.

A. **Do I have Exhibit 20?**

Q. Yeah, it's your article Gerrymandering or Geography?

A. **Okay.**

Q. I believe it's Exhibit 20. You have it in front of you there. It should be in that stack.

A. **Yes. I do.**

Page 197

Q. All right. I'm going to draw your attention to page 4.

A. **Okay.**

Q. And don't you refer to Wisconsin as a republican gerrymander here on page 4?

A. **Yes.**

Q. Okay. Didn't Wisconsin also exhibit a pro republican efficiency gap of 15 percent?

A. **Are you referring to the data in the table here?**

Q. Yeah.

A. **So I -- I am not measuring exactly efficiency gap here. I'm using a slightly different methodology.**

Q. Right.

A. **I would estimate that it's probably fairly close to the efficiency gap. Yes.**

Q. Okay. All right. On page 6.

A. **Yes.**

Q. Okay. In two of your remodels, the more thorough ones, don't you find that democratic gerrymanders result in a bigger advantage than republican gerrymanders?

A. **I wouldn't characterize it that way because the difference between those coefficients is not statistically significant.**

Q. The republican ones are not bigger than the

Page 198

democratic ones, right?
A. True.
Q. Doesn't this suggest that both parties can significantly benefit themselves through gerrymandering regardless of political geography?
A. This suggests that in 2012 both parties did benefit themselves through gerrymander.
Q. Regardless of political geography?
A. Holding constant political geography.
Q. Yeah. Using presidential election results, isn't it true that the pro republican bias under bipartisan and court gerrymanders largely disappears according to your work on page 6?
A. Let me refresh my -- yes. And I think this speaks to the variation that I see in effects of both gerrymandering and geography when the overall election environment is different than it was in the 2012 congressional environment.
Q. Doesn't this suggest that there's no inherent bias in favor of either side when a plan is drawn without partisan intent?
A. It shows that there is not necessarily a bias in favor of one side or another across all possible election results.
Q. Okay. All right. Let's go to Exhibit 21. We'll

Page 199

cruise through your articles real fast here.
A. Okay.
Q. I draw your attention again to page 6 here. Okay. And don't you find that geography on this page, page 6, don't you find that geography produced a bias of only two percent in 2014?
A. That is the estimate that I come up with, yes.
Q. Don't you also find that urbanization doesn't have a statistically significant impact on bias in the -- in your 2014 model on page --
A. Yes.
Q. -- 7?
A. Yes.
Q. Is it right that you determined the bias supposedly due to geography simply by assuming that it's a bias in states with court-drawn or bipartisan maps?
A. In this article. Yes.
Q. As you put it, you, quote, assumed the average bias observed in bipartisan states, parens, seven percent and two percent in 2012 and 2014 is the overall bias due to geography?
A. Can you tell me --
Q. Page 14. I'm sorry, I'm on page 14. Let me -- and I'll start over --
A. Okay.

Page 200

Q. -- to be fair. Okay. Page 14 -- well, that doesn't make any sense.
MR. STEPHANOPOULOS: It's a footnote.
THE WITNESS: Is this footnote 3?
BY MR. EARLE:
Q. Footnote 3. Okay. Good. So the question again, I will reword it. All right.
Is it right that you determined the bias supposedly due to geography simply by assuming it's the bias in states with court-drawn or bipartisan maps?
A. Yes.
Q. Okay. As you put it, you, quote, assumed the average bias observed in bipartisan states in 2014 and two thousand -- let me repeat that.
As you put it you, quote, assumed the average bias observed in bipartisan states is the overall bias due to geography, correct?
A. In the context of this article. Yes.
Q. In coming up with this estimate, it's correct that you don't control for any aspects of the state's demographics, urbanization, or political environment; isn't that right?
A. For this estimate. That's true, yes.
Q. All right. Let's go to Exhibit 23. Draw your

Page 201

attention to page 2. Don't you agree that when parties have complete control of redistricting, they, quote, pack members of the opposed party into a small number of ideologically homogeneous districts creating some safe incumbents and create a large number of districts that favor their own party?
A. Yes, that is how I characterize most partisan gerrymanders or the general operation of partisan gerrymanders.
Q. Okay. I think we're done. Just we -- you owe us the documents.
A. Okay.
Q. And --
MR. STEPHANOPOULOS: Give us two seconds?
MR. KEENAN: Yeah, let's take a short break. I'll even think if I have anything to ask. I may not have anything.
(Break taken 3:24 p.m. to 3:26 p.m.)
MR. EARLE: Do you have anything?
MR. KEENAN: I'm not going to have anything.
MR. EARLE: I think we're done.
MR. KEENAN: We'd like to sign.
MR. EARLE: We've asked for an expedited copy.
(Deposition ended at 3:27 p.m.)

STATE OF WISCONSIN }
} SS:
COUNTY OF WALWORTH }

I, LAURA L. KOLNIK, Registered Professional Reporter and Notary Public in and for the State of Wisconsin, do hereby certify that the foregoing proceedings were taken before me on the ____ day of ______________, 20_____.

That the appearances were as noted initially.

That before said witness testified, he was first duly sworn by me to testify the truth, the whole truth and nothing but the truth relative to said cause.

I further certify that I am neither counsel for, related to, nor employed by any of the parties to the action in which this proceeding was taken; and, further, that I am not a relative or employee of any attorney or counsel employed by the parties hereto, nor financially interested, or otherwise, in the outcome of this action.

That the foregoing proceedings are true and correct as reflected by my original machine shorthand notes taken at said time and place.

Dated this _____ day of ____________, ______

____
LAURA L. KOLNIK, RPR/RMR/CRR
Notary Public
State of Wisconsin
My commission expires
February 23, 2018

Measuring the Compactness of Political Districting Plans

Measuring the Compactness of Political Districting Plans
Author(s): Roland G. Fryer Jr., and Richard Holden
Reviewed work(s):
Source: *Journal of Law and Economics*, Vol. 54, No. 3 (August 2011), pp. 493-535
Published by: The University of Chicago Press for The Booth School of Business of the University of Chicago and The University of Chicago Law School
Stable URL: http://www.jstor.org/stable/10.1086/661511
Accessed: 17/05/2012 03:46

Measuring the Compactness of Political Districting Plans

Roland G. Fryer, Jr. *Harvard University*
Richard Holden *University of New South Wales*

Abstract

We develop a measure of compactness based on the distance between voters within the same district relative to the minimum distance achievable, which we coin the relative proximity index. Any compactness measure that satisfies three desirable properties (anonymity of voters, efficient clustering, and invariance to scale, population density, and number of districts) ranks districting plans identically to our index. We then calculate the relative proximity index for the 106th Congress, which requires us to solve for each state's maximal compactness—a problem that is nondeterministic polynomial-time hard (NP hard). The correlations between our index and the commonly used measures of dispersion and perimeter are −.37 and −.29, respectively. We conclude by estimating seat-vote curves under maximally compact districts for several large states. The fraction of additional seats a party obtains when its average vote increases is significantly greater under maximally compact districting plans relative to the existing plans.

1. Introduction

The architecture of political boundaries is at the heart of the political process in the United States.[1] When preferences over political candidates are sufficiently

We are grateful to Alberto Alesina, Roland Benabou, Rosalind Dixon, Edward Glaeser, Emir Kamenica, Lawrence Katz, Gary King, Glenn Loury, Barry Mazur, Franziska Michor, Peter Michor, David Mumford, Barry Nalebuff, Ariel Pakes, Andrei Shleifer, Andrew Strominger, Jeremy Stein, and seminar participants at Brown University (applied math), Harvard University (labor economics), the National Bureau of Economic Research Summer Institute (law and economics), and the University of Vienna (math) for helpful discussions and suggestions. Shiyang Cao, Alexander Dubbs, Laura Kang, Eric Nielsen, and Andrew Thomas provided excellent research assistance. Financial support was provided by the Alphonse Fletcher Sr. Fellowship. Fryer thanks the Erwin Schrödinger International Institute for Mathematical Physics in Vienna for its hospitality.

[1] Article 1, section 4, of the U.S. Constitution provides that "[t]he Times, Places and Manner of holding Elections for Senators and Representatives shall be prescribed in each State by the Legislature thereof; but the Congress may at any time by Law make or alter such Regulations, except as to the Places of choosing Senators."

[*Journal of Law and Economics*, vol. 54 (August 2011)]

heterogeneous, altering the landscape of political districts can have large effects on the composition of elected officials. Prior to the 2003 Texas redistricting, the congressional delegation comprised 17 Democrats and 15 Republicans; after the 2004 elections there were 11 Democrats and 21 Republicans.[2] Politically and racially motivated districting plans are believed to be a significant reason for the lack of adequate racial representation in state and federal legislatures, and there is a debate as to whether the creation of majority/minority districts to ensure some level of minority representation has led to fewer minority-friendly policies (see Shotts [2002] for an excellent overview and critique).

There are several factors that weigh on the constitutionality of districting plans: (1) equal population (the Supreme Court first established this principle for congressional districts in *Wesberry v. Sanders* (376 U.S. 1 [1964]), (2) contiguity (which is a requirement in 49 state constitutions), and (3) compactness. This last consideration—distinct from the mathematical notion of a finite subcover of a topological space—refers to how oddly shaped a political district is. The Supreme Court has acknowledged the importance of compactness in assessing districting plans for nearly half a century.[3] Yet, despite its importance as a factor in adjudicating gerrymandering claims, the court has made it clear that no manageable standards have emerged (see the judgment of Justice Antonin Scalia in *Vieth v. Jubelirer*, 541 U.S. 267 [2004]). There is no consensus on how to adequately measure compactness.[4]

In this paper, we propose a simple index of compactness based on the average physical distance between voters and show that this index has a number of attractive features. The index is the ratio of distance between voters in the same political district under a given plan and the minimal such distance achievable by any possible districting plan. The greater this ratio, which we call the relative

[2] In the United States, political boundaries are typically redrawn every 10 years, after the decennial census. The 2003 middecade redistricting in Texas is a notable exception. The Supreme Court recently held that this was not unconstitutional in *League of United Latin American Citizens v. Perry*, 548 U.S. 399 (2006).

[3] The apportionment acts of 1842, 1901, and 1911 contained a compactness requirement. In *Davis v. Bandemer* (476 U.S. 173 [1986]), Justices Lewis Powell and John Paul Stephens pointed to compactness as a major determinant of partisan gerrymandering, and Justices Byron White, William Brennan, Harold Blackmun, and Thurgood Marshall cited it as a useful criterion. Nineteen state constitutions still contain a compactness requirement (Barabas and Jerit 2004).

[4] An important argument against the use of compactness as a districting principle is that it may disadvantage certain population subgroups. As Justice Scalia put it in *Vieth v. Jubelirer* (541 U.S. 267, 290), "Consider, for example, a legislature that draws district lines with no objectives in mind except compactness and respect for the lines of political subdivisions. Under that system, political groups that tend to cluster (as is the case with Democratic voters in cities) would be systematically affected by what might be called a natural packing effect. See *Bandemer*, 478 U.S. 159 (O'Connor, J., concurring in judgment)." First, the courts use compactness as one of several criteria. Second, it is an open question whether more compact districting plans have a positive or negative effect on racial or political representation.

proximity index (RPI), the less compact a district.[5] The index satisfies three desirable properties: (1) voters are treated equally (anonymity), (2) increasing the distances between voters within a political district leads to a larger value of the index (clustering), and (3) the index is invariant to the scale, population density, and number of districts in a state (independence). In Appendix A, we show that any compactness index that satisfies these properties ranks districting plans identically to the relative proximity index.

The RPI has several advantages over existing measures of compactness. First, it is the only compactness index that permits meaningful comparisons across states. Second, the index does not assume (implicitly or otherwise) that voters are uniformly distributed across political districts. Many previously proposed measures adopt a geometric approach (using the perimeter length of political districts, for example) and fail to consider the distribution of voters within a state. Third, our measure is constructed at the state level. Some measures apply to political districts.[6] Yet the districting problem is fundamentally about partitioning; the shape of one element of the partition affects the shapes of the other elements. Analyzing individual pieces of a larger partition in isolation can be misleading. Fourth, although our index is simple, it is based on desirable properties that compactness measures should satisfy. Existing measures have been proposed in a relatively ad hoc fashion. At a minimum, our approach is a more principled way of narrowing the field of competing measures.

We apply the index to the districting plans of the 106th Congress using tract-level data from the U.S. census. In doing so, we are required to calculate each state's maximal compactness. This number is the denominator of our index. But calculating this number by brute force, enumerating the set of all feasible partitions and maximizing compactness over this set, is impossible.[7] Existing algorithms to solve similar problems in computer science and computational biology work only for small samples (≈ 100) or do not require that partitions have the same size. We develop an algorithm for approximating this partitioning problem that is suitable for very large samples and guarantees nearly equal populations in each partition. The algorithm is based on power diagrams—a generalization of classic Voronoi diagrams—which have been used extensively in algebraic and tropical geometry (Passare and Rullgard 2004; Richter-Gebert,

[5] For the empirical analysis and characterization of the optimally compact districting plan we use Euclidean distance. But since many of our results are proven in an arbitrary metric space, one can extend much of the analysis here by using driving distance or what many legal scholars refer to as "communities of interest."

[6] See Young (1988), however, and Section 2.2.

[7] A back-of-the-envelope calculation reveals that, for California alone, the cardinality of this set is larger than the number of atoms in the observable universe.

Sturmfels, and Theobald 2003), condensed matter physics, and toric geometry and string theory (Diaconescu, Florea, and Grassi 2002).[8]

The empirical results we obtain on the compactness of districting plans are interesting and in some cases quite surprising. The five states with the most compact districting plans are Idaho, Nebraska, Arkansas, Mississippi, and Minnesota. The five least compact states are Tennessee, Texas, New York, Massachusetts, and New Jersey. The districting plan that solves the minimum-partitioning problem is more than 40 percent more compact than the typical districting plan. States that are more compact tend to be states with a larger share of minorities and a larger difference between the percentages who vote Republican and Democrat. The latter is intuitive: states with more to gain from altering the design of political districts tend to do it more. Whether or not a state is forced to submit its districting plans to the Department of Justice (under section 5 of the Voting Rights Act) is also highly correlated with compactness. With only 43 observations, these estimates are not statistically significant. The rank correlations between the RPI and the most popular indexes of compactness, dispersion, and perimeter are −.37 and −.29, respectively.

We conclude our analysis by estimating a counterfactual of the 2000 congressional elections in California, New York, Pennsylvania, and Texas using optimally compact districts derived from our algorithm. To better understand the impact that a strict policy of maximal compactness might have on those elected, we estimate a seat-vote curve for the actual and hypothetical districting plans of each state. Seat-vote curves are a common tool that political scientists use to analyze the partisan consequences of districting plans. These curves are characterized by two things: bias and responsiveness. Bias reports, when the vote is split, twice the difference between the seat share the Democrats get and 50 percent. Responsiveness is the fraction of seats the Democrats get if the average vote goes up 1 percent. Responsiveness can be interpreted as a measure of the nature of democracy in the state. For instance, if Responsiveness is 1, then representation is proportional to the share of the vote. If it is greater than 1, it is majoritarian, and if it were to be infinity, then it would be winner take all.

The results of this exercise are quite illuminating. California, New York, Pennsylvania, and Texas all have substantially more responsive seat-vote curves under our new partition, but Bias is unchanged. These results show that maximally compact districts would have a statistically significant effect on voting outcomes, making election outcomes more responsive to actual votes.

The structure of the paper is as follows. Section 2 provides a brief legal history of compactness and an overview of existing measures. Section 3 presents the

[8] Power diagrams are a powerful tool to partition Euclidean space into cells by minimizing the distance between points in a cell and the centroid of that cell. We prove that maximally compact districts are power diagrams and that the line separating two adjacent districts is perpendicular to the line connecting their centroids, and all such lines separating three adjacent districts meet at a single point. It follows that the resulting districts are convex polygons.

relative proximity index and provides a brief discussion of its properties. Section 4 implements the index using data from the 106th Congress. Section 5 provides a counterfactual estimate of the congressional elections in four large states using the partitions derived from our index. Section 6 concludes with a discussion of potential extensions and generalizations of our approach. There are five appendixes. Appendix A contains an axiomatic derivation of the RPI, showing that any index that satisfies our three axioms will rank districting plans identically to the RPI. Appendix B provides further technical details, including a formal description of the algorithm used to compute maximally compact districts and proofs of all technical results. Appendix C provides a guide to programs to calculate the RPI, Appendix D contains figures comparing actual district maps and those obtained from our algorithm, and Appendix E contains figures comparing seat-vote curves.

2. Background and Previous Literature

2.1. *A Brief Legal History of Compactness*

Compactness has played a fundamental role in the jurisprudence of gerrymandering, both racial and political. Since *Gomillion v. Lightfoot* (364 U.S. 339 [1960]), where the court struck down Alabama's plan to redraw the boundaries of the city of Tuskegee, the court has recognized compactness as a relevant factor in considering racial gerrymandering claims. In *Gomillion* the court referred to the proposed district as "an uncouth 28-sided figure" (364 U.S. 340). Although *Gomillion* is considered by many to be a jurisprudential high-water mark, the role of compactness in considering racial gerrymandering claims has been affirmed in other decisions.[9] As Justice Sandra Day O'Connor put it, "We believe that reapportionment is one area in which appearances do matter" (*Shaw v. Reno,* 509 U.S. 603, 647 [1993]).

Compactness has also played an important role in partisan gerrymandering claims. It has been recognized by the court as a traditional districting principle. In *Davis v. Bandemer,* Justices Powell and Stevens described compactness as a major criterion (478 U.S. 173), and Justices Byron White, Brennan, Blackmun, and Marshall described it as an important criterion (106 S. Ct. 2797, 2815). In *Vieth,* the plurality acknowledged compactness as a traditional districting principle. Justice Anthony Kennedy, in his concurring opinion, stated that compactness is an important principle in assessing partisan gerrymandering claims: "We have explained that 'traditional districting principles,' which include 'compactness, contiguity, and respect for political subdivisions,' are 'important not because they are constitutionally required . . . but because they are objective

[9] In *Shaw v. Reno* (509 U.S. 630 [1993]), the court upheld a challenge to North Carolina's redistricting plan on the basis that the ill compactness of the districts was indicative of racial gerrymandering. See also *Thornburg v. Gingles* (478 U.S. 30 [1986]) or *Growe v. Emison* (278 U.S. 109 [1993]).

factors that may serve to defeat a claim that a district has been gerrymandered on racial lines.' . . . In my view, the same standards should apply to claims of political gerrymandering, for the essence of a gerrymander is the same regardless of whether the group is identified as political or racial" (541 U.S. 127, 335). Despite different views about what a judicially manageable standard is or might be, the court has been unanimous that it must include some notion of compactness.

2.2. Existing Measures of Compactness

There is a large literature in political science on the measurement of compactness. Niemi et al. (1990) provide a comprehensive account of the various measures that have been proposed (see also Young 1988).[10] Niemi et al. (1990) classify existing measures into four categories: (1) dispersion measures, (2) perimeter measures, (3) population measures, and (4) other miscellaneous measures.[11] The important takeaway is that all of these measures either fail to account for the population distribution or are not invariant to geographical size. As such, meaningful comparisons across states or time cannot be made.

One class of dispersion measures are based on length versus width of a rectangle that circumscribes the district (Harris 1964; Eig and Setizinger 1981; Young 1988). A second uses circumscribing figures other than rectangles and considers the area of these figures.[12] At least two moment-of-inertia measures have been suggested. Schwartzberg (1966) and Kaiser (1966) consider the variance of the distances from each point in the district to the district's areal center. Boyce and Clark (1964) consider the mean distance from the areal center to a point on the perimeter reached by equally spaced radial lines.

A second set of measures are those based on perimeters. The sum of perimeter lengths was suggested by Adams (1977), Eig and Setizinger (1981), and Wells (1982), but this measure is potentially intractable for reasons highlighted in the classic work of Mandelbrot (1967) on the length of the coastline of Great Britain. In fact, a measure based on fractal dimensions was proposed by Knight (2004). Various authors have proposed measures that compare the perimeter to the area of the district. Cox (1927) considers the ratio of the district area to that of a circle with the same perimeter.[13]

There are three population-based measures. Hofeller and Grofman (1990) propose two: the ratio of the district population to the convex hull of the district and the ratio of the district population to the smallest circumscribing circle.

[10] Some of these measures were originally proposed for purposes other than those involving legislative districts but were later applied by other authors to that issue. We cite the original authors.

[11] We draw heavily on their summary and classification.

[12] Reock (1961) proposes a circle, Geisler (1985) a hexagon, Horton (1932) and Gibbs (1961) a circle with diameter equal to the district's longest axis, and still others use the smallest convex figure (see Young 1988).

[13] For variants of Cox (1927), see Attneave and Arnoult (1956), Horton (1932), Schwartzberg (1966), or Pounds (1972).

Weaver and Hess (1963) suggest the population moment of inertia, normalized to lie in the unit interval.

Niemi et al.'s (1990) final miscellaneous category includes three measures: the absolute deviation of district area from average area in the state (Theobald 1970), a measure based on the number of reflexive and nonreflexive interior angles (Taylor 1973), and the sum of all pairwise distances between the centers of subunits of the district, weighted by subunit population (Papayanopolous 1973). Finally, Mehrotra, Johnson, and Nemhauser (1998) use a branch-and-price algorithm to compute a districting plan for South Carolina. Their objective function is how far people are from a graph-theoretic measure of the center of the district.

3. The Relative Proximity Index

3.1. *Basic Building Blocks*

Let **S** denote a collection of states with typical element $S \in \mathbf{S}$. A finite set S, whose elements we call individuals or voters, is a metric space with associated distance function $d_{ij} \geq 0$, which measures the distance between any two elements $i, j \in S$. Let $V_S = \{v_1^S, \ldots, v_n^S\}$ denote a finite partition of S into elements $v_i \in V_S$, which we shall refer to as voting districts, or districts. We will routinely refer to the partition V_S as a districting plan for state S and allow n to represent a generic integer. We restrict voting districts to be equal in size, up to integer rounding.[14,15] Let $\mathcal{V}_S$ denote the set of all partitions of S that satisfy this restriction. We say that a districting plan V_S is feasible if and only if $V_S \in \mathcal{V}_S$.

Definition 1. A compactness index for a state S is a map $c\colon V_S \mapsto \mathbb{R}_+$.

3.2. *The Relative Proximity Index*

The RPI is the ratio of two components. The numerator sums the pairwise squared distance between voters within each district in a state, as given by the actual districting plan in the state. The denominator is that same sum but for the districting plan that minimizes the sum.

Consider voter i in element $v \in V_S$ and define

$$\pi(V_S) = \sum_{v \in V} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2. \tag{1}$$

[14] This was first held as a requirement by the Court in *Baker v. Carr* (369 U.S. 186 [1962]) and is becoming a very strict constraint. For instance, a 2002 Pennsylvania redistricting plan was struck down because one district had 19 more people (not even voters) than another. The 2004 Texas redistricting had each district with the same number of people up to integer rounding. Yet the population may grow at drastically different rates across political districts between redistrictings. For instance, in the 2000 census, a typical state had a 23 percent difference in the populations of its smallest and largest districts.

[15] In symbols, $|v_i^S| \in \{\lfloor |S| / |V_S| \rfloor, \lceil |S| / |V_S| \rceil\}$ for all $v_i^S \in V_S$, where $\lceil x \rceil = \inf\{n \in \mathbb{Z} \,|\, x \leq n\}$ and $\lfloor x \rfloor = \sup\{n \in \mathbb{Z} \,|\, n \leq x\}$.

Similarly, let $V_S^* = \arg\min_{V_S \in \mathcal{V}_S} \{\pi(V_S)\}$. The RPI, for a partition of state S, V_S, is given by

$$\text{RPI} = \frac{\pi(V_S)}{\pi(V_S^*)}.$$

The RPI is well defined if $\pi(V_S^*) \neq 0$, which holds so long as all voters are not located at the same point.

In the nondegenerate case, the RPI ranges from 1 to infinity; higher numbers indicate less compactness. The index has an intuitive interpretation: a value of 3 implies that the current districting plan is roughly three times less compact than a state's maximal compactness.

3.3. *A Constructive Example*

Consider the state depicted in Figure 1. The nodes represent voters. There are two voting districts separated by the bold dashed line. Voters are spread evenly across the state; each adjacent voter is 1 kilometer apart. Voter 1 is 1 kilometer away from voters 2 and 4, $\sqrt{2}$ kilometers away from voter 5, $\sqrt{5}$ kilometers away from voter 6, and so on.

There are two steps involved in calculating the RPI. First, we calculate the numerator. For voter 1 the sum of squared distances is 5, since she is 1 kilometer away from voter 2 and 2 kilometers away from voter 3—and they are the only other voters in her district. For voter 2 the total is $1^2 + 1^2 = 2$, and for voter 3 it is $1^2 + 2^2 = 5$. Voters 4, 5, and 6 are symmetric to voters 1, 2, and 3, respectively. Thus, the numerator of our index is $2(5 + 2 + 5) = 24$.

The second step in calculating the RPI is to account for state-specific topography. This will represent the denominator of our index. There are nine other feasible partitions in addition to $\{\{1, 2, 3\}, \{4, 5, 6\}\}$.[16] We perform the same calculation as above for each of those partitions and then take the minimum of these 10 values. The minimizing partition is $\{\{1, 4, 5\}, \{2, 3, 6\}\}$, although $\{\{1, 2, 4\}, \{3, 5, 6\}\}$ achieves the same value. That value turns out to be $2(1^2 + 2 + 1^2 + 2 + 1^2 + 1^2) = 16$. The index is thus $24/16 = 3/2$.

The example provides a snapshot of the RPI and previews some of its properties. For instance, because the index is calculated relative to a state-specific baseline, neither the size of states nor their population density can solely alter the index. If we increased the distance between any two nodes in Figure 1 to 2 kilometers, the index would not change. Similarly, if we imputed 10 more individuals to each node—thinking of them in terms of neighborhoods rather than households—the index would be unaltered.

[16] They are $\{\{1, 2, 4\}, \{3, 5, 6\}\}$, $\{\{1, 2, 5\}, \{3, 4, 6\}\}$, $\{\{1, 2, 6\}, \{3, 4, 5\}\}$, $\{\{1, 3, 4\}, \{2, 5, 6\}\}$, $\{\{1, 3, 5\}, \{2, 4, 6\}\}$, $\{\{1, 3, 6\}, \{2, 4, 5\}\}$, $\{\{1, 4, 5\}, \{2, 3, 6\}\}$, $\{\{1, 4, 6\}, \{2, 3, 5\}\}$, and $\{\{1, 5, 6\}, \{2, 3, 4\}\}$.

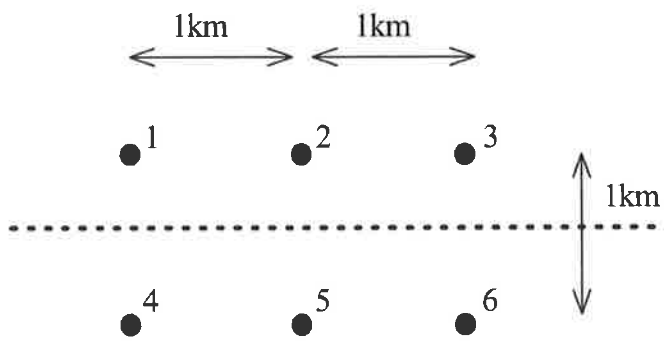

Figure 1. A simple example

3.4. Three Desirable Properties

Any desirable index of compactness should satisfy three properties. (Formal mathematical statements of these properties are provided in Appendix A.)

Anonymity. The index does not depend on the identity of any given voter.

Invariance. The index does not depend on a state's population density, physical size, or number of districts.

Clustering. If two states with the same number of voters, the same number of voting districts, and the same value for the minimum-partitioning problem have different total intradistrict distances, then the state with the larger value is less compact.

It is straightfoward to see from the above example that the RPI satisfies these properties. All voters are weighted equally, so anonymity is satisfied. The denominator of the RPI scales the index so that invariance is satisfied. Finally, clustering is satisfied because the numerator sums pairwise squared distances. In fact, we can say something much stronger:

Theorem 1. Any compactness index that satisfies anonymity, invariance, and efficient clustering ranks districting plans identically to the RPI.

Proof. See Appendix A.

The result is proved by noting that by transforming a given state (expanding the set of individuals and number of districts, for example) it can be compared to another state. Anonymity and independence ensure that this can be done in a way that does not alter the compactness index, and clustering then allows a comparison of two districting plans based on their total intracluster pairwise distances.

4. Implementing the Relative Proximity Index

In this section, we apply the RPI to the districting plans of the 106th Congress. The challenge with calculating our index is computing the denominator, which requires finding a districting plan that minimizes the distance between voters. This is a complex combinatorial problem for which existing algorithms are inadequate. We solve this problem by showing that optimal districting plans are akin to so-called power diagrams[17] and then modifying an algorithm presented in Aurenhammer, Hoffmann, and Aronov (1998) to create a power diagram. The key ingredient in the algorithm is the centroid, or geometric center, of existing districts,[18] a point that is provided in census data from the GeoLytics database. We apply our algorithm to the data from the 2000 census and calculate both the optimal districting plan following that census and the relative proximity index for the actual districting plans employed to elect the 106th Congress.

4.1. The Minimum-Partitioning Problem

Calculating the denominator of the relative proximity index is a complicated combinatorial problem. When partitioning n voters into d districts, the number of feasible partitions is $\{(n-1)!/[(n/d-1)!(n-n/d)!]\}^{d-1}$. So, for California alone, using data at the tract level, $n = 6{,}800$ and $d = 53$. The cardinality of the set of feasible partitions is $78.4 \times 10^{59{,}351}$. Technically speaking, the problem is nondeterministic polynomial-time hard (NP hard).

Similar problems arise in fields such as applied mathematics (computer vision), computer science and operations research (the k-way equipartition problem), and computational biology (gene clustering). The celebrated Mumford-Shah functional is a candidate functional designed to segment images (Mumford and Shah 1989). The structure of the functional contains two penalty functions: one to ensure that the continuous approximation is close to the discrete problem and another to penalize perimeter length. While the Mumford-Shah functional is a powerful tool for myriad problems, it cannot guarantee even nearly equal population size across districts.

If our objective function were simply distance, rather than distance squared, the problem would be precisely the k-way equipartition problem, which has received considerable attention in computer science and is related to a literature in computational biology employing minimum-spanning trees to partition sim-

[17] Power diagrams are a generalization of Voronoi diagrams due to Aurenhammer (1987). Voronoi diagrams are convex polygons with the important feature that each contains a so-called generator point such that that all other points within the polygon are closer to that generator point than to generator points of adjacent polygons.

[18] More precisely, a centroid is the intersection of all straight lines that divide the district into two parts of equal moment about the line.

ilar genes into clusters.[19] Good algorithms for the k-way equipartition problem when sample sizes are small (≈ 100) can be found in Ji and Mitchell (2005) and Mitchell (2003). This restriction makes these algorithms impractical for our purposes.

Below, we develop an algorithm to approximate the minimum-partitioning problem for large samples, based on power diagrams (a concept we make precise below), that guarantees nearly equal populations in each partition and runs in $O[n \log(n')]$ time, where n' is the number of voters and n is the number of districts in a state.

4.2. Optimally Compact Districting Plans and Power Diagrams

In this section, we show that optimally compact districting plans are power diagrams, a generalization of Voronoi diagrams, which were introduced into computational geometry by Aurenhammer (1987). Consider a set of generator points $m_1, \ldots, m_n$ in a finite dimensional Euclidean space. The power of a point (voter) $x \in S$ with respect to a generator point m_i, which is some arbitrary point, is given by the function $\text{pow}_\lambda(x, m_i) = \|x - m_i\|^2 - \lambda_i$, where $\|\cdot\|$ is the Euclidean norm.[20] The total number of voters assigned to generator point m_i is called its capacity, denoted K_{m_i}. A power diagram is an assignment of voters to generator points such that point x is assigned to generator point m_i if and only if $\text{pow}_\lambda(x, m_i) < \text{pow}_\lambda(x, m_j)$ for all $j \neq i$. Roughly speaking, voters are placed in the district whose centroid they are closest to. Let the points assigned to generator point m_i be denoted D_i, which is referred to as a cell. Note that no two D_i's can intersect, and furthermore, every $x \in S$ is in some D_i, and hence $\{D_1, \ldots, D_n\}$ is a partition of S. Note also that the dividing line between cells D_i and D_j in a power diagram satisfies $\|x - m_i\|^2 - \|x - m_j\|^2 = \lambda_i - \lambda_j$.

Definition 2. An optimally compact districting plan for state S is a feasible districting plan, V_S, with an associated total distance $\sum_{v \in V_S} \sum_{i,j \in v} (d_{ij})^2$ such that there does not exist another feasible districting plan, V_S', with an associated total distance $\sum_{v \in V'_S} \sum_{i,j \in v} (d_{ij})^2$ such that $\sum_{v \in V'_S} \sum_{i,j \in v} (d_{ij})^2 < \sum_{v \in V_S} \sum_{i,j \in v} (d_{ij})^2$.

We can now state our second key result:

Theorem 2. Optimally compact districting plans are power diagrams.

Proof. See Appendix B.

This theorem follows from three lemmas that partially characterize an optimal

[19] Without the constraint that each district must have an equal number of voters, the problem is the min-sum k-clustering problem, which was shown by Sahni and Gonzales (1976) to be nondeterministic polynomial-time (NP) complete. An approximation for it in a general metric space that runs in $n^{O(1/\epsilon)}$ time has been found by Bartal, Charikar, and Raz (2001). It is also closely related to the classic graph-partitioning problem, which is also known to be NP hard.

[20] When $\lambda_i = \lambda$ for all i, then the power diagram is a Voronoi diagram. Power diagrams are thus a generalization of Voronoi diagrams.

districting plan and establish that these characteristics imply a power diagram. The first lemma shows that our objective function is equivalent to a variant of the k-means objective function. This is important because it allows us to focus attention on district centroids.

The second lemma shows that any pair of districts are separated by a line perpendicular to a line connecting their centroids. This separating line is the locus of points at which the powers of the two centroids are equal. It represents all points at which one is indifferent between placing voters in one district or the other. Finally, we establish that all such lines separating any three adjacent districts meet at a single point; they are concurrent.

To see that these properties imply a power diagram, recall that a power diagram is a set of lines dividing a Euclidean space into a finite number of cells. The line separating two adjacent cells is such that the power of the points along this locus is equal to their respective centroids. And the power of a point is measured as a function of the difference between a point and the centroid of its district, which we have already established is equivalent to our objective function. It is important to note that if the line separating two adjacent districts were not perpendicular to the line connecting their centroids, then one could not be indifferent between points being in one district or the other everywhere along the line. This holds for all such pairs of districts, which implies concurrent lines. Taken together, these imply that optimally compact districtings are power diagrams.[21] Notice that, since all subsets of a convex set formed by drawing straight lines are convex, it follows that the resulting districts must be convex polygons.

Theorem 2 provides an important insight for building an algorithm, allowing us to use all we know about a partial characterization of optimally compact districts. There are three important caveats. First, we have not yet proven that there is a unique power diagram for every set of starting values. Second, we are able to map optimal districting plans into power diagrams only when distance is quadratic, because this guarantees that optimal districting involves straight lines. Mathematically, this is an obvious limitation. Practically, however, it boils down to assuming that courts punish outliers in a district more. Given this assumption, we are hard pressed to find a principled reason for courts to prefer higher order exponents. Third, power diagrams do not guarantee a global optimum to the minimum-partitioning problem because their structure depends on exogenously given starting values.

Figure 2A depicts the optimally compact districting plan for a hypothetical state. There are nine voters, arranged so that the state is a lattice. The stars

[21] Aurenhammer, Hoffmann, and Aronov (1998) prove a closely related theorem, taking squared distance from the centroid as the objective function. Their proof proceeds by showing that if an algorithm can be designed to find a power diagram, then it is an optimal partition. By contrast, we provide a constructive proof based on the perpendicular- and concurrent-line lemmas. We could, of course, state our lemma on the equivalence of the objective functions and then appeal to their result, but our current proof provides more information about optimal districtings.

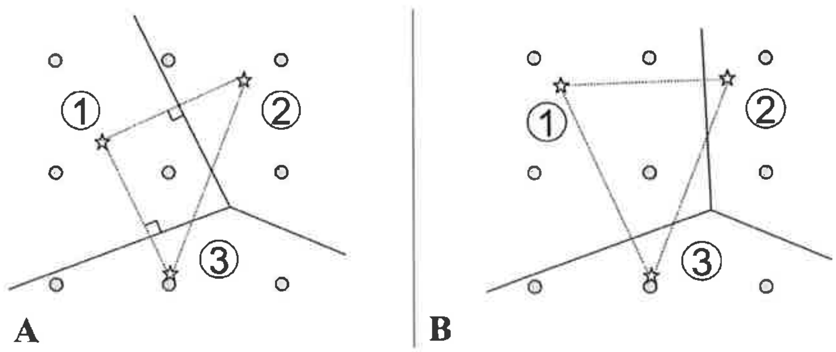

Figure 2. Good and bad generator points

represent the centroids of the resulting districts. Note that the line separating districts 1 and 2 is perpendicular to a line connecting their centroids (the same is true for districts 1 and 3 and for 2 and 3). This is an illustration of the perpendicular-line lemma alluded to above. The concurrent-line lemma is also illustrated by the intersection of the lines separating districts 1, 2, and 3 at a single point. The partition depicted is indeed the globally optimal partition. Once one knows that, the centroids of the districts are easy to compute.

In our problem, however, we do not know the optimal districts in advance, and so we must choose generator points that will not in general be the centroids of the optimal districting plan. An important part of the approximation problem is selecting and improving upon the generator points. To illustrate this point, consider Figure 2B, which chooses alternative generator points than those used to partition in Figure 2A. The generator point used for district 1 differs from that used in Figure 2A, resulting in four voters being placed in district 1 and only two in district 2, thereby violating the equal-size constraint.

4.3. An Algorithm Based on Power Diagrams

The algorithm we propose is a modification of the second algorithm presented in Aurenhammer, Hoffmann, and Aronov (1998). Since we know by theorem 2 that local optima of the RPI are power diagrams, we search within the set of power diagrams for one that is a feasible districting. However, as power diagrams are generated around sites, which we call $z_1, \ldots, z_n$, it is necessary to update the locations of the sites as well as the design of the districts.

We provide a complete formal treatment in Appendix B and here give a heuristic description of the algorithm. The algorithm takes the centroids of existing districts as starting generator points and computes a power diagram. Power diagrams do not require partitions (cells) to be even roughly equal, so, after constructing the diagram, the algorithm adjusts the district boundaries until the number of voters within each district is equal up to integer rounding. We then recalculate the centroids of the new districts and check to see if any pair of individuals can switch districts and reduce the objective function (total squared distances). Our modification of Aurenhammer, Hoffmann, and Aronov's algorithm continues to check until there are no more pairs that can be switched and reduce the objective function by a predetermined value. The algorithm then repeats itself—recalculating centroids, drawing power diagrams, adjusting boundaries, and so on—until it reaches a value within preset bounds for a stopping rule.

4.4. The Compactness of Political Districting Plans of the 106th Congress

The ideal data to estimate the relative proximity index would contain the geographical coordinates of every household in the United States, its political district, some measure of distance between any two households within a state, and a precise definition of communities of interest. This information is not available.

In lieu of this, we use tract-level data from the 2000 U.S. census from the GeoLytics database, which contain the latitude and longitude of the geographic centroid of each tract, the political district each centroid is in, and its total population.[22] Census tracts are small, relatively permanent statistical subdivisions of a county. The spatial size of census tracts varies widely depending on the density of settlement, but they do not cross county boundaries. Census tracts usually have between 2,500 and 8,000 persons and, when first delineated, are designed to be homogeneous with respect to population characteristics, economic status, and living conditions. Our main interest in using this level of aggregation (relative to blocks or block groups) is that census tracts are more likely to contain some notion of communities of interest.

An important consideration in the application of the RPI is how to handle tracts of different densities. The equal-representation constraint—districting plans must have the same number of individuals in each district up to integer rounding—is predicated on individuals, not tracts. Our algorithm, described below, addresses this issue by allowing one to divide tracts into arbitrarily small units. There is an important trade-off between computational burden and the variance in population across districts; the burden will lessen with technological progress. For ease of implementation, we have chosen not to split any tracts. As a robustness check, we split tracts of small states into four smaller parts and assigned them to the same longitude and altered their latitude by .001 degrees. In all cases, accuracy (and computing time) were substantially increased with little effect on the RPI.

To calculate the RPI for each state, we begin with the numerator of the index,

[22] For roughly 5,000 census tracts, information on congressional district was not provided. In these cases, we mapped the coordinates of the centroid of the tract and manually keypunched the congressional district to which it belonged.

$\sum_{v \in V} \sum_{i,j \in v} (d_{ij})^2$, where i and j are population centroids of tracts and v are voting districts. We weight the total distances by the population density of each tract. An identical calculation is performed for the denominator, but V is constructed by our power diagram algorithm.

The empirical results we obtain on the compactness of districting plans are displayed in Table 1. The maximum deviations from equal partitions in the actual data and those resulting from our algorithm are an indication of the degree to which the equal-size constraint holds. The bootstrapping technique that we used for the mean RPI is described below. It is important to realize that for every state, the elements of our partitions are more balanced than what appears in the actual districting plans. Further, the largest deviation from equal partitions in the actual data (Florida, .46) is substantially larger than our largest deviation (California, .22).

Table 1 illustrates that the five states with the most compact districting plans are Idaho, Washington, Arkansas, Mississippi, and New Hampshire. The five most compact states are Idaho, Nebraska, Arkansas, Mississippi, and Minnesota. The five least compact states are Tennessee, Texas, New York, Massachusetts, and New Jersey. The districting plan that solves the minimum-partitioning problem is more than 40 percent more compact than the typical districting plan. The rank correlations between the RPI and the most popular indexes of compactness, dispersion and perimeter, are −.37 and −.29, respectively.

Axiom 3 (invariance to scale, population density, and number of districts; see Appendix A) ensures that the RPI can be compared across states, but it does not guarantee that the distribution of RPI values across states is the same. It is entirely plausible that it is easier (a lower percentile of the distribution of RPI values from feasible partitions) to obtain a given value of RPI for Texas than, say, Florida. Thus, gleaning an understanding of how sensitive RPI values are for a given state is difficult.

To try to address this issue, we calculated 200 RPI values for each state by randomly generating starting values for the algorithm. Table 1 reports the means and associated standard deviations from this process and in what percentile in the distribution our original RPI value lies, if the distribution of RPI values is assumed to be normal. In all but one case, our original estimates are higher than the mean of the simulated distribution, and in most cases, under the normality assumption, we are at the far extreme of the right tail of the distribution. There are four notable exceptions: Oklahoma, Oregon, Rhode Island, and Wisconsin. In these states, our estimate of the RPI is at the median or below in the simulated distribution. This is likely due to the fact that the current partitions of these states generate starting values that are highly nonoptimal. To obtain maximal compactness in these states, a significant restructuring is likely needed.

To understand what state demographic characteristics are correlated with compactness, we estimate a state-level ordinary least squares regression where the dependent variable is the RPI and the independent variables are the percentages

Table 1
The Relative Proximity Index, 2000

State	RPI	Max Deviation: Actual	Max Deviation: Algorithm	Mean RPI	SD RPI	Percentile
Alabama	1.21	.27	.05	.99	.03	1.00
Arizona	1.34	.20	.15	1.27	.04	.97
Arkansas	1.08	.14	.05	.78	.01	1.00
California	1.49	.17	.04	.96	.03	1.00
Colorado	1.59	.15	.05	1.28	.02	1.00
Connecticut	1.36	.02	.01	1.09	.35	.78
Florida	1.39	.46	.07	.83	.08	1.00
Georgia	1.24	.14	.09	.90	.01	1.00
Hawaii	1.59	.09	.04	1.48	.02	1.00
Idaho	.97	.10	.02	.80	.02	1.00
Illinois	1.43	.29	.11	.98	.07	1.00
Indiana	1.49	.20	.06	1.05	.02	1.00
Iowa	1.38	.06	.05	1.29	.01	1.00
Kansas	1.11	.08	.05	.95	.01	1.00
Kentucky	1.51	.14	.05	1.22	.01	1.00
Louisiana	1.15	.13	.05	.79	.43	.80
Maine	1.39	.04	.03	1.15	.01	1.00
Maryland	1.52	.22	.04	1.25	.02	1.00
Massachusetts	1.87	.10	.05	1.54	.01	1.00
Michigan	1.24	.13	.04	.99	.02	1.00
Minnesota	1.05	.16	.05	.90	.02	1.00
Mississippi	1.02	.18	.05	.87	.01	1.00
Missouri	1.38	.23	.05	1.01	.16	.99
Nebraska	1.01	.05	.04	.89	.23	.70
Nevada	1.38	.08	.05	1.19	.01	1.00
New Hampshire	1.10	.01	.00	1.09	.00	.95
New Jersey	2.27	.21	.05	1.69	.02	1.00
New Mexico	1.23	.06	.04	1.14	.01	1.00
New York	1.83	.21	.10	1.45	.45	.80
North Carolina	1.33	.28	.04	1.15	.09	.97
Ohio	1.62	.13	.05	1.42	.01	1.00
Oklahoma	1.24	.09	.05	1.42	.36	.31
Oregon	1.26	.09	.04	1.21	.28	.56
Pennsylvania	1.81	.25	.22	1.27	.05	1.00
Rhode Island	1.18	.03	.02	1.18	.01	.55
South Carolina	1.22	.21	.04	1.27	.02	.00
Tennessee	2.91	.25	.04	2.59	.04	1.00
Texas	1.90	.30	.22	1.24	.07	1.00
Utah	1.46	.06	.04	1.40	.01	1.00
Virginia	1.38	.22	.07	1.14	.04	1.00
Washington	1.17	.15	.06	.77	.03	1.00
West Virginia	1.68	.06	.05	1.61	.01	1.00
Wisconsin	1.40	.11	.08	1.22	.58	.62

Note. Relative proximity index (RPI) values are calculated using tract-level data from the 2000 census. Max Deviation is calculated as 1 minus the total population of the largest congressional district divided by the total population of the smallest congressional district. Mean RPI is calculated as the mean of 200 repetitions of the RPI, each having different starting values.

of the populations that are black, Asian, or Hispanic; population density; difference in presidential vote shares between Democrats and Republicans; and whether the state is required to submit its districting plans to the Department of Justice under the preclearance provision of section 5 of the Voting Rights Act (not shown in tabular form).[23] States that are more compact tend to be states with a larger share of blacks and a larger difference between the percentages who vote Republican and Democrat. The latter is intuitive: states with more to gain from altering the design of political districts tend to do it more. Whether or not a state is forced to submit its districting plans is also highly correlated with compactness. Consistent with axiom 2 (efficient clustering; see Appendix A), the RPI is uncorrelated with population density. It is important to note that none of these partial correlations are statistically significant because of small samples.

Beyond the technical considerations, perhaps the best evidence in favor of our approach can be illustrated visually. The figures in Appendix D present side-by-side comparisons of congressional district maps for actual districting plans and those obtained from our algorithm.[24] Figures D1 and D2 illustrate this comparison for the least and most compact states, Tennessee and Idaho, respectively. The districts in Tennessee, under the current plan, resemble the salamander-shaped districts drawn by Eldridge Gerry that gave rise to the name "gerrymandering." Under the algorithm, however, Tennessee is transformed into a neat set of convex polygons. Idaho is at the other extreme. Because the state need only be cut into two equal parts, the existing cut and our preferred cut are very similar. Further, our partition provides a more equal distribution of voters across the districts, which explains why the calculated RPI is slightly less than 1.

These figures illustrate three key points. First, the geometric properties discussed above (the perpendicular- and concurrent-line lemmas and the convexity of political districts) are immediately apparent. Second, those states that rank relatively high (low) in terms of the RPI appear to quite different (similar) to the partition resulting from our algorithm. Third, Figures D3 and D6 (Hawaii and Nevada) suggest that communities of interest are an important consideration. In the actual plans, Honolulu and Las Vegas are their own districts, while the rest of the state is contained in another. The issues faced by residents of the outer islands might well be more similar to each other than they are to those of residents in Honolulu. This serves to highlight why compactness is only one factor that weighs on the redistricting question. The RPI in its current implementation ignores this consideration. An RPI with a more general notion of

[23] The states that are subject to the preclearance provision are Alabama, Alaska, Arizona, Georgia, Louisiana, Mississippi, South Carolina, Texas, and Virginia.

[24] For a complete set of maps, see Roland Fryer, Papers (http://www.economics.harvard.edu/faculty/fryer/papers_fryer).

distance or carefully selected starting values for the power diagram can address this issue.

4.5. *How Good an Approximation?*

One wonders how good an approximation our algorithm provides to an exact solution to the minimum-partitioning problem. We have two ways to address this question. The first is to note that the compter science literature on power diagrams and algorithms based on them (see, for example, Aurenhammer, Hoffmann, and Aronov 1998) shows that thse algorithms typically perform very well (to within a few percentage points of the actual optimum). This can be shown by taking hypothetical data sets to which the exact solution can be found (because they are sufficiently small) and then comparing the performance of the algorithm. Yet it is not clear how performance on these algorithms scales.

One might also wonder whether the use of tract-level data (rather than finer grained block-level data) leads to markedly less precision. To address this, we ran several smaller states at the block level. The average RPI calculated at the block level is slightly higher than in the tract-level analysis reported in Table 1. For instance, Nebraska has an RPI of 1.01 in the tract-level data and 1.33 using blocks. The key issue with block-level analysis is our inability to calculate RPI for medium or large states. On computers with eight high-speed processors and 16 gigabytes of RAM (such as the one we used in our analysis), we estimate that large states such as Texas and California would take several years each to finish.[25]

5. Election Counterfactuals

Thus far, we have derived an index of compactness, shown how one implements the index, and provided some basic facts about the most and least compact districting plans and what correlates with these plans. We conclude our analysis with some suggestive evidence on the impact of maximally compact districting plans on election outcomes in four large states.

In winner-take-all election contests, such as elections for representatives to the U.S. Congress and for electoral votes for the U.S. presidency, the winner is determined by which candidate receives the plurality of the votes. In most of these cases, only the top two parties need to be considered, which yields an easy condition for an election win in a district.

Assuming that there are n districts, labeled $i \in [1, \ldots, n]$, let ϕ_i denote the proportion of the two-party vote received by the candidate from the first

[25] Currently, large clusters or supercomputers can run at above 1.5 petaflops (a petaflop is 10^{15} floating point operations per second), and the IBM Sequoia project is projected to run at 20 petaflops by 2011. That is roughly the power of 2,000,000 laptops, or around 11,000 times faster than the machine on which we conducted our analysis. Thus, analysis of our index at the block level will be feasible soon.

party (in examples to follow, the Democratic Party). The candidate's victory can then be expressed as $s_i = w_i \mathbb{I}(\phi_i > \frac{1}{2})$, where w_i denotes how many seats are determined by the vote: one for single-member districts or three or more for the electoral college, for example. Two important summary statistics are the average district vote, $\Phi = (1/n)\sum_{i=1}^{n} \phi_i$, and the seat share, $S = \sum_{i=1}^{n} s_i / \sum_{i=1}^{n} w_i$.

Many other statistics can be generated using the vote and seat outcomes directly, but we are particularly interested in partisan bias and responsiveness. Namely, Bias $= 2E(S|\Phi = .5) - 1$ estimates the deviation from the median share of seats if each side receives an identical average district vote, and Responsiveness $= (dS/d\Phi)|\Phi$ estimates how a small shift in the average district vote would translate into a shift in the share of seats. This estimate is taken at either the observed average district vote or the median vote. Bias measures the degree to which an evenly divided state would elect an uneven slate of representatives, and Responsiveness is the fraction of seats the Democrats get if the average vote goes up 1 percent.

5.1. Data and Statistical Framework

Our empirical strategy has four steps. First, we estimate a cross-sectional regression of Democratic vote shares on controls such as past election results and incumbency using the 2000 congressional districting plan. The regression is at the voter tabulation district (VTD) level, a subdivision of congressional districts. Second, using the optimally compact congressional districting plans we devised in Section 4, we reassign voter districts to new congressional districts. Not only will this change how voter district results are aggregated to the congressional district level, it will also change some of the controls for each voter district. Third, we use the coefficient estimates and the estimate of residual variance from the voter district regression to simulate outcomes under both the actual districting plan and the optimally compact districting plan. Finally, we aggregate VTD-level results up to the congressional districts in each simulation and compare the distribution of simulations across the two districting plans.

We use VTD-level election return data from U.S. elections for the 105th and 106th Congresses for four large states: California, New York, Pennsylvania, and Texas. These states were chosen because of their large numbers of congressional districts (roughly 30 or greater) and the availability of vote shares by VTD. There are approximately 300 VTDs in a typical congressional district, although there is substantial variation. In our data, for instance, California has 7,000 VTDs for 50 districts, Texas has 8,000 for 30, Pennsylvania has 9,000 for 20, and New York contains 13,000 for 30.

The intuition behind our approach is straightforward. Consider Figure D7, which depicts the existing districting plan of New York and the plan derived from our algorithm. To fix ideas, concentrate on the western portions of the state. There are roughly 433 VTDs in each congressional district in New York. Suppose an election takes place. Currently, a congressional representative is chosen by aggregating the votes from the VTDs within each district. In Figure D7, this amounts to adding votes from roughly 433 voting centers in districts 27–31. Now suppose we want to estimate how the choice of representatives would change if the districting plan were drawn to maximize compactness. To do this, we simply take note of which VTDs are in the new partitions and aggregate within each new district. In short, we disaggregate down to the VTD level, take note of the new districting lines, and then aggregate up taking these boundaries into account. As before, the winner of the new districts (in Figure D7 this now amounts to districts 4, 6, 8, and 17) is determined by aggregating the votes from VTDs.

There are a few complications. First, we need to assign candidates to the new districts in a reasonable manner. Second, we need to take into account the results of previous elections and whether the candidate is an incumbent—both of these factors weigh heavily on the prediction of future elections. Third, we need to think about how to get standard errors on our estimates.

To formalize the intuition above, we employ techniques from elementary Bayesian statistics developed in Gelman and King (1994). We provide a terse synopsis of their approach below. The crux of the Gelman-King method is a linear model with two distinct error components of the form

$$\phi_i = X\beta + \gamma_i + \varepsilon_i. \tag{2}$$

The vector X consists of an intercept term, results from the previous election, and an incumbent dummy.

To derive precise predictions in this framework, more structure has to be placed on the error terms. Let $\gamma_i \sim N(0, \sigma_\gamma^2)$ represent the systematic error component, an expression of the unobserved variables that applied before the election campaign began and would be identical if the election were to be run again. This might include the result in the previous election, the race of the candidates, or a relevant change in election law. The unpredictability of the behavior of voters is also a source of systematic error.

The second source of error is a random component that can be explained by random events during the election, such as the weather on election day or the reaction of the public to an unintentional gaffe. Let $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$.

There are two key assumptions in the Gelman-King method. First, errors are expressed in terms of two parameters: σ^2, the sum of the individual variances σ_γ^2 and σ_ε^2, and λ, the proportion of the total variance attributed to the systematic component; $\lambda = \sigma_\gamma^2/(\sigma_\gamma^2 + \sigma_\varepsilon^2)$. Second, the counterfactual assumes that the regrouping of voters into new districts will not have a systematic effect on voting behavior.

5.1.1. Estimating λ and σ^2

In practice, a districting map is constant over a series of elections. Thus, λ and σ^2 are found by taking the mean of individual estimators from each year.

In each year, σ^2 is the variance of the random error term in equation (2), and λ, the fraction of the error attributed to systematic error, is estimated by including the results of the previous election as an explanatory variable in the current one. By calculating this for each election that did not follow a redistricting (that is, in which the electoral map is identical to that of the previous election) and taking the mean, we have an estimator for λ.[26]

5.1.2. Generating Hypothetical Future Elections

To predict the properties of a subsequent election using the same districting plan, a series of hypothetical elections are simulated using the estimates for β and σ^2. A new set of explanatory variables **X** is used to demonstrate the conditions at the election. Since no information can be derived about the nature of the systematic error component beforehand, one error term is used, $\omega = \gamma + \varepsilon$, with variance σ^2. Thus, a single hypothetical election is then generated by drawing from

$$\phi_{\text{hyp}} = \mathbf{X}_{\text{hyp}}\beta + \delta_{\text{hyp}} + \omega, \tag{3}$$

where β is the posterior distribution, with mean $\hat{\beta} = (X'X)^{-1}X'\phi$ and (with a normality assumption) variance $\Sigma_\beta = \sigma^2(X'X)^{-1}$. The δ term is used to produce hypothetical elections whose average district vote is desired to be different from the original. Integrating out the conditional parameters β and γ, one obtains the marginal distribution:

$$\phi_{\text{hyp}}|\phi \sim N[\lambda\mathbf{v} + (\mathbf{X}_{\text{hyp}} - \lambda\mathbf{X})\hat{\beta} + \delta,\ (\mathbf{X}_{\text{hyp}} - \lambda\mathbf{X})\Sigma_\beta(\mathbf{X}_{\text{hyp}} - \lambda\mathbf{X})'^2]\sigma^2 I.$$

To evaluate the election system, let $\mathbf{X}_{\text{hyp}} = \mathbf{X}$; to evaluate under counterfactual conditions, set $\mathbf{X}_{\text{hyp}}$ to the desired explanatory variables.

5.1.3. Comparing Districting Plans

With the above statistical model in hand, we can predict elections under different partitions of a state into voting districts. The procedure is as follows. First, we estimate the model in equation (2). Second, having generated a new map through our algorithm, we determine the values for the explanatory variables for each district (for example, incumbency), either by aggregating and averaging the previous values in each precinct or by making sensible predictions for their value. In terms of vote shares, we simply aggregate the VTDs in the new partitions. For incumbency, we assign each incumbent to the latitude and longitude of the centroid of his or her district. Under the new districting plan, if there is one such incumbent per district, he or she becomes the incumbent used in the model. In the rare cases where there is more than one incumbent assigned to a district under a new districting plan, we break the tie by choosing the incumbent

[26] Ideally, one would have historical votes for many years to tease out the systematic error component. We have only 2 years of such data.

closest to the resulting centroid and moving the other incumbent to another district to keep the numbers constant. Finally, with our new map we simulate the model 1,000 times; deriving the relevant parameters is straightforward.

5.2. Analyzing Seat-Vote Curves

Using the methodology described above, the figures in Appendix E provide seat-vote curves for California, New York, Pennsylvania, and Texas under each state's actual districting plan and the plan that maximizes its compactness. The vertical axes depict the proportion of seats won by Democrats. The horizontal axes depict the share of votes that the Democrats earned in the election. Each figure reports two interesting quantities: Vote is the average district vote the Democrats received in the election, and Seats is the fraction of seats the Democrats received in the election (not the hypothetical seat share). The dark lines represent our estimate of the seat-vote curve, and the two lines parallel to them are 95 percent confidence intervals. One can see that there is a marked difference between the seat-vote curves estimated from the actual data and those estimated from the partition developed by our algorithm in California and New York. The slope of the curve is significantly steeper in both states. The slopes in Texas and Pennsylvania are also slightly steeper, but the difference is much less dramatic.

To get a better sense of the magnitudes involved, Table 2 presents our estimates of Bias and Responsiveness for the actual partition of our four states and those gleaned from the algorithm. We also report the t-statistic on the difference between them. Under maximally compact districting, measures of Bias are slightly smaller in all states except Pennsylvania, although none of the differences are statistically significant. In terms of responsiveness, however, there are large and statistically significant differences between the existing partitions and those that are maximally compact. New York, in particular, has a fivefold increase, from .482 to 2.51. In other words, under the current partition, a 1 percent increase in vote share for Democrats results in a .482 percent increase in seats. When districting is maximally compact, however, a 1 percent increase in vote share results in a 2.51 percent increase in seats. The next largest change is in California—increasing from 1.086 to 1.731. Pennsylvania and Texas show smaller increases, which are statistically significant at the 10 percent level.

6. Concluding Remarks

There will be continued debate about the design of districting plans. We have developed a simple but principled measure of compactness. Our measure can be used to compare districting plans across states and time, a feature not found in existing measures, and our algorithm provides a way of approximating the most compact plan. Further, the impact that a maximally compact districting plan can have on the responsive of votes is encouraging. These are first steps

Table 2
Partisan Bias and Responsiveness: Actual versus Maximally Compact Districtings

State	Bias			Responsiveness		
	Actual	Algorithm	*t*-Statistic	Actual	Algorithm	*t*-Statistic
California	.028	.007	.469	1.086	1.731	−4.327*
	(.010)	(.045)		(.069)	(.132)	
New York	.103	.018	1.051	.482	2.51	−6.540*
	(.014)	(.080)		(.036)	(.308)	
Pennsylvania	−.0027	.031	−.363	1.138	1.562	−1.800⁺
	(.021)	(.076)		(.128)	(.198)	
Texas	.062	.039	.334	.8872	1.305	−1.717⁺
	(.024)	(.064)		(.103)	(.221)	

Note. Estimates are based on voter tabulation district–level election return data for the 105th and 106th Congresses.
⁺ Statistically significant at the 10% level.
* Statistically significant at the 5% level.

toward a more scientific understanding of districting plans and their effects. Extensions and generalizations abound.

Perhaps the most obvious extension is to consider higher dimensional spaces, generalized distance functions, and communities of interest. Aurenhammer and Klein (2000) provide a comprehensive survey of Voronoi diagrams and how to incorporate generalized notions of distance, including p-norms, convex and airlift distances, and nonplanar spaces. These extensions are not only mathematically interesting and elegant: they have real-world content. Consider the following thought experiment. Suppose there is a city on a hill.[27] On the west side is a mild, long incline toward the rest of the city, which is in a plane. On the east side is a steep cliff, either impassable or with just a narrow, winding road that very few people use. While the next residential center to the east is much closer to the hilltop on a horizontal plane, it is much farther in terms of all sorts of distances that we think might matter: transportation time, intensity of social interactions, sets of shared local public goods and common interests, and so forth. Thus, for all practical purposes, one probably wants to include the hilltop in a western district rather than an eastern one. More general notions of distance can handle this. A similar situation arises when there is a natural boundary (for example, a river or highway) that effectively segregates or reduces communication between two population centers that are geographically very close. Conversely, there could be something (such as a tunnel or subway) that makes two non-connected regions effectively close to each other, or there may be other notions of communities and shared interest that lend themselves to a natural clustering. It is imperative to note that the derivation of our index assumed only a general metric space—many of these ideas fit squarely within our framework. The empirical application of the index, however, required us to only consider Euclidean

[27] We are grateful to Roland Benabou for this illustrative example.

distances. The challenge ahead is to incorporate more general notions of distance into an empirically tractable algorithm.

Appendix A

An Axiomatic Derivation of the Relative Proximity Index

A1. Three Properties

We now describe three properties that any compactness index should satisfy and formally discuss each in turn.

A1.1. Axiom 1: Anonymity

Axiom 1, an anonymity condition in the same spirit as that typically used in social choice theory (Arrow 1970), requires that all individuals be treated equally. That is, any compactness index should not depend on the particular identities (race, political affiliation, wealth, and so forth) of voters. Consider a state S with associated partition V and compactness index $c(V, S)$. For any bijection h: $S \to S$ and compactness index $c_h(V, S)$, $c_h(V, S) = c(V, S)$.

A1.2. Axiom 2: Clustering

Compactness is fundamentally a mathematical partitioning problem—deciding who to group with whom in a political district. Clustering is the quintessential objective (Bartal, Charikar, and Raz 2001).[28] Our second axiom requires that if two states with the same number of voters and voting districts and the same value for the minimum-partitioning problem have different weighted intradistrict distances, then the state with the larger value is less compact.

Let $\gamma_k = \sum_{i,j \in v} \alpha_{ij} (d_{ij})^{\delta}$ for $k = \{1, \ . \ . \ . \ , n\}$ and let $g(\gamma_1, \ . \ . \ . \ , \gamma_n) : \mathbb{R}^n \to \mathbb{R}$ be a monotonic, increasing function. Consider two states, S_1 and S_2, and partitions V and V', respectively, such that S_1 and S_2 have the same number of voters and the same number of districts, and

$$\min_{V \in \nu_{S_1}} g_{S_1}(\gamma_1, \ . \ . \ . \ , \gamma_n) = \min_{V \in \nu_{S_2}} g_{S_2}(\gamma_1, \ . \ . \ . \ , \gamma_n).$$

Then

$$g_{S_1}(\gamma_1, \ . \ . \ . \ , \gamma_n) > g_{S_2}(\gamma_1, \ . \ . \ . \ , \gamma_n) \Rightarrow c(V, S_1) > c(V', S_2).$$

A1.3. Axiom 3: Independence

Our final axiom requires that any measure of the compactness of a state be insensitive to its physical size, population density, and number of districts. This is vital for making cross-state comparisons of districting plans. Before stating the property formally, we need some further notation. We say that a state $\hat{S}$ is

[28] Other common objectives are distance from the geographic centroid of each partition or distance from a representative (typically the center of a cluster and not necessarily the center of the partition).

an n replica of S if and only if $\forall i \in S, \exists j_1, \ldots, j_n \in \hat{S}$ such that $d_{ij} = 0, \forall i$ and $d_{j_i j_k} = 0, \forall i, k$. It is also useful to have a shorthand for the realized value of the minimum-partitioning problem. Consider two partitions of state S, V and V', with ρ and ρ' elements, respectively. Let $V_S^{\min_\rho}$ and $V_S^{\min_{\rho'}}$ be the respective minimizing partitions.

Consider $S, \hat{S} \in S$ with cardinality $|S|$ and $|\hat{S}|$, respectively.

Scale. If $d_{ij} = \lambda d_{ij}$ for all $i, j \in S, \hat{S}$ then $c(V, S) = c(V, \hat{S})$ for all V.

Density. If $|\hat{S}| = \lambda |S|$ and $\hat{S}$ is a λ replica of S, then $c(V, S) = c(V, \hat{S})$ for all V.

Number of Districts.

$$\text{If } \frac{\sum_{v \in V_S^\rho} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2}{V_S^{\min_\rho}} = \frac{\theta \sum_{v \in V_S^{\rho'}} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2}{V_S^{\min_{\rho'}}} \Rightarrow c(V, S) = \theta c(V', S).$$

Density independence means that if we replicate a state by multiplying the number of people in each household by λ, the index of compactness is unaltered. For instance, when comparing two voting districts (Cambridge, Mass., and New York City, for example) that differ in their population density, the index provides the same cardinal measure of compactness.

Scale independence provides a similar virtue, permitting comparisons across states that differ in the distances between individuals (Massachusetts and Texas, say), allowing one to increase the distances between all individuals in a state by a constant with no resulting change in the index. Independence with respect to the number of districts is also vital in making cross-state comparisons.

A2. Uniqueness Result

Let $O_c = (\mathbb{R}_+, \succeq)$ denote the ordered set generated by the relative proximity index c, and let $O_{\hat{c}}$ denote the ordered set over elements $V_S \in \mathcal{V}_S$ generated by any other compactness index. We say that two indexes, c and $\hat{c}$, are ordinally isomorphic if $O_c = O_{\hat{c}}$. We are now equipped to state our main result. The proof of this follows.

Theorem 1.

1) The relative proximity index satisfies anonymity, clustering, and independence.

2) Suppose that $\delta = 2$ and that $g_{S_i}(\cdot)$ is symmetric for all i; then any compactness index that satisfies anonymity, clustering, and independence is ordinally isomorphic to the relative proximity index.

A2.1. Proof of Theorem 1.1

That the RPI satisfies the three axioms follows from five simple lemmas that we now state and prove.

Lemma 1. The relative proximity index satisfies anonymity.

Proof. Consider a partition V of state S and an associated compactness index

$c(V, S)$. Now consider a bijection $h: S \rightarrow S$. The term $\sum_{v \in V_S} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2$ is unchanged since h is a bijection, and hence there are the same number of points in each element of V, and they are at the same points. For identical reasons the denominator of the RPI does not change, and hence $c(V, S) = c_h(V, S)$ for any bijection h.

Lemma 2. The relative proximity index satisfies clustering.

Proof. Let there be two partitions, V_S^1 and V_S^2, such that

$$\sum_{v \in V_S^1} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2 > \sum_{v \in V_S^2} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2. \tag{A1}$$

Clustering requires that

$$c(V_S^1, S) > c(V_S^2, S).$$

Suppose, by way of contradiction, that expression (A1) holds, and

$$c(V_1, S) < c(V_2, S). \tag{A2}$$

That is,

$$\frac{\sum_{v \in V_S^1} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2}{\min\limits_{V \in \mathcal{V}_S} \sum_{v \in V} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2} < \frac{\sum_{v \in V_S^2} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2}{\min\limits_{V \in \mathcal{V}_S} \sum_{v \in V} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2}. \tag{A3}$$

The denominators are identical, and hence the supposition requires that

$$\sum_{v \in V_S^1} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2 < \sum_{v \in V_S^2} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2, \tag{A4}$$

a contradiction. Q.E.D.

Lemma 3. The relative proximity index satisfies density independence.

Proof. Consider S and $\hat{S}$, with $|S|$ and $|\hat{S}|$, respectively, and with $\hat{S}$ a λ replica of S. We need to show that $\text{RPI}(V, S) = \text{RPI}(V, \hat{S})$ for all $V \in \mathcal{V}_S$, $V \in \mathcal{V}_{\hat{S}}$. That is,

$$\frac{\sum_{v \in V_S} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2}{\min\limits_{V \in \mathcal{V}_S} \sum_{v \in V} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2} = \frac{\sum_{v \in V_{\hat{S}}} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2}{\min\limits_{V \in \mathcal{V}_{\hat{S}}} \sum_{v \in V} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2}$$

for all $V \in \mathcal{V}_S$, $V \in \mathcal{V}_{\hat{S}}$. By the definition of a λ replica, the right-hand side of the above equation is simply

$$\frac{\lambda \sum_{v \in V_S} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2}{\lambda \min\limits_{V \in \mathcal{V}_S} \sum_{v \in V} \sum_{i \in v} \sum_{j \in v} (d_{ij})^2},$$

which is clearly equal to the left-hand side for any partition. Q.E.D.

Lemma 4. The relative proximity index satisfies scale independence.

Proof. Scale independence requires that for two states, S and $\hat{S}$, with $d_{jk} = \lambda d_{jk}$ for all $j, k \in S, \hat{S}$. Then $c(V, S) = c(V, \hat{S})$ for all $V \in \mathcal{V}_S$, $V \in \mathcal{V}_{\hat{S}}$. That is,

$$\frac{\sum_{v\in V_S}\sum_{i\in v}\sum_{j\in v}(d_{ij})^2}{\min\limits_{V\in \nu_S}\sum_{v\in V}\sum_{i\in v}\sum_{j\in v}(d_{ij})^2} = \frac{\sum_{v\in V_{\hat{S}}}\sum_{i\in v}\sum_{j\in v}(d_{ij})^2}{\min\limits_{V\in \nu_S}\sum_{v\in V}\sum_{i\in v}\sum_{j\in v}(d_{ij})^2}$$

for all $V \in \mathcal{V}_S$, $V \in \mathcal{V}_{\hat{S}}$. Scale independence means that the right-hand side of the above equation is simply

$$\frac{\sum_{v\in V_S}\sum_{i\in v}\sum_{j\in v}(\lambda d_{ij})^2}{\min\limits_{V\in \nu_S}\sum_{v\in V}\sum_{i\in v}\sum_{j\in v}(\lambda d_{ij})^2} = \frac{\lambda^2\sum_{v\in V_S}\sum_{i\in v}\sum_{j\in v}(d_{ij})^2}{\lambda^2\min\limits_{V\in \nu_S}\sum_{v\in V}\sum_{i\in v}\sum_{j\in v}(d_{ij})^2},$$

which is clearly equal to the left-hand side for any partition.

Lemma 5. The relative proximity index satisfies number-of-districts independence.

Proof. The proof follows immediately from the definition of independence with respect to number of districts. Q.E.D.

We can now prove theorem 1.2. It is proved by transforming a given state so that it can be compared to another state. Anonymity and independence ensure that this can be done in a way that does not alter the compactness index, and clustering then allows a comparison of two districting plans to be made based on their total intracluster pairwise distances.

A2.2. Proof of Theorem 1.2

From theorem 1.1 we have $\mathrm{RPI}(V, S_m) > \mathrm{RPI}(\hat{V}, S_n) \Rightarrow c(V, S_m) > c(\hat{V}, S_n)$ for any m, n. Suppose that theorem 1.2 is not true. This implies that

$$c(V, S_m) > c(\hat{V}, S_n) \quad \text{and} \quad \mathrm{RPI}(V, S_m) < \mathrm{RPI}(\hat{V}, S_n) \tag{A5}$$

or

$$c(V, S_m) < c(\hat{V}, S_n) \quad \text{and} \quad \mathrm{RPI}(V, S_m) > \mathrm{RPI}(\hat{V}, S_n)$$

for some m, n.

If $S_m = S_n$, then the argument is straightforward. Begin with the first pair of inequalities. Note that equality implies that $\mu_{ij} = \mu$ for all i, j and that symmetry of g combined with equality implies that g is additively separable in its arguments. Then by equality and clustering we have

$$\sum_{v\in V_{S_m}}\sum_{i\in v}\sum_{j\in v}(d_{ij})^2 > \sum_{v\in \hat{V}_{S_n}}\sum_{i\in v}\sum_{j\in v}(d_{ij})^2 \Rightarrow c(V, S_m) > c(\hat{V}, S_n),$$

since $\mathrm{RPI}(V, S_m) < \mathrm{RPI}(\hat{V}, S_n)$ and

$$S_m = S_n \Rightarrow \min_{V\in \nu_{S_m}} \sum_{v\in V_{S_m}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2 = \min_{V\in \nu_{S_n}} \sum_{v\in \hat{V}_{S_n}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2,$$

we have

$$\sum_{v\in V_{S_m}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2 < \sum_{v\in \hat{V}_{S_n}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2.$$

By clustering this implies that $c(V, S_m) < c(\hat{V}, S_n)$, a contradiction. Identical reasoning rules out the case in which

$$c(V, S_m) < c(\hat{V}, S_n) \quad \text{and} \quad \mathrm{RPI}(V, S_m) > \mathrm{RPI}(\hat{V}, S_n).$$

Now consider the case in which $S_m \neq S_n$, and suppose that S_m contains γ_m districts and S_n contains γ_n districts. Consider the following transformation of state n. First, make a λ replica of S_n and a μ replica of S_m so that the number of voters is the same as in the transformed state S_m. Note that $c(V, S_m)$ and RPI (V, S_m) are unchanged because of independence. In a slight abuse of notation we will continue to use V and S_m in reference to the μ-replicated state. Second, expand or contract the state in the sense that the distance between any two points—say, d_{ij}—in state S_n is αd_{ij} in state $S_{n'}$. Note that any partition of state n is a well-defined partition of state $S_{n'}$ as it contains the same voters, scaled by α. Choose α such that

$$\alpha = \frac{|n| \min_{V\in \nu^{\gamma_m}_{S_n}} \sum_{v\in \hat{V}_{S_n}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2}{\mu\, |m| \min_{V\in \nu_{S_m}} \sum_{v\in V_{S_m}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2},$$

where $|n|$ and $|m|$ are the numbers of voters in states S_n and S_m, respectively, and the γ_m superscript denotes a partition into γ_m elements. Note that

$$\min_{V\in \nu_{S_m}} \sum_{v\in V_{S_m}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2 = \min_{V\in \nu^{\gamma_m}_{S_{n'}}} \sum_{v\in V_{S_n}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2. \tag{A6}$$

Third, select a feasible partition of $S_{n'}$ with γ_m elements, and denote this par tition $\hat{V}'$. Suppose that

$$\sum_{v\in \hat{V}'_{S_{n'}}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2 = \theta \sum_{v\in \hat{V}_{S_n}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2$$

and that

$$\min_{V\in \nu^{\gamma_m}_{S_n}} \sum_{v\in \hat{V}_{S_n}} \sum_{i\in v} \sum_{j\in v} f(d_{ij}) = \beta \min_{V\in \nu^{\gamma_n}_{S_n}} \sum_{v\in \hat{V}_{S_n}} \sum_{i\in v} \sum_{j\in v} f(d_{ij}).$$

Hence,

$$\frac{\sum_{v\in \hat{V}'_{S_n}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2}{\min_{V\in \nu^{\gamma_m}_{S_n}} \sum_{v\in \hat{V}_{S_n}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2} = \frac{\theta}{\beta} \frac{\sum_{v\in \hat{V}_{S_n}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2}{\min_{V\in \nu^{\gamma_n}_{S_n}} \sum_{v\in \hat{V}_{S_n}} \sum_{i\in v} \sum_{j\in v} (d_{ij})^2}.$$

By independence,

$$c(\hat{V}', S_{n'}) = \frac{\theta}{\beta} c(\hat{V}, S_n)$$

and

$$\mathrm{RPI}(\hat{V}', S_{n'}) = \frac{\theta}{\beta} \mathrm{RPI}(\hat{V}, S_n).$$

From expression (A5),

$$c(V, S_m) > \frac{\beta}{\theta} c(\hat{V}', S_{n'}) \quad \text{and} \quad \mathrm{RPI}(V, S_m) < \frac{\beta}{\theta} \mathrm{RPI}(\hat{V}', S_{n'}). \tag{A7}$$

But since $S_{m'}$ and $S_{n'}$ have the same number of voters, the same number of districts, and equation (A6) holds, it follows that expression (A7) implies that c violates clustering.

Identical reasoning rules out the case in which

$$c(V, S_m) < c(\hat{V}, S_n) \quad \text{and} \quad \mathrm{RPI}(V, S_m) > \mathrm{RPI}(\hat{V}, S_n),$$

and hence the proof is complete. Q.E.D.

Appendix B

Proofs and Description of the Algorithm

B1. Proof of Theorem 2

Let districts of state S be denoted $D_1, \ldots, D_d$. A districting plan is feasible if $|D_i| = n$ for all $i \in \{1, \ldots, d\}$. The set of feasible districtings is $\mathcal{V}$. Let the centroid of district D_i be m_i, so $m_i = \frac{1}{n}\sum_{x \in D_i}(x)$. Define the functions

$$\psi(D_i) = \sum_{x \in D_i} ||x - m_i||^2, \qquad \Psi(D_1, \ldots, D_d) = \sum_{i=1}^{d} \psi(D_i).$$

We say that districting is optimally compact if it minimizes $\Psi(D_1, \ldots, D_d)$ over all $(D_1, \ldots, D_d) \in \mathcal{V}$. For $z_1, \ldots, z_d \in \mathbb{R}^2$, let

$$\psi_{z_i}(D_i) = \sum_{x \in D_i} ||x - z_i||^2, \qquad \Psi_{z_1, \ldots, z_d}(D_i) = \sum_{i=1}^{d} \psi_{z_i}(D_i).$$

A power diagram with sites $z_1, \ldots, z_d$ is a partition of $\mathbb{R}^2$ into districts $D_1, \ldots, D_d$ such that for fixed constants $\lambda_1, \ldots, \lambda_d \in \mathbb{R}$,

$$D_i = \left\{q \in \mathbb{R}^2 : i = \arg\min_j \left[||q - z_j||^2 - \lambda_j\right]\right\}.$$

It is clear that a power diagram is described by its edges and that if x is on the same side as D_i of any complete set of linear separators between D_i and other districts, then $x \in D_i$, and otherwise not. The edges of D_i are described by the set of $q \in \mathbb{R}^2$ such that $\|q - z_i\|^2 - \lambda_i = \|q - z_j\|^2 - \lambda_j$, or $\|q - z_i\|^2 - \|q - z_j\|^2 = \lambda_i - \lambda_j$.

Lemma 6. The function $\Psi(D_1, \ldots, D_d)$ is proportional to the RPI for $(D_1, \ldots, D_d) \in \mathcal{V}$, so minimizing one is equivalent to minimizing the other. Specifically,

$$\sum_{i=1}^{d} \sum_{x \in D_i} \sum_{y \in D_i} \|x - y\|^2 = 2n \sum_{i=1}^{d} \sum_{x \in D_i} \|x - m_i\|^2.$$

Proof.

$$\begin{aligned}
\sum_{i=1}^{d} \sum_{x \in D_i} \sum_{y \in D_i} \|x - y\|^2 &= \sum_{i=1}^{d} \sum_{x \in D_i} \sum_{y \in D_i} (\|x\|^2 + \|y\|^2 - 2x \times y) \\
&= \sum_{i=1}^{d} \sum_{x \in D_i} \Big(n\|x\|^2 - 2nm_i \times x + \sum_{y \in D_i} \|y\|^2\Big) \\
&= \sum_{i=1}^{d} \Big[\sum_{x \in D_i} (n\|x\|^2 - 2nm_i \times x) + n \sum_{y \in D_i} \|y\|^2\Big] \\
&= \sum_{i=1}^{d} \Big[\sum_{x \in D_i} (2n\|x\|^2 - 2nm_i \times x)\Big] \\
&= \sum_{i=1}^{d} \Big[2n \sum_{x \in D_i} (\|x\|^2 - m_i \times x)\Big] \\
&= \sum_{i=1}^{d} 2n\Big[\sum_{x \in D_i} (\|x\|^2) - n\|m_i\|^2\Big] \\
&= \sum_{i=1}^{d} \Big\{2n\Big[\sum_{x \in D_i} (\|x\|^2 - 2m_i \times x + \|m_i\|^2)\Big]\Big\} \\
&= \sum_{i=1}^{d} \Big[2n\Big(\sum_{x \in D_i} \|x - m_i\|^2\Big)\Big] \\
&= 2n \sum_{i=1}^{d} \sum_{x \in D_i} \|x - m_i\|^2.
\end{aligned}$$

Q.E.D.

Lemma 7. For all $(D_1, \ldots, D_d) \in \mathcal{V}$,

$$(m_1, \ldots, m_d) = \arg \min_{(z_1, \ldots, z_d)} \Psi_{z_1, \ldots, z_d}(D_1, \ldots, D_d).$$

Proof. It suffices to show that substituting m_i for z_i minimizes the expression on the right. Its first-order condition with respect to z_i is

$$\forall D_i,\ 2\sum_{x\in D_i}(x - z_i) = 0 \Rightarrow z_i = \frac{1}{n}\sum_{x\in D_i} x = m_i.$$

Q.E.D.

Lemma 8. In an optimally compact districting, every pair of adjacent districts is separated by a line perpendicular to a line connecting their centroids.

Proof. Let $(D_1, \ldots, D_d)$ be optimally compact. Without loss of generality we can prove the lemma for districts D_1 and D_2. By isometry we can assume that $m_1 = (0, 0)$ and $m_2 = (\xi, 0)$. Pick $v_1 = (x_1, y_1) \in D_1$ and $v_2 = (x_2, y_2) \in D_2$. Let $D_1' = D_1 \cup \{v_2\} - \{v_1\}$ and $D_2' = D_2 \cup \{v_1\} - \{v_2\}$. By the optimality of $(D_1, \ldots, D_d)$ and the optimality lemma,

$$\psi(D_1) + \psi(D_2) \leq \psi(D_1') + \psi(D_2') \leq \psi_{m_1}(D_1') + \psi_{m_2}(D_2')$$

$$\Rightarrow \|v_1 - m_1\|^2 + \|v_2 - m_2\|^2$$

$$\leq \|v_1 - m_2\|^2 + \|v_2 - m_1\|^2$$

$$\Rightarrow -2v_1 \times m_1 - 2v_2 \times m_2$$

$$\leq -2v_1 \times m_2 - 2v_2 \times m_1$$

$$\Rightarrow (v_2 - v_1) \times (m_1 - m_2) \leq 0$$

$$\Rightarrow (x_2 - x_1) \times (-\xi) + (y_2 - y_1) \times 0 \leq 0$$

$$\Rightarrow x_1 \leq x_2.$$

Since v_1 and v_2 are arbitrary, we can pick them such that v_1 is the point in D_1 with greatest x_1 and v_2 is the point in D_2 with least x_2, which shows that there is a line of the form $x = c$ for $c \in \mathbb{R}$ separating the two districts. Isometrics preserve perpendicularity, so applying one moving m_1 and m_2 away from $(0, 0)$ and $(\xi, 0)$ leaves the separator between D_1 and D_2 perpendicular to the segment connecting m_1 and m_2. Q.E.D.

Lemma 9. Let $(D_1, \ldots, D_d)$ be optimal. For every three districts, there exist three concurrent lines, each of which separates two of the three districts, with one line separating each pair of districts.

Proof. Without loss of generality, we prove this lemma for the three districts D_1, D_2, and D_3. By the straight-line lemma, there exist linear separators between D_1 and D_2, D_2 and D_3, and D_3 and D_1 perpendicular to the lines connecting their centroids. We can characterize these lines by the equations $\|r - m_1\|^2 - \|r - m_2\|^2 = \mu_{1,2}$, $\|s - m_2\|^2 - \|s - m_3\|^2 = \mu_{2,3}$, and $\|t - m_3\|^2 - \|t - m_1\|^2 = \mu_{3,1}$ for free variables r, s, $t \in \mathbb{R}^2$. If the lines are concurrent, that means that there exists $q \in \mathbb{R}^2$ satisfying all three equations. Adding them together gives $\mu_{1,2} + \mu_{2,3} + \mu_{3,1} = 0$. Therefore, if the lines are concurrent, then for all r, s, and t on the lines,

$$\|r - m_1\|^2 - \|r - m_2\|^2 + \|s - m_2\|^2 - \|s - m_3\|^2 + \|t - m_3\|^2 - \|t - m_1\|^2 = 0.$$

Assume there is no choice for $\mu_{1,2}$, $\mu_{2,3}$, and $\mu_{3,1}$ such that the lines are concurrent. Then, for all r, s, and t on the three edges,

$$\|r - m_1\|^2 - \|r - m_2\|^2 + \|s - m_2\|^2 - \|s - m_3\|^2 + \|t - m_3\|^2 - \|t - m_1\|^2 \neq 0.$$

If any one of $\mu_{1,2}$, $\mu_{2,3}$, or $\mu_{3,1}$ induces an optimal separator at both the values v_1 and v_2 in $\mathbb{R}^2$, then it must also do so at the value $\lambda v_1 + (1 - \lambda)v_2$ for $\lambda \in [0, 1]$. So the expression above is either strictly greater or strictly less than zero for all permissible values of r, s, and t. We assume without loss of generality that it is greater. Then there exist $v_1 \in D_1$, $v_2 \in D_2$, and $v_3 \in D_3$ such that when they are substituted for r, s, and t, respectively, the above expression reaches a positive infimum. The expression cannot be at an infimum unless the extreme values of r, s, and t are specifically chosen to be in D_1, D_2, and D_3, respectively; otherwise $\|r - m_1\|^2 - \|r - m_2\|^2$, for example, could be decreased by moving r in the direction $m_1 - m_2$ while still separating D_1 and D_2. Therefore,

$$\|v_1 - m_1\|^2 - \|v_1 - m_2\|^2 + \|v_2 - m_2\|^2 - \|v_2 - m_3\|^2 + \|v_3 - m_3\|^2 - \|v_3 - m_1\|^2 > 0 \Leftrightarrow \|v_1 - m_1\|^2 + \|v_2 - m_2\|^2 + \|v_3 - m_3\|^2 > \|v_1 - m_2\|^2 + \|v_2 - m_3\|^2 + \|v_3 - m_1\|^2.$$

Let $D_1' = D_1 \cup \{v_3\} - \{v_1\}$, $D_2' = D_2 \cup \{v_1\} - \{v_2\}$, and $D_3' = D_3 \cup \{v_2\} - \{v_3\}$. Then,

$$\psi(D_1) + \psi(D_2) + \psi(D_3) > \psi_{m_1}(D_1') + \psi_{m_2}(D_2') + \psi_{m_3}(D_3') > \psi(D_1') + \psi(D_2') + \psi(D_3').$$

This contradicts the optimality of $D_1, \ldots, D_d$, and the lemma follows. Q.E.D.

Proof of Theorem 2. We prove that any optimal districting is a power diagram with sites equal to their centroids, $m_1, \ldots, m_d$. For any pair of districts D_i and D_j, we can pick $\mu_{i,j}$ such that $\|q - m_i\|^2 - \|q - m_j\|^2 = \mu_{i,j}$ is a linear separator between the districts, and if we add a third district D_j, we can similarly pick $\mu_{j,k}$ and $\mu_{k,i}$ such that the districting lines are concurrent, or $\mu_{i,j} + \mu_{j,k} + \mu_{k,i} = 0$. Note that $\mu_{a,b} = -\mu_{b,a}$. We prove that there exist constants $\lambda_1, \ldots, \lambda_d$ such that $\lambda_i - \lambda_j = \mu_{i,j}$ by induction. This is obviously true when $n = 2$. Assume that it is true for districts $D_1, \ldots, D_k$. For i, $j < k + 1$,

$$\mu_{i,k+1} = \mu_{i,j} + \mu_{j,k+1} = \lambda_i - \lambda_j + \mu_{j,k+1}$$

$$\Rightarrow \lambda_i - \mu_{i,k+1} = \lambda_j - \mu_{j,k+1}.$$

Thus, $\lambda_i - \mu_{i,k+1}$ is constant over choice of i; call the constant λ_{k+1}. That makes $\mu_{i,k+1} = \lambda_i - \lambda_{k+1}$ for any i, and the induction is complete. Clearly any $x \in D_i$ is on the m_i side of a boundary line between D_i and another district, so it follows that optimal districtings are power diagrams. Q.E.D.

B2. Algorithm Details

The algorithm we propose is a modification of the second algorithm presented in Aurenhammer, Hoffmann, and Aronov (1998). Since we know by theorem 2 that local optima of the RPI are power diagrams, we search within the set of power diagrams for one that is a feasible districting. However, as power diagrams are generated around sites, which we call $z_1, \ldots, z_n$, it is necessary to update the locations of the sites as well as the design of the districts.

First we explain the Aurenhammer, Hoffmann, and Aronov (1998) algorithm for finding a power diagram that minimizes $\Psi_{z_1,\ldots,z_d}(D_1, \ldots, D_d)$, with $|D_i| \approx n$ for all i. Since a power diagram is defined by its sites and their weights, $\lambda_1, \ldots, \lambda_d$, assuming fixed sites each district D_i is a function of $\lambda_1, \ldots, \lambda_d$, or $D_i = D_i(\lambda_1, \ldots, \lambda_d)$. We suppress this dependence for simplicity. Let

$$\xi(\lambda_1, \ldots, \lambda_d) = \sum_{i=1}^{d} (n - |D_i|) \times \lambda_i + \Psi_{z_1,\ldots,z_d}(D_1, \ldots, D_d).$$

Aurenhammer, Hoffmann, and Aronov (1998) simplify the problem by continuing as if each D_i does not change locally with respect to each λ_i everywhere, as this is true almost everywhere (at all but finitely many points). Therefore, $|D_i|$ and $\Psi_{z_1,\ldots,z_d}(D_1, \ldots, D_d)$ are locally constant with respect to λ_i, so

$$\frac{\partial \xi}{\partial \lambda_i} = n - |D_i|.$$

Let $\Lambda = (\lambda_1, \ldots, \lambda_d)$. Using some choice of Λ_0, we can update it by gradient descent:

$$\Lambda_{t+1} = \Lambda_t + \varepsilon_t \times \nabla\xi(\Lambda_t).$$

In our implementation we set Λ_0 to be the zero vector. It remains to pick the step sizes $\{\varepsilon_t\}_{t\geq 0}$. To do this, one first determines an overestimate of the minimum value of ξ; call it $\bar{\xi}$. This can be done by setting $\bar{\xi} = \Psi_{z_1,\ldots,z_d}(D_1, \ldots, D_d)$ for any feasible districting $(D_1, \ldots, D_d)$. We use the notation $D_i(\Lambda_t)$ to mean one of the districts induced by the power diagram weights contained in the vector Λ_t, and let

$$\varepsilon_t = \frac{\bar{\xi} - \xi(\Lambda_t)}{\sum_{i=1}^{d} |D_i(\Lambda_t)|^2}.$$

This step size is iterated until the minimum is either reached or missed, which happens when $\sum_{i=1}^{d} |D_i(\Lambda_t)| \times |D_i(\Lambda_{t+1})| > 0$. Then $\bar{\xi}$ is updated by solving the equation

$$\frac{\bar{\xi} - \xi(\Lambda_t)}{\sum_{i=1}^{d} |D_i(\Lambda_t)|^2} = \frac{\bar{\xi} - \xi(\Lambda_{t+1})}{\sum_{i=1}^{d} |D_i(\Lambda_{t+1})|^2}.$$

The size ε_{t+1} is chosen accordingly. This algorithm is repeated until the $|D_i|$'s are within some predetermined error bound around n.

Once optimal districts $D_1, \ldots, D_d$ for sites $z_1, \ldots, z_d$ are chosen, by lemma 7 (see Appendix Section B1) the function $\Psi_{z_1, \ldots, z_d}(D_1, \ldots, D_d)$ is improved by moving the z_i's to the centroids of the D_i's and keeping the $\lambda_1, \ldots, \lambda_d$ constant. Yet not all of the D_i's are necessarily of size n, so they need to be adjusted by the above procedure. This process is repeated until moving the $z_1, \ldots, z_d$ still leaves the sizes of the D_i's within the prescribed error bound.

Note that the algorithm described in Aurenhammer, Hoffmann, and Aronov (1998) tends to fail when one of the districts is randomly set to zero. Our solution to this issue was to move z_i to a random new location if $|D_i|$ became zero during any point in the process. Random new locations were chosen using a uniform distribution function ranging from the minimum to the maximum of the longitude and the latitude of the state in question.

Appendix C

A Guide to Programs

All programs to compute feasible districtings minimizing the RPI are written for Matlab. There are two main programs, Main.m and Compute_Index.m, and support programs District.m, getRandGP.m, Psi.m, Weighted_Assign.m, Weighted_FirstTryAssign.m, and Weighted_PowerDiagram.m. We briefly describe each of the main programs below.

Main.m and Compute_Index.m are both shell programs that call District.m, the actual algorithm, and store its output in text files. Typing Compute_Index(File Name, Iterations) reads demographic data about a state from a text file—say, "indiana.out"—and creates a new districting Iterations times. The file should have the latitudes and longitudes of the census tracts of the states in columns 2 and 3, respectively, the federal information processing standards (FIPS) code of the state repeated in every entry of column 4, the current districts of all census tracts in column 5, and the populations of all census tracts in column 6. Compute_Index.m generates two output files. The first, in this case "indiana.out .output," contains the latitudes and longitudes of the census tracts in the first two columns and their new district numbers in the subsequent columns. Each column after the second represents a different iteration of the algorithm. The second output file, in this case "indiana.out.stats," contains statistics from each iteration of the algorithm on a different row. The first column has the RPIs, the second has the accuracy of the districting, and the third has the accuracy of the current districting. Accuracy is measured as

$$\max_{i \in 1, \ldots, d} \left| \frac{|D_i| - n}{n} \right| .$$

Compute_Index.m has the following hard-coded parameters that are passed to District.m: outside_tol_ratio, tol_ratio, outside_bail, and bail. The parameters tol_ratio and bail are the stopping criteria for the subroutine Weighted_Assign.m, which creates the best districting around randomly initiated sites. If the accuracy falls below tol_ratio or the number of iterations of the gradient-descent procedure rises above bail, the algorithm terminates. Likewise, outside_tol_ratio and outside_bail are the stopping criteria for the larger districting algorithm. If the accuracy of the districting falls below outside_tol_ratio or the number of times the sites are moved rises above outside_bail, the algorithm terminates. The set values for outside_tol_ratio, tol_ratio, outside_bail, and bail are, respectively, .9 times the real accuracy, whichever is the lesser of .9 times the real accuracy or .05, 35 times the number of districts in the state, and 35 times the number of districts in the state.

Main(File Name) reads a list of states and iterations for each state to be run by Compute_Index.m. The file is of the following form:

states	bootstraps
alabama	4
arizona	7
arkansas	3
california	1

Names of states and numbers of iterations are separated by tabs. If "arizona" is written in this file, Compute_Index.m will open a file called "arizona.out." Main.m creates an additional file called "index.txt" that lists the FIPS code for every state next to the best RPI the algorithm has found for it such that the accuracy for the districting corresponding to that RPI is better than the state's current accuracy.

This procedure yields an RPI greater than one and an accuracy better than the current accuracy nearly all of the time for all states other than Connecticut, Idaho, Minnesota, and Nebraska, which already are well districted and usually require quite a few bootstraps to improve on the current districting.

Appendix D

Congressional District Map Comparisons for the 106th Congress

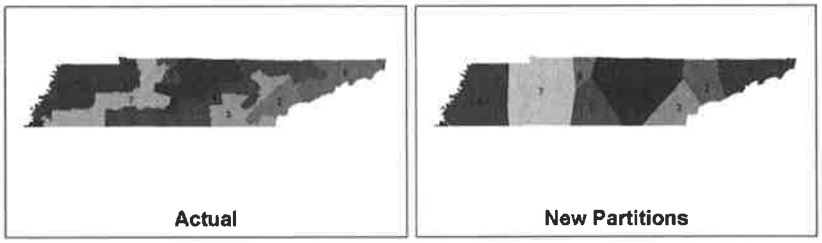

Figure D1. Tennessee

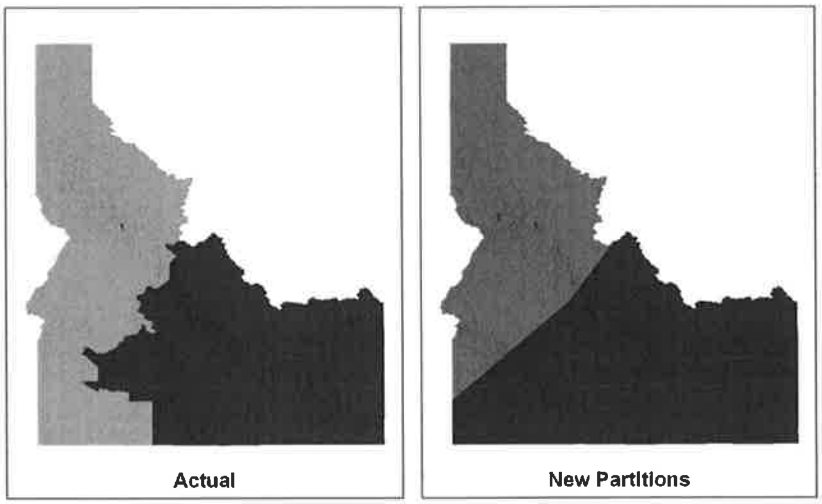

Figure D2. Idaho

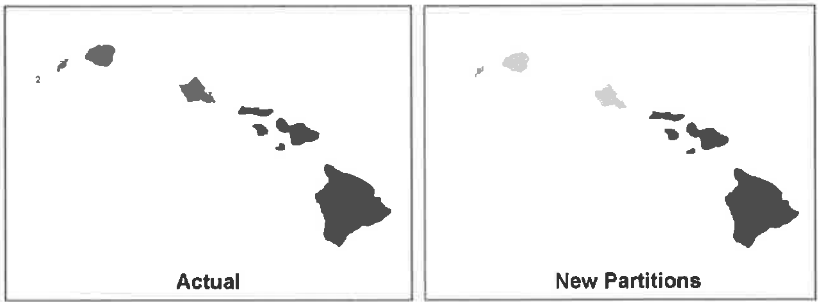

Figure D3. Hawaii

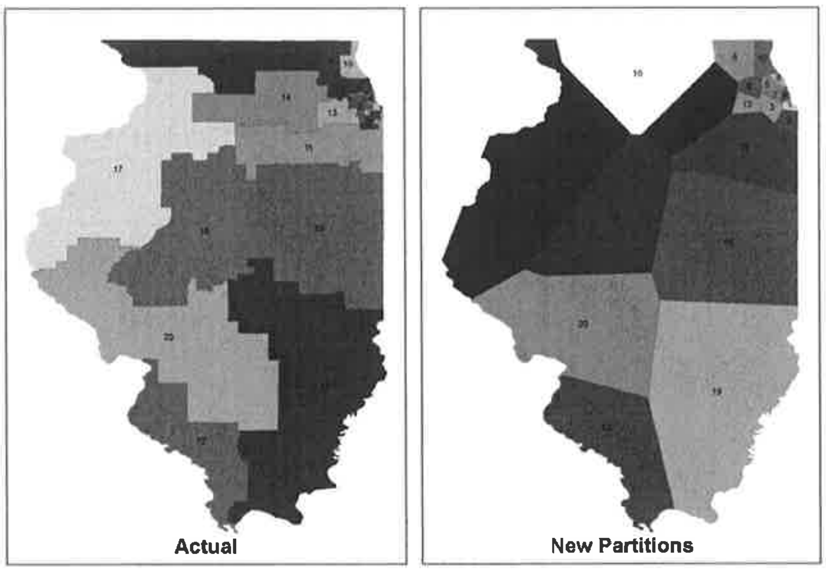

Figure D4. Illinois

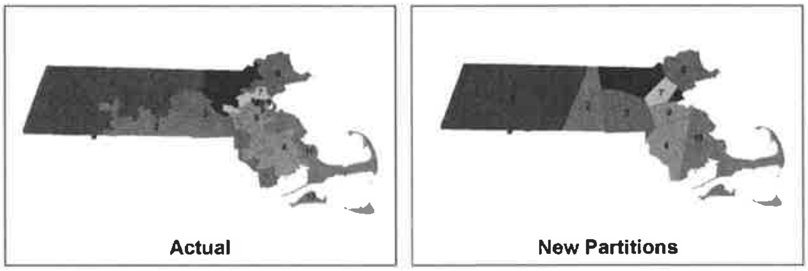

Figure D5. Massachusetts

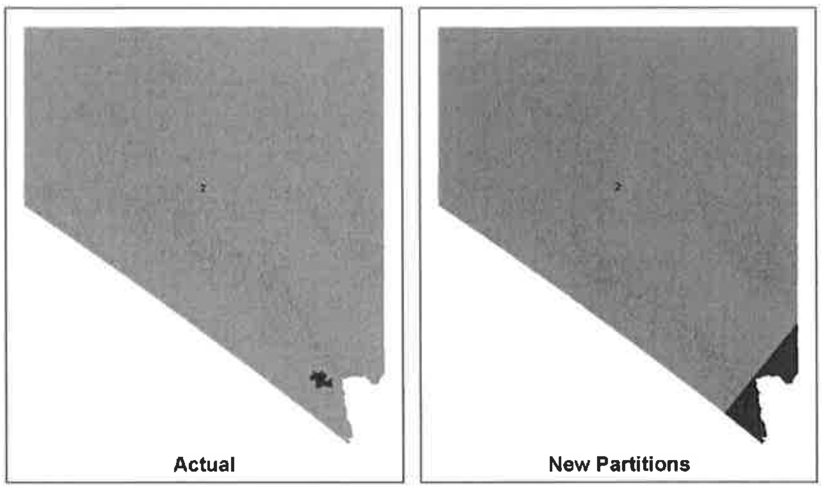

Figure D6. Nevada

Figure D7. New York

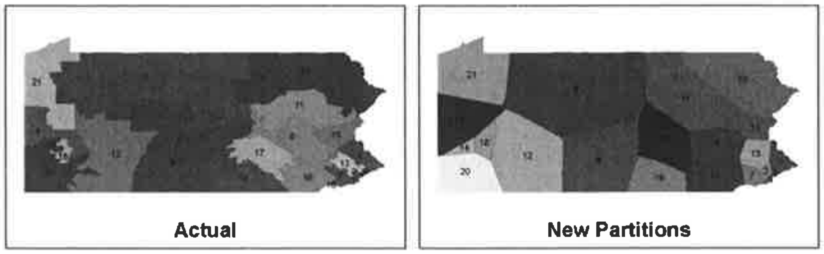

Figure D8. Pennsylvania

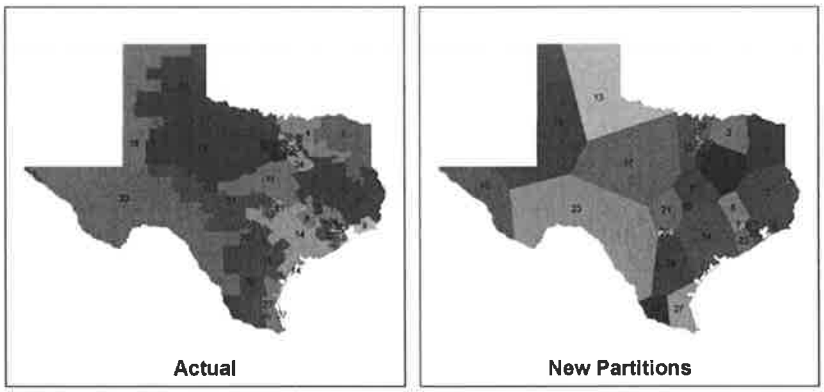

Figure D9. Texas

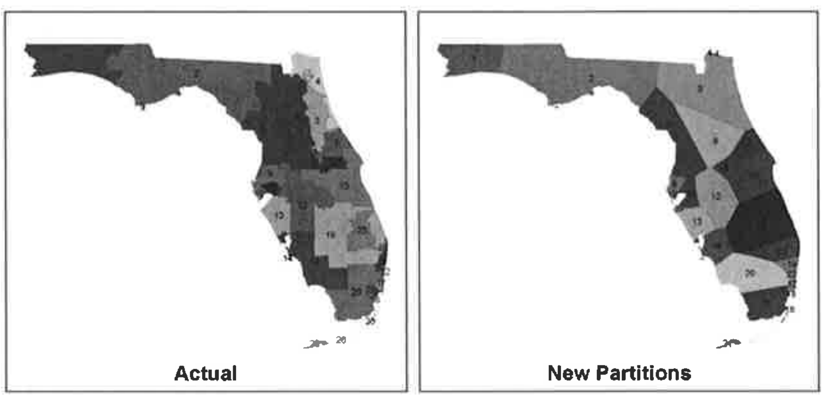

Figure D10. Florida

Appendix E

Comparison of Actual and Maximally Compact Seat-Vote Curves

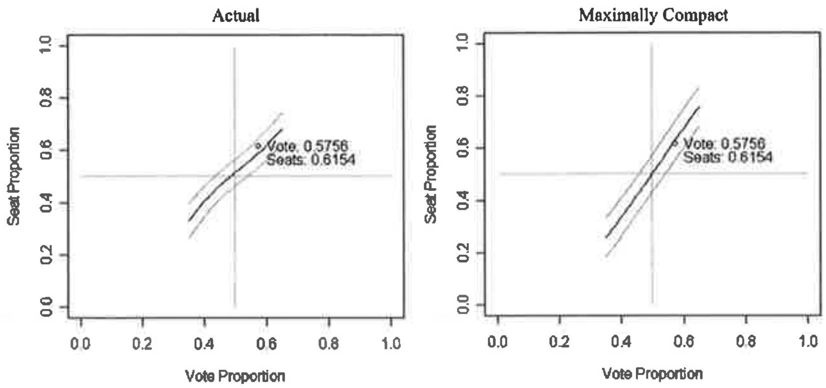

Figure E1. California

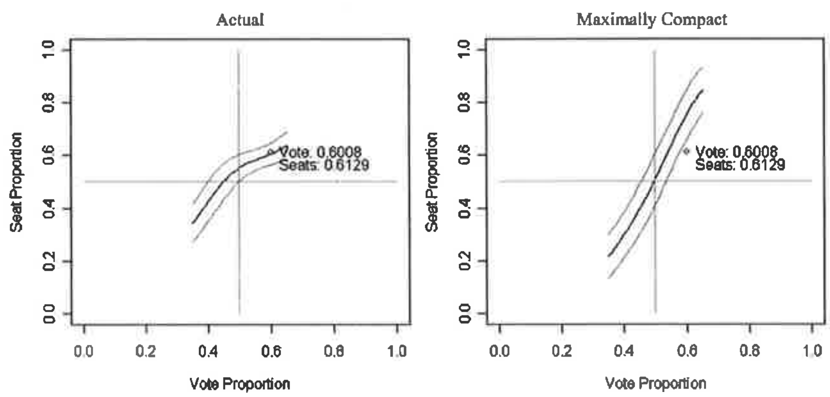

Figure E2. New York

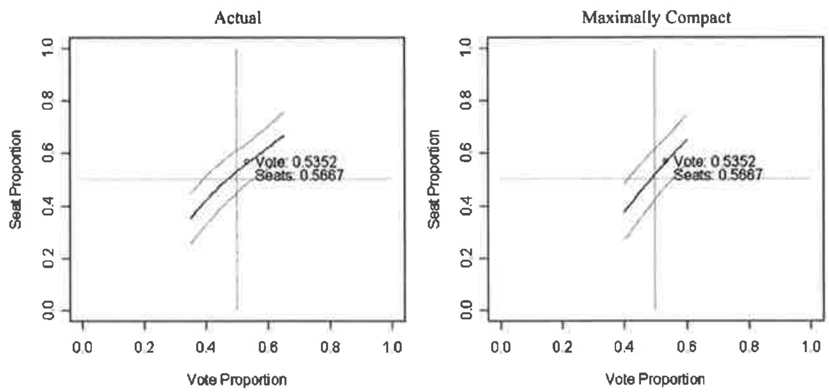

Figure E3. Texas

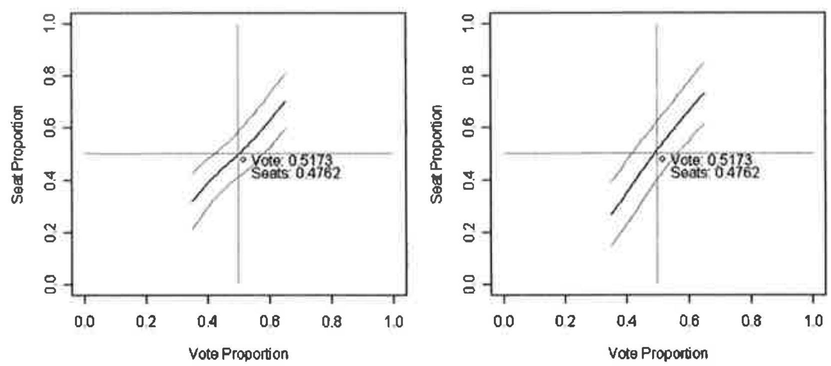

Figure E4. Pennsylvania

References

Adams, Bruce. 1977. A Model State Reapportionment Process: The Continuing Quest for "Fair and Effective Representation." *Harvard Journal on Legislation* 14:825–904.

Arrow, Kenneth. 1970. *Social Choice and Individual Values.* Cowles Foundation Monograph Series. New Haven, Conn.: Yale University Press.

Attneave, Fred, and Malcolm D. Arnoult. 1956. The Quantitative Study of Shape and Pattern Perception. *Psychological Bulletin* 53:452–71.

Aurenhammer, Franz. 1987. Power Diagrams: Properties, Algorithms and Applications. *SIAM Journal on Computing* 16:78–96.

Aurenhammer, Franz., F. Hoffmann, and B. Aronov. 1998. Minkowski-Type Theorems and Least-Squares Clustering. *Algorithmica* 20:61–76.

Aurenhammer, Franz, and Rolf Klein. 2000. Voronoi Diagrams. Pp. 201-90 in *Handbook* *of Computational Geometry*, edited by J. Sack and G. Urrutia. Amsterdam: Elsevier Science.

Barabas, Jason, and Jennifer Jerit. 2004. Redistricting Principles and Racial Representation. *State Politics and Policy Quarterly* 4:415–35.

Bartal, Yair, Moses Charikar, and Danny Raz. 2001. Approximating min-sum k-Clustering in Metric Spaces. Pages 11–20 in *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*. New York: Association for Computing Machinery.

Boyce, Ronald R., and W. A. V. Clark. 1964. The Concept of Shape in Geography. *Geographical Review* 54:561–72.

Cox, E. P. 1927. A Method of Assigning Numerical and Percentage Values to the Degree of Roundness of Sand Grains. *Journal of Paleontology* 1:179–83.

Diaconescu, D. E., B. Florea, and A. Grassi. 2002. Geometric Transitions, del Pezzo Surfaces and Open String Instantons. *Advanced Theoretical and Mathematical Physics* 6:643–702.

Eig, Larry M., and Michael V. Seitzinger. 1981. *State Constitutional and Statutory Provisions Concerning Congressional and State Legislative Redistricting*. Report No. 81–143A. Washington, D.C.: Congressional Research Service.

Geisler, R. Gene. 1985. Measuring Gerrymandering of Legislative Districts. Unpublished manuscript.

———. 1994. A Unified Method of Evaluating Electoral Systems and Redistricting Plans. *American Journal of Political Science* 38:514–54.

Gibbs, J. P. 1961. A Method for Comparing the Spatial Shape of Urban Units. Pp. 99–106 in *Urban Research Methods*, edited by J. P. Gibbs. New York: Van Nostrand.

Harris, Curtis C., Jr. 1964. A Scientific Method of Districting. *Behavioral Science* 9:219–25.

Hofeller, Thomas, and Bernard Grofman. 1990. Comparing the Compactness of California Congressional Districts under Three Different Plans: 1980, 1982, and 1984. Pp. 281–88 in *Political Gerrymandering and the Courts*, edited by Bernard Grofman. New York: Agathon.

Horton, R. E. 1932. Drainage Basin Characteristics. *Transactions of the American Geophysical Union* 13:350–61.

Ji, Xiaoyun, and John E. Mitchell. 2005. Finding Optimal Realignments in Sports Leagues Using a Branch-and-Cut-and-Price Approach. *International Journal of Operations Research* 1:101–22.

Kaiser, Henry F. 1966. An Objective Method for Establishing Legislative Districts. *Midwest Journal of Political Science* 10:200–213.

Knight, Jeffrey L. 2004. GIS Based Compactness Measurement Using Fractal Analysis. Department of Economics, Brown University, Providence, R.I.

Mandelbrot, Benoit B. 1967. How Long Is the Coast of Britain? Statistical Self-Similarity and Fractal Dimension. *Science* 156:636–38.

Mehrotra, Anuj, Ellis L. Johnson, and George L. Nemhauser. 1998. An Optimization Based Heuristic for Political Districting. *Management Science* 44:1100–14.

Mitchell, John E. 2003. Realignment in the National Football League: Did They Do It Right? *Naval Research Logistics* 50:683–701.

Mumford, David, and Jayant Shah. 1989. *Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems*. Cambridge, Mass.: Harvard University Press.

Niemi, Richard G., Bernard Grofman, Carl Carlucci, and Thomas Hofeller. 1990. Measuring Compactness and the Role of a Compactness Standard in a Test for Partisan and Racial Gerrymandering. *Journal of Politics* 52:1155–81.

Papayanopolous, L. 1973. Quantitative Principles Underlying Legislative Apportionment. *Annals of the New York Academy of Sciences* 19:181–91.

Passare, Mikael, and Hans Rullgard. 2004. Amoebas, Monge-Ampere Measures, and Triangulations of the Newton Polytope. *Duke Mathematics Journal* 121:481–507.

Pounds, Norman J. G. 1972. *Political Geography.* 2nd ed. New York: McGraw-Hill.

Reock, Ernest C., Jr. 1961. Measuring Compactness as a Requirement of Legislative Apportionment. *Midwest Journal of Political Science* 5:70–74.

Richter-Gebert, Jurgen, Bernd Sturmfels, and Thorsten Theobald. 2003. First Steps in Tropical Geometry. Unpublished manuscript. Technische Universität München, Zentrum Mathematik, Munich. http://arxiv.org/abs/math/0306366.

Sahni, S., and T. Gonzales. 1976. P-Complete Approximation Problems. *Journal of the ACM* 23:555–66.

Schwartzberg, Joseph E. 1966. Reapportionment, Gerrymanders, and the Notion of Compactness. *Minnesota Law Review* 50:443–52.

Shotts, Kenneth W. 2002. Gerrymandering, Legislative Composition, and National Policy Outcomes. *American Journal of Political Science* 46:398–414.

Taylor, Peter J. 1973. A New Shape Measure for Evaluating Electoral District Patterns. *American Political Science Review* 67:947–50.

Theobald, H. Rupert. 1970. Equal Representation: A Study of Legislative and Congressional Appointment in Wisconsin. Pp. 71–260 in *The State of Wisconsin Blue Book,* edited by H. Rupert Theobald and Patricia V. Robbins. Madison: Wisconsin Legislative Reference Bureau.

Weaver, James B., and Sidney W. Hess. 1963. A Procedure for Nonpartisan Districting: Development of Computer Techniques. *Yale Law Journal* 72:288–308.

Wells, David I. 1982. How to Inhibit Gerrymandering: The Purpose of Districting Criteria. *National Civic Review* 71:183–87.

Young, H. P. 1988. Measuring the Compactness of Legislative Districts. *Legislative Studies Quarterly* 13:105–15.

Research & Politics
http://rap.sagepub.com/

Gerrymandering or geography? How Democrats won the popular vote but lost the Congress in 2012

Nicholas Goedert

Research & Politics 2014 1:
DOI: 10.1177/2053168014528683

The online version of this article can be found at:
http://rap.sagepub.com/content/1/1/2053168014528683

Published by:
SAGE
http://www.sagepublications.com

EXHIBIT 20
WIT: Goedert
DATE: 12-15-15
Halma-Jilek Reporting, Inc.

Research Article

Research and Politics
April-June 2014: 1–8

DOI: 10.1177/2053168014528683
rap.sagepub.com
SAGE

Gerrymandering or geography? How Democrats won the popular vote but lost the Congress in 2012

Nicholas Goedert

Abstract

This article assesses whether the antimajoritarian outcome in the 2012 US congressional elections was due more to deliberate partisan gerrymandering or asymmetric geographic distribution of partisans. The article first estimates an expected seats–votes slope by fitting past election results to a probit curve, and then measures how well parties performed in 2012 compared to this expectation in each state under various redistricting institutions. I find that while both parties exceeded expectations when controlling the redistricting process, a persistent pro-Republican bias is also present even when maps are drawn by courts or bipartisan agreement. This persistent bias is a greater factor in the nationwide disparity between seats and votes than intentional gerrymandering.

Keywords

Congress, legislative elections, gerrymandering, 2012 American elections

Leading into the 2012 general election in the United States, much of the media's prognostication focused on the possibility that President Barack Obama might win reelection with a majority of the Electoral College yet a minority of the popular vote. In retrospect, Obama won a comfortable popular vote victory, but the same election saw a parallel "antimajoritarian" outcome in the House Representatives: Republicans won just 49.4% of the aggregated two-party vote and yet won 54% of the seats.

On the surface, Republican partisan gerrymandering appears to explain this disparity. The argument that Democrats underperformed in their seat share due to Republican control of redistricting in many large states is relatively simple. Firstly, it is certainly true that Republicans controlled this process in more states, representing more seats. In addition, in each of these states, Democrats won fewer seats than any reasonable allocation of the popular vote would suggest was "fair." For example, Republicans won a large majority of the seats in Pennsylvania, North Carolina, and Michigan, despite losing the mean popular vote by district in each state.

However, the problem for Democrats might actually be more fundamental: the current geographic distribution of partisans now leaves Democrats at a disadvantage so long as congressional representation is based on contiguous geographic districts. It is unsurprising that Republicans won more than their fair share of seats in states where they drew the maps. However, Democrats also underperformed under bipartisan maps, and gained only small advantages from their own maps, suggesting their main issue is not gerrymandering, but districting itself.

The observation that Republicans appear to have a natural advantage in the geographic dispersion of their voters is not just a recent one. Erikson studied this phenomenon in northern districts in the 1960s, concluding that "the tendency toward a Republican gerrymander in the distribution of constituency vote" was "the 'natural' state of affairs" and "more an accident of geography than the intentional creation of Republican legislatures" (Erikson, 1972: 1241–1243).

In the 1970s, this bias seemed to reverse to the benefit of Democrats, largely due to overwhelming Democratic control of districting in the South (see e.g. Brunell, 1999; McGhee, 2012). In recent years, however, Erikson's thesis has received

Washington University in St. Louis, USA

Corresponding author:
Nicholas Goedert, Postdoctoral Research Associate, Washington University in St. Louis, Campus Box 1063, One Brookings Drive, St. Louis, MO 63130, USA.
Email: ngoedert@wustl.edu

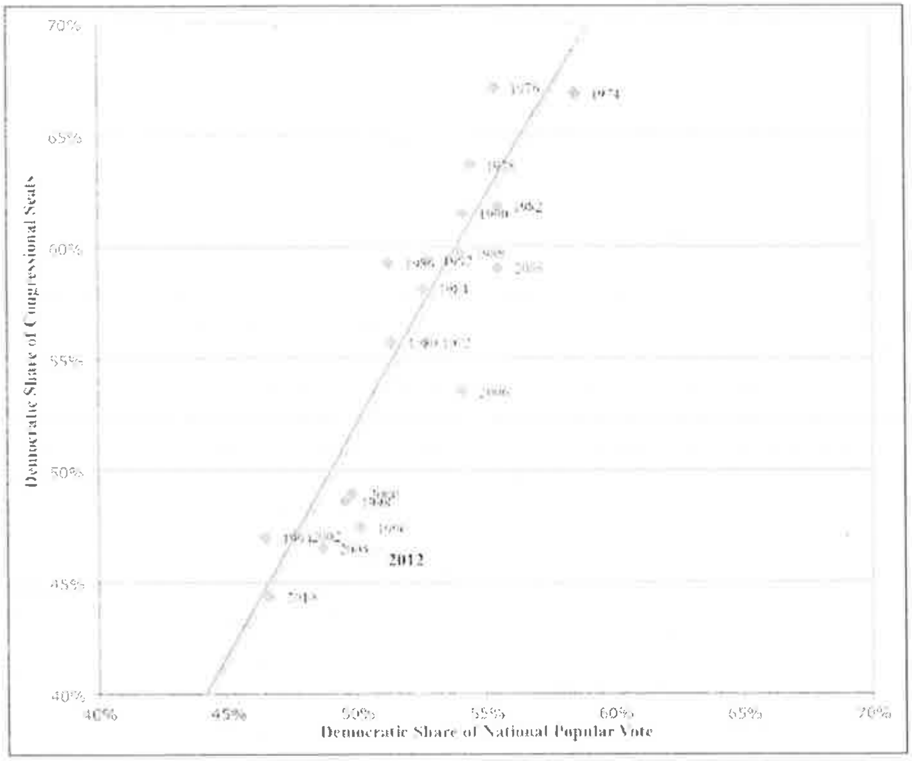

Figure 1. Seats–votes curve in US congressional elections, 1972–2012.

renewed attention. Hirsch, for example, examines the 2000 redistricting cycle and asserts that "Democratic concentrations in urban areas make it easier for Republicans to gerrymander successfully…[and] relatively harder for Democrats to gerrymander successfully" (Hirsch, 2003: 196).[1] Chen and Rodden (2013) use random districting simulations of Florida and other states to argue that the Democratic Party is disadvantaged even under neutral districting methods, tracing this bias back to urban population shifts during the industrial revolution. In addition, through a case study of several ideologically neutral proposals to redistrict Virginia, Altman and McDonald conclude "there may be some modest truth to the claim that urban Democrats are inefficiently concentrated within their urban communities from a redistricting standpoint" (Altman and McDonald, 2013: 830).

Several recent trends, however, might cast doubt on the lasting relevance of Erikson's assertion. These include more sophisticated and varied redistricting institutions and tools and changing demographic patterns, particularly the dramatic rise in Hispanic population. This note takes a first cut at adjudicating this question as applied to the 2012 election results.

Estimating the seats–votes curve

To assess the bias in maps of individual states, we must first establish how a "fair" map might translate the popular vote for individual candidates into seats. It has been almost universally observed that electoral systems employing single-member districts yield seat majorities that exaggerate vote majorities (Lijphart, 1999; McDonald, 2009; Rae, 1967). To the extent that this exaggeration is not biased to favor one party, it is often seen as a feature of such systems rather than a bug, creating governing mandates out of what would otherwise be the confusion of unstable plurality coalitions. The exaggeration tends to take the shape of a probit or logit function, although the slope (i.e. the sensitivity) of the curve has been found to vary widely among electoral systems (e.g. King and Browning, 1987; Taagepera and Shugart, 1989; Tufte, 1973).

Tufte (1973) proposed that a system of districting must pass two tests to be "minimally democratic." Firstly, it must be *responsive* such that an increase in votes for one party will translate into an increase in seats, and secondly, it must be *unbiased* in treating both parties alike. We therefore start from the premise that a fair assignment of seats to parties will be not be biased in favor of one party, but also will not require proportional representation. Rather, we will assume that a party should expect to win a proportion of seats in line with historical patterns found in modern congressional elections.

The "fair expectation" for seats given a vote share is thus estimated by imputing a responsiveness slope that is average for all congressional elections since the nationwide implementation of equal-population districts. Figure 1 shows the relationship between national vote share and seats won in congressional elections since 1972, as well as a fit line using both probit (solid line) and ordinary least squares (OLS; dashed line). Within the observed range, these two methods yield almost identical results, indicating that a 1% increase in vote share will produce about a 2% increase in seat share. Thus, winning 55% of the vote will generally yield about 60% of the seats.[2] The estimated 2012 result (not included in the fit line) falls far below this line, demonstrating the Democrats' underperformance compared with historical averages.

The probit curve has a slope coefficient of 0.026, representing responsiveness, and a constant of −0.040 (where the independent variable is the Republican percentage point advantage in the aggregated popular vote, and the dependent variable is share of seats won). This coefficient of 0.026 is used throughout the analysis to represent the "expected" responsiveness of the seats–votes curve, equivalent to the ρ term in King and Browning's (1987) model.[3]

In lieu of using national election data to measure the responsiveness of congressional seats to votes, we can alternately estimate this slope using state-by-state election data from the same 1972–2010 period, using mean two-party vote share by district as the independent variable, and statewide seat share as the dependent variable, similar to the 2012 results presented below. This method (detailed in Table A1 of Supplementary Material) yields a slope coefficient of 0.0234. In addition, unopposed races in the South, particularly in the first two decades, distort this result: the coefficient estimate is 0.0271 if the South is excluded.[4] Using this method, we can also include fixed effects for decade, none of which are significant. Although the bias in congressional maps appears to vary over time, there is little variation in responsiveness, either within this period or when comparing the last 40 years to earlier decades in the 20th century. Imputing the lowest slope value under this method (0.0234) still yields substantively very similar results (shown in Table A2 of Supplementary Material).

Methodology for vote share and seat share

Drawing on the 2012 election results, I have calculated each party's mean vote share across each state's congressional districts, using mean rather than the aggregate share so that each district is weighted equally regardless of turnout and unopposed races can be included. Where a candidate ran completely unopposed, I have assigned that candidate's party 100% of the vote; where a candidate ran against only minor parties, I have assigned the opposing party the vote share of the minor candidates. I then compare the mean vote share with the expected seat share under a "fair" map with zero bias and a historically average seats–votes curve. For example, Michigan Democrats won a mean vote share of 53%, which, when we apply the slope estimate above, translates into winning 56% of seats. In actuality, however, Democrats won only 5 of Michigan's 14 seats (36%), 20% less than the expected number of seats in that state.

Each state is coded for redistricting control by Republicans, Democrats, or some other institution (e.g. commission, court, bipartisan agreement) to assess whether Republicans exceeded their expected seat share more when they controlled the redistricting process. Table 1 shows bias results for five categories of states, with negative numbers in the last column indicating the degree of pro-Republican bias. The first three subheads show states with at least six congressional districts with maps drawn by Republicans, Democrats, and bipartisan agreement/courts, respectively, while the last two subheads show states with the largest Hispanic populations and those in the Deep South, categories that will be analyzed separately.

Seats won versus seats expected by redistricting control

If the overall pro-Republican bias in the national election outcome was due predominantly to Republicans controlling the districting process in more states, we should expect to observe opposing biases of similar magnitudes in individual states when Republicans and Democrats controlled the process. In addition, we would expect little or no bias in states where maps were drawn by courts or bipartisan agreement. At first glance, neither of these hypotheses seems true.

In every state districted by Republicans, Democrats won fewer seats than their historical expectation, and in six cases they underperformed by 20% or more. It appears as though Republicans gained dramatic benefits across the board from holding the reins of districting.

In contrast, Democrats only slightly exceeded their expected seat share in the three states—Illinois, Massachusetts, and Maryland—where they controlled the process, gaining just a fractional seat above expectation in each. For instance, Illinois Democrats won a smaller majority in their delegation than Republicans won in Pennsylvania or Ohio, despite winning a much larger vote share. Although winning all of Massachusetts' nine districts may seem a wildly inequitable distribution, by winning 76% of the mean vote Massachusetts Democrats could expect to win 91% of the seats under a "fair" map. If John Tierney had won 1% less in his MA-6 race, Democrats would have slightly underperformed their expected share.[5] While Democrats underperformed by an average of 19% under Republican gerrymanders, they only exceeded expectation by 5% under these Democratic gerrymanders.

In addition, we observe bias even where we should expect none in the redistricting process. Democrats also fell short of expectation in several states with bipartisan or court-drawn maps. For example, despite a constitutional

Table 1. Seats won versus mean vote share by gerrymandering party: 2012 congressional elections.

Republican gerrymanders

State	CDs	Dem. Vote share	Dem. Seats won	Dem. seats Expected	Won–Exp. Difference
Indiana	9	46%	22%	42%	−20%
Michigan	14	53%	36%	56%	−20%
Missouri	8	43%	25%	36%	−11%
North Carolina	13	51%	31%	52%	−21%
Ohio	16	48%	25%	46%	−21%
Pennsylvania	18	51%	28%	51%	−23%
Tennessee	9	38%	22%	28%	−6%
Virginia	11	49%	27%	48%	−21%
Wisconsin	8	51%	38%	52%	−15%
Total	**106**	**48%**	**28%**	**47%**	**−19%**

Democratic gerrymanders

State	CDs	Dem. Vote share	Dem. Seats won	Dem. seats Expected	Won–Exp. Difference
Illinois	18	56%	67%	63%	4%
Massachusetts	9	76%	100%	91%	9%
Maryland	8	64%	88%	76%	11%
Total	**35**	**63%**	**80%**	**75%**	**5%**

Bipartisan or court gerrymanders

State	CDs	Dem. Vote share	Dem. Seats won	Dem. seats Expected	Won–Exp. Difference
Colorado	7	50%	43%	50%	−7%
Florida	27	48%	37%	46%	−9%
Kentucky	6	39%	17%	29%	−12%
Minnesota	8	57%	63%	63%	−1%
New Jersey	12	57%	50%	65%	−15%
New York	27	67%	78%	81%	−3%
Washington	10	53%	60%	57%	3%
Total	**97**	**55%**	**54%**	**61%**	**−7%**

High Hispanic population states

State	CDs	Dem. Vote share	Dem. Seats won	Dem. seats Expected	Won–Exp. Difference
Arizona	9	48%	56%	45%	10%
California	53	60%	72%	70%	2%
New Mexico	3	54%	67%	59%	8%
Nevada	4	51%	50%	53%	−3%
Texas	36	43%	33%	37%	−3%
Total	**105**	**53%**	**56%**	**56%**	**0%**

Deep South states

State	CDs	Dem. Vote share	Dem. Seats won	Dem. seats Expected	Won–Exp. Difference
Alabama	7	35%	14%	21%	−7%
Georgia	14	39%	36%	28%	8%
Louisiana	6	32%	17%	20%	−4%
Mississippi	4	39%	25%	29%	−4%
South Carolina	7	41%	14%	31%	−17%
Total	**38**	**37%**	**24%**	**26%**	**−2%**

amendment prohibiting Republican legislators from using partisanship to draw maps in Florida, the GOP nevertheless managed to win 17 seats with 51.4% of the vote, surpassing expectation by 2.5 seats. Even under bipartisan gerrymandering in New York, in which Democrats won 21 of 27 seats, their vote share suggested they should have won 22. Across the seven states with bipartisan or court gerrymanders, Republicans exceeded expectation by an average of 7%.[6]

So how many seats did this underlying disadvantage cost the Democrats? If we imagine that these bipartisan or court maps were unbiased, and that Democrats and Republicans received equal benefit from their own maps (for example, a 12% advantage as an average), this would have yielded 16 or 17 additional seats, likely getting the Democrats within a couple seats of the majority. By contrast, the disparity between the number of seats gerrymandered by Republicans compared to Democrats likely costs Democrats about nine seats.[7] This initial analysis reveals that geography is a slightly greater factor than intentional gerrymandering in explaining why Democrats won fewer seats than expected from their vote share.

If there is any area of the country where the geographic distribution of partisans has *not* led to an underrepresentation of Democrats, we might expect to observe it where Democratic voting strength does not hew as closely to the black/white or urban/rural divide. In particular, we find this pattern interrupted in areas with very large Hispanic populations, as Hispanics tend to be both less saturated in their support for Democrats and more geographically dispersed than African-Americans living in large urban areas. In the five states with the highest proportion of Hispanics (Arizona, California, New Mexico, Nevada, and Texas), Democrats won a seat share very close to expectation in each state, despite not controlling the process in any of them. It is possible that non-partisan commissions in California and Arizona may have contributed to greater fairness, but the ease of drawing geographically large, majority Hispanic districts in these states, (e.g. AZ-4, CA-16, CA-51, and TX-23) might have also mitigated the advantage Republicans have in other regions given the distribution of their voters.

The final subhead of Table 1 depicts results from five states in the Deep South. In these states, voting is highly racial polarized and, unlike most of the rest of the nation, much of the African-American population is rural. In addition, amendments to the Voting Rights Act (VRA) have been interpreted to require the drawing of African-American-majority or African-American-influence districts across rural parts of these states, with district maps requiring Department of Justice preclearance under the VRA. Past research has suggested that this may constrain maps to resemble Republican gerrymanders even when drawn by another party (Goedert, 2012; Hill, 1995; Lublin, 1999), and we do see that results in these states are slightly biased against Democrats with one exception.[8] Because we therefore might expect these states to be much differently impacted by both urbanization and the gerrymandering party compared to the rest of the nation, they are excluded from the regression analysis below.

Regression results

To more directly approach Chen and Rodden's (2013) argument that Democrats are disadvantaged due to their heavy concentration in cities, I analyzed these results using an OLS regression, including 2010 US Census data on race and urbanization. Table 2 depicts regression results with each state weighted by number of districts, excluding five Deep South states and states with only one or two districts. The dependent variable is the difference between Democratic seats won and the number of seats expected given their vote share. A high positive value is a map distorted in favor of Democrats, while a high negative value is a map distorted in favor of Republicans. Dummy variables are assigned for partisan redistricting procedures; the excluded category is bipartisan or court-drawn maps. In addition, controls are included in some models for the percent of the population that lives in urban areas or that is African-American or Hispanic. The "Hispanic Dummy" in Model 1 is a "1" for the five most heavily Hispanic states.

Model 1 reaffirms the three central conclusions from Table 1. Firstly, the effect of partisan control of the districting process is significant and in the expected direction. Secondly, as we can see from the negative and significant constant, which captures the bias in states with a bipartisan or court-drawn map and without a large Hispanic population, maps are distorted in favor of Republicans even when we control for partisan gerrymanders. Finally, this distortion is not present in the case of the most heavily Hispanic states.

Model 2 tests the effect of minority population proportions, includes controls for state size and overall partisanship of the state, and also yields a closer test of the Chen and Rodden (2013) hypothesis by including the urbanization variable. Chen and Rodden hypothesize that the distortion is due to population shifts toward urban areas. If this were true, we would expect more distortion against Democrats in heavily urbanized states. Consistent with Table 1 and Model 1, a larger Hispanic population reduces bias against Democrats, but the size of the African-American population has no significant effect on distortion, and we see no effect for urbanization.[9]

Model 3, including only states with more than six districts, paints a different picture, showing a significant negative coefficient for urbanization. Among larger states, which likely include both urban and rural areas, heavily urbanized states (e.g. New Jersey and Pennsylvania) are more often heavily distorted against Democrats than more rural states (e.g. Minnesota and Wisconsin) after

Table 2. Regression results.

Democrat % seats won minus % seats expected	Hisp. dum Model 1	Model 2	>6 CDs
Democratic gerrymander	9.13*	10.1**	16.6***
	(4.63)	(4.79)	(4.75)
Republican gerrymander	−11.2***	−4.08	−13.6**
	(2.89)	(3.81)	(4.86)
Percent African-American	–	−0.41	−0.29
		(0.26)	(0.24)
Percent Hispanic	–	0.58**	0.77***
		(0.22)	(0.24)
Urbanization	–	0.046	−0.72**
		(0.22)	(0.34)
Democratic vote	–	0.32	0.11
		(0.21)	(0.24)
Number of seats	–	−0.29*	−0.16
		(0.16)	(0.18)
Hispanic dummy	9.95***	–	–
	(3.11)		
Constant	−5.52**	−25.5	45.0
	(2.26)	(15.8)	(29.2)
Observations	33	33	21
R-squared	0.557	0.641	0.829

Notes: Standard errors in parentheses. Data points weighted by state size. ***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

Table 3. Seats won versus mean vote share by gerrymandering party: 2012 presidential vote (summary).

Summary table

	CDs	Dem. Vote share	Dem. Seats won	Dem. seats Expected	Won–Exp. Difference
Republican gerrymanders	106	50%	28%	50%	−22%
Democratic gerrymanders	35	61%	80%	72%	8%
Bipartisan or court gerrymanders	97	57%	63%	64%	−1%
High Hispanic population	105	55%	57%	59%	−2%
Deep South states	38	43%	21%	36%	−15%

controlling for the gerrymandering party. Furthermore, the coefficients for partisan maps increase when we limit the sample to larger states, possibly indicating the greater flexibility parties have in drawing districts in such states.[10]

Robustness check: Presidential election results

Although the current congressional map has thus far only seen one cycle of election results, there has been another election held across all 435 of these districts that we can use to test the robustness of this paper's finding: the 2012 presidential election. Despite winning with 52.0% of the two-party popular vote, Obama won only 209 congressional districts, further suggesting pro-Republican bias. We can substitute Obama's margin for the congressional election result to measure bias under the various redistricting regimes.

The results of replicating Table 1 using presidential election results are summarized in Table 3 and are detailed in Table A3 of Supplementary Material. In the case of partisan maps and heavily Hispanic states, the average bias is very similar to the bias under the actual congressional election results. Notably, the difference in bias between Republican and Democratic gerrymanders remains the same at 14%. However, the pro-Republican bias under bipartisan and court gerrymanders largely disappears. There are likely two explanations for this difference. Firstly, President Obama won three districts in Minnesota and five districts in New York with 52% or less of the vote, which might be described as luck. However, this result also suggests that the asymmetry in the geographic distribution of partisans is not constant across states and regions. In some "bluish" states, the more conservative areas such as upstate New York and rural Minnesota may be only marginally Republican. These districts may be won by Republicans in a nationally tied electoral environment but captured by Democrats in a climate somewhat more favorable to them, such as Obama's 4% popular vote victory. In contrast, in the Deep South where the more conservative regions are deeper "red," probably exaggerated by VRA considerations, the bias against Democrats is actually exacerbated as their vote majority increases.

Conclusion

Both the state-by-state results and aggregated regression analysis suggest that while deliberate partisan gerrymandering produces additional seats for the districting party, partisan gerrymandering is not a sufficient explanation for the overall antimajoritarian outcome. Instead, pro-Republican bias is observed under all districting regimes. In addition, the regression results offer possible support for the Chen and Rodden (2013) thesis that urbanization has created bias while also forecasting its possible demise if patterns of rapid Hispanic population growth continue.

It is important to note the limits to these conclusions. Firstly, while asymmetric population distributions are a plausible explanation for persistent bias, and one supported by previous research, they are not the only possible cause. For example, one might claim that incumbency could give Republicans advantages in more marginal districts (see McGhee, 2012). This article does not attempt to isolate that cause.[11]

This analysis does not imply that Democrats are doomed to the minority even for the next decade. It does indicate they are unlikely to retake the House in an essentially tied national election. Yet national elections are not usually this close: Democrats reversed a Republican gerrymander in Pennsylvania, Virginia, Ohio, and Michigan in 2006 or 2008 (all states with aggressive Republican maps). The 2012 maps leave the Democratic Party several openings; for example, Republicans now sit in five Pennsylvania districts won by Obama in 2008. To win these seats, Democrats will need the electorate to look like 2006 or 2008, but this is far from unprecedented: Democrats won the popular vote by at least 5 points in 12 of the last 20 cycles. But given the unequal concentrations of vote share in most states, not just those with Republican gerrymanders, a Democratic majority will be a bit more difficult than it should be.

Supplementary Material

The entire Supplementary Material is available at: http://bit.ly/1jOtnma

Declaration of conflicting interest

The author declares that there is no conflict of interest.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Notes

1. While Hirsch argues that the combination of redistricting control and geographic imbalance biased the 2002 election results against the Democrats by 25 seats, he does not distinguish between these two factors in that estimate, and he argues that almost all bias can be located within four states with Republican-controlled districting.
2. The linear method estimates an average slope of 2.02 for the past 40 years, compared with Tufte's (1973) average of 2.09 for the preceding 70 years. Tufte justifies using a linear estimate, as opposed to probit or logit, because the major-party vote shares rarely fall outside of the 35–65% range. However, as vote shares in several states in the 2012 election fall outside of this range, a curve that will deal more appropriately with extreme values is needed for our purpose.
3. The constant in this regression represents approximately a 3–4% bias in favor of Democrats over this period. When broken down by decade (shown in Table A1 of Supplementary Material), the bias estimate aligns with past research in showing Democratic bias in the 1970s and 1980s, shifting toward Republican bias in the 2000s (e.g. King and Gelman, 1991; McGhee, 2012), possibly due to the same gerrymandering and geography trends observed here for 2012. The sign of this bias is reversed under the state elections data method (also in Table A1 of Supplementary Material), with the difference likely attributable to the method of imputation for unopposed races. If estimated using a logit function on the national data, the slope coefficient is .0415, with all results substantive unchanged.
4. In cases where a candidate runs unopposed and no votes are collected, no votes are added to the national total, but a 100% vote share is imputed into the state result. This will lead to a difference in responsiveness between the methods where unopposed incumbents are predominantly one party. About 70% of such races in the data set occur in the South, with about two-thirds of those being Democrats in the 1970s and 1980s.
5. Obviously, Democrats could not have hoped to perform better in Massachusetts than they did. At the state level, however, this example illustrates the national phenomenon of Democrats failing to maximize their vote by oversaturating their support in certain areas. In addition, Democrats controlled the process in Arkansas, but won none of its four seats; earning an average of 35% of the vote across this state would have predicted winning one seat under a "fair" map.
6. This average disparity is extremely close to the 6% disparity observed nationwide, as the Democrats' 1% popular vote advantage is estimated to correspond to 52% of seats expected, compared to 46% of seats actually won.
7. Reducing the Republican bias by 7% in the 238 seats under Republican, Democratic, or Bipartisan control in Table 1 nets

the Democrats (238 × .07) = 16.7 seats. If we instead assume the level of bias shown in Table 1, but allocate 70 seats to both Democratic and Republican control (rather than 35 and 106, respectively), this reduces the number of Republican seats in Republican-controlled maps by (36 × .19) = 6.8 seats, and increases the number of Democrats in Democratic-controlled maps by (35 × .05) = 1.75 seats (for a total of 8.6 seats).

8. The exception here is Georgia, which is biased toward the Democrats despite being districted in 2011 by Republicans. This is likely attributable to the novel strategy of "minority influence" districts employed in a Democratic gerrymander in the 2000s, a strategy upheld in *Georgia v. Ashcroft* (2003), combined with the need to avoid retrogression from this map to achieve VRA clearance in the next decade.
9. Because of the inclusion of other controls with continuous values in Models 2 and 3, the value of the constant is no longer inherently meaningful.
10. The coefficients for Republican gerrymanders between models are different at $p < .05$, but not significantly different for Democratic gerrymanders.
11. This explanation seems less plausible given that Mitt Romney won 52% of congressional districts despite losing the national popular vote by 4%.

References

Altman M and McDonald MP (2013) A half-century of Virginia redistricting battles: Shifting from rural malapportionment to voting rights to public participation. *University of Richmond Law Review* 47: 771–831.

Brunell TL (1999) Partisan bias in U.S. congressional elections, 1952–1996. *American Politics Research* 27(3): 316–337.

Chen J and Rodden J (2013) Unintentional gerrymandering: Political geography and electoral bias in legislatures. *Quarterly Journal of Political Science* 8(3): 239–269.

Erikson R (1972) Malapportionment, gerrymandering, and party fortunes in congressional elections. *American Political Science Review* 66(4): 1234–1245.

Goedert N (2012) Southern redistricting under the VRA: A model of partisan tides. In: *the 2012 state politics & policy conference*, Houston, TX. Available at: http:// 2012sppconference.blogs.rice.edu/files/2012/02/VRA-Partisan-Tides-SPPC-version1.pdf.

Hill KA (1995) Does the creation of majority black districts aid Republicans? An analysis of the 1992 elections in eight southern states. *Journal of Politics* 57(2): 384–401.

Hirsch S (2003) The United States House of Unrepresentatives: What went wrong in the latest round of congressional redistricting. *Election Law Journal* 2(2): 179–216.

King G and Browning RX (1987) Democratic representation and partisan bias in congressional elections. *American Political Science Review* 81(4): 1251–1273.

King G and Gelman A (1991) Systemic consequences of incumbency advantage in U.S. House Elections. *American Journal of Political Science* 35(1): 110–138.

Lijphart A (1999) *Patterns of Democracy: Government Forms and Performance in Thirty-Six Countries*. New Haven, CT: Yale University Press.

Lublin D (1999) *The Paradox of Representation*. Princeton, NJ: Princeton University Press.

McDonald M (2009) Seats to votes ratios in the United States. Unpublished Paper, George Mason University. Available at: http://www.democracy.uci.edu/files/democracy/docs/conferences/mcdonaldgmu.doc (accessed 31 March 2014).

McGhee E (2012) The history of the House votes-seats discrepancy, in two graphs. *The Monkey Cage* (Political Science blog). Available at: http://themonkeycage.org/blog/2012/11/19/the-history-of-the-house-votes-seats-discrepancy-in-two-graphs/ (accessed 31 March 2014).

Rae D (1967) *The Political Consequences of Electoral Laws*. New Haven, CT: Yale University Press.

Taagepera R and Shugart MS (1989) *Seats & Voters: The Effects & Determinants of Electoral Systems*. New Haven, CT: Yale University.

Tufte ER (1973) The relationship between seats and votes in two-party systems. *The American Political Science Review* 67(2): 540–554.

The Case of the Disappearing Bias:
A 2014 Update to the "Gerrymandering or Geography" Debate

Nicholas Goedert
Lafayette College
goedertn@lafayette.edu
November, 2015

ABSTRACT: This note observes that the pro-Republican bias in the relationship between seats and votes that characterized the 2012 U.S. congressional elections largely disappeared in the 2014 elections, where Republicans won a six-point victory in the national popular vote but only a handful of additional seats. Replicating analysis from an earlier article on the 2012 elections, I find that the source of the decline in bias supports two theories about the effects of gerrymandering and geography on the U.S. Congress. First, bias declined most sharply in states where maps were drawn by Republicans, suggesting these maps were drawn specifically to maximize seats during a tied national election environment. And second, pro-Republican bias present in bipartisan maps almost entirely disappears, as does the previously observed effect of urbanization on bias, further supporting existing theories about the asymmetric geographic dispersion of partisans.

The 2014 midterm elections were by most measures an unmitigated success for the Republican party. In addition to holding 55 Senate seats and 31 Governorships, Republicans won 247 seats in the House of Representatives, the party's largest majority since the Great Depression. But these 247 seats represent a surprisingly small gain considering the difference in the national popular vote for Congress between 2012 and 2014. Two years earlier, Republicans won a 33-seat majority despite losing the popular vote by 1%; in 2014, winning the popular vote by almost 6% yielded only an additional 13 seats.

And projections from scholars suggest that the modest Republican House gains may have indeed been surprising to given the overall size of the Republican wave on other fronts. The October 2014 issue of *PS: Political Science and Politics* included five short articles predicting the results of the upcoming elections. On the whole, these predictions were quite accurate in estimating a median Republican gain of 14 seats in the House (Campbell 2014). But while correctly or slightly over-predicting the Republican gains in House, all three articles addressing Senate races predicted the Republican would pick up fewer than the nine Senate seats they did (see Abramowitz 2014; Highton, McGhee and Sides 2014; Lewis-Beck and Tan 2014). Additionally, Abramowitz estimates that a six point Republican lead in the Congressional general ballot should result in a 17 seat gain in the House but a 7 seat gain in the Senate.

As discussed in my previous article "Gerrymandering or Geography?: How Democrats Won the Popular Vote but Lost the Congress in 2012" (2014), the 2012 congressional election result was strongly biased in favor of the Republicans due to a combination of the asymmetric geographic dispersion of partisan and intentional gerrymandering that the Republican party dominated following the 2010 census. But it seems shortsighted to only judge the overall bias of a map with respect to a single, closely contested election. Indeed, recent scholarship such as Stephanopoulos and McGhee (2015) has expanded on the notion that bias should be judged with respect to 50/50 election by measuring vote efficiency in maps across a range of election environments (see also McGhee 2014). This note replicates my 2012 analysis using the recent election data, and finds that these same factors play a much less certain role in inducing bias during in the Republican popular vote wave of 2014, despite the same maps being in effect.

We observe declining bias in both Republican and bipartisan gerrymanders. This result highlights two aspects of the debate over districting bias in the current cycle of congressional districting. First, bias is the product of the interaction of districts with the national election environment, and not stable across all elections. Maps that appear biased when the election is close may also appear fair when one party wins by a sizeable margin (and vice-versa). And second, the absence of bias in 2014, just like the presence of bias in 2012, is explainable by a combination of intentional gerrymandering and the asymmetric distribution of partisans.

National Seats-Votes Curve

Goedert (2014) observed that an historically average seat/votes curve over the past 40 years of U.S. congressional elections can be approximated by a line with a slope of about 2, or a probit curve with a slope of 0.026 (where the IV is the Republican advantage in the national popular vote, and the DV is Republican share of seats won). This largely matches the findings over the previous century by Tufte (1973). Figure 1 replicates the same table in Goedert 2014 with the addition of a data point for 2014. While 2012 lies far below both the linear (dashed) and probit (solid) expectation lines, indicating strong Republican bias in the result, 2014 falls much closer to expectation, despite the historically strong Republican seats total. Based on the historical average from 1972-2010, Republicans won 22 more seats than expected in 2012, but only 5 more than expected in 2014.

[Figure 1 about here]

Given the steep decline in Republican bias on the national level, we should also expect to see this bias disappear in many states whose delegations tilted toward Republicans in 2012.

Where should we expect to see bias decline most dramatically? It would be in states where (1) the partisan allocation of seats was biased toward Republicans in 2012; (2) the vote share for Republican increased in 2014; and (3) this increase led to few or no additional seats for the GOP in 2014. In moving from an evenly matched election to a moderate Republican wave, we would expect marginally-Democratic seats to be most likely to flip to Republicans; states with many such seats would see Republican bias increase in 2014, while states with none of these seats would see bias decrease. In other words, we are most likely looking at states that included very few swing or slightly left-leaning districts. Such a pattern would certainly be predicted in the case of Republican gerrymanders, and thus we predict the greatest decline in bias is states with Republican maps. However, the "asymmetric dispersion" theory would also predict this pattern of few lean-leaning swing seats in situations where the geographic dispersion of partisans (most states excepting those with high Hispanic populations) would tend to preclude their creation. So states with bipartisan gerrymanders should also see some decline in the bias generated from asymmetric partisan dispersion, but less than Republican gerrymanders, which *deliberately* avoid these districts.

In contrast, we would *not* expect to see bias decline in states containing a lot of slightly Democratic seats that would be vulnerable during a wave like 2014. This would include states with marginally Democratic regions (e.g. rural Hispanics-majority districts) or gerrymanders that would deliberately create them (drawn by Democrats). While Republican bias should not *decrease* in these states, it is unclear whether it should *increase;* this would depend on the partisan balance of the state compared to the size of the wave. The reason for this ambiguity is that the few Democratic gerrymanders in the current decade tended to occur in states that already consistently vote heavily Democratic, including Massachusetts and Maryland. It is possible that the Democratic vote is strong enough in these states that even a maximally Democratic gerrymander would not require drawing many marginally Democratic seats, or that the size of even the 2014 wave would not be enough to overcome their existing partisan lean.

Breakdown by Gerrymandering Regime

Table 1 replicates the same table from Goedert 2014, breaking down individual states by the party responsible for gerrymandering at the start of the decade, with separate categories for states with very high Hispanic population and deep South states most affected by Voting Rights Act constraints (as discusses in that article).[1]

[Table 1 about here]

As shown in Table 1, it appears that bias has responded exactly as hypothesized. We immediately see the biggest difference in the Republican gerrymanders, where Democratic vote share fell most steeply (from an average of 48% to 43%), but Republicans collectively gained only one seat. The result is that the pro-GOP bias generated from these maps was reduced by more than half. And the change was quite consistent across states: bias fell by at least 5% in eight of the nine states. In 2012, six of these states saw a Republican bias of at least 20%; in 2014, none of them do. It is still notable that Republican gerrymanders remained biased as a whole, as Republicans of course still win virtually all of the seats absent those few deliberately packed with Democrats. The decline in bias is largely due to Republicans winning seats the had already won in 2012, but by larger margins. It may be that bias in swing states Republican gerrymanders could be entirely reversed toward the Democrats under a strong Democratic tide (as was seen in states such as Pennsylvania and Ohio during the 2008 wave election), but this drastic outcome is unlikely during a Republican wave unless the tide was so strong as to make even packed Democratic seats competitive.

Bipartisan maps also see bias decline, though to a lesser extent and less predictably than Republican maps. Overall, these maps went from having a 7% Republican bias to less than 2%, now appearing collectively very close to fair. Republicans gained 4% in vote share in these states, and three additional seats, all in New York; overall both parties won about half the vote and half the seats.

In contrast, we might expect Republicans to gain several seats in Democratic gerrymanders, which generally try to draw slightly pro-Democratic districts to maximize their seat share in close elections. And we see evidence of this in Illinois, the most notable Democratic gerrymander of this decade, where Republicans defeated two incumbents in 2014, destroying the bias that map generated in 2012. Maryland remains highly biased toward the Democrats, largely because the incumbent in MD-06 survived a shockingly close race by 1%. And the all-Democratic delegation in Massachusetts remained, but their dominant mean vote share predicted Democrats would win every district in the state anyway. Overall, these states remained slightly biased toward Democrats as they had in 2014.[2] The summarized results in Table 2 suggest that both the intentional gerrymandering and geographic dispersion sources of bias declined by 5 percentage points between 2012 and 2014, from 12% to 7% in the case of gerrymandering, and from 7% to 2% in the case of geography.[3]

The previous article hypothesized that states with the largest Hispanic populations may not have displayed the same Republican bias as other states because Democratic-leaning Hispanics (especially in more rural areas), may have made the drawing of Democratic leaning districts more natural in these states. Conversely, we might expect these same districts to be more vulnerable to a moderate Republican wave. And indeed, Republicans gained a seat in each Arizona, Nevada, and Texas in 2014.[4] However, overall bias actually moved slightly in favor of Democrats, largely because Democrats were extremely fortunate to win all seven races decided in California by less than 5%. Bias did not change substantially in the Deep South states because Republican vote share changed very little; we might speculate that vote choice in this region is relatively inelastic.

[Table 2 about here]

Regression Analysis of Urbanization

In the previous article, a regression analysis showed that Republican bias correlated with urbanization among medium and large states in the 2012 elections, as a test of Chen and Rodden's (2013) theory that urban population patterns generate Republican bias in legislative maps even under neutral districting procedures. Table 3 replicates that analysis for 2014, with starkly different results. Both the effect of urbanization increasing Republican bias and the effect of Hispanic population decreasing it are reduced to statistically insignificant levels in 2014. The urbanization coefficient declines in 2014 because the forces that created bias in an evenly balanced election (many urban seats won overwhelmingly by Democrats, and less urban seats won narrowly by Republicans) are not as present in an election favoring Republicans. In 2014, those urban seats are still won by Democrats, but less overwhelmingly, while the Republican seats stay Republican by a larger margin. And when urbanization is no longer significantly associated with bias, the lack of bias among heavily-Hispanic states is no longer exceptional, as it was in 2012.

And the effects of partisan gerrymandering also becomes less significant. Although the coefficients on Democratic and Republican gerrymanders decrease only slightly, the uncertainty around them increases: partisan gerrymandering was a less consistent predictor of bias during the Republican wave in 2014 compared to the close election in 2012, a result consistent with state-by-state examples in Table 1. Note that the *difference* in these coefficients is not significant between 2012 and 2014. However, this is consistent with the general sense that while there is strong evidence of Republican bias due to both gerrymandering and geography, the conclusions we can draw in either direction on either count are much murkier in the case of 2014.

[Table 3 about here]

Conclusion

After a startling deviation from historical norms in 2012, the relationship of seats to votes in the 2014 congressional elections returned to a state much closer to expectation. While this evidence remains purely anecdotal based on two consecutive elections, the contrast between them provides further insight as to when to expect to find bias in congressional maps. In particular, the steep decline in bias in Republican-drawn maps suggests they were drawn specifically to maximize seat expectation in a nationally tied election. Additionally, the similar decline in bias in bipartisan maps in a pro-Republican wave election supports the theory that districts are sometimes unintentionally drawn resembling Republican gerrymanders, including many slightly right-leaning seats along with several heavily Democratic seats, due to the geographic dispersion of partisans. This is further supported by the contrasting effect (or lack there of) of urbanization on the bias across these elections.

Finally, the stark differences in results across temporal proximate and superficially similar elections highlights the importance of considering the national election environment, and its potential for wide variation, in evaluating gerrymanders and voting systems. When evaluating the respective effects of intentional gerrymandering and geographic dispersion, it is important to consider the range of possible electoral environments. Partisan gerrymanders may be drawn to be most effective (and this most biased) when then national electoral environment is close. But this same circumstance of a tied national election may also yield significant Republican bias due to geographic dispersion, making Democratic gerrymanders seem less effective, and Republican maps more effective, than they would under a different overall environment. So simply evaluating the context of a close national election may not tell the full story.

Moreover, many pundits have predicted a sustained and unbreakable lock on the House of Representatives through the remainder of the decade as a result of the bias observed in 2012.

But the Republican wave in 2014 demostrates that observation is not constant across time, and just as they did in 2008, Democrats could potentially eliminate this bias, both due to gerrymandering and geography, through a wave in their favor in 2016 or beyond.

Tables and Figures

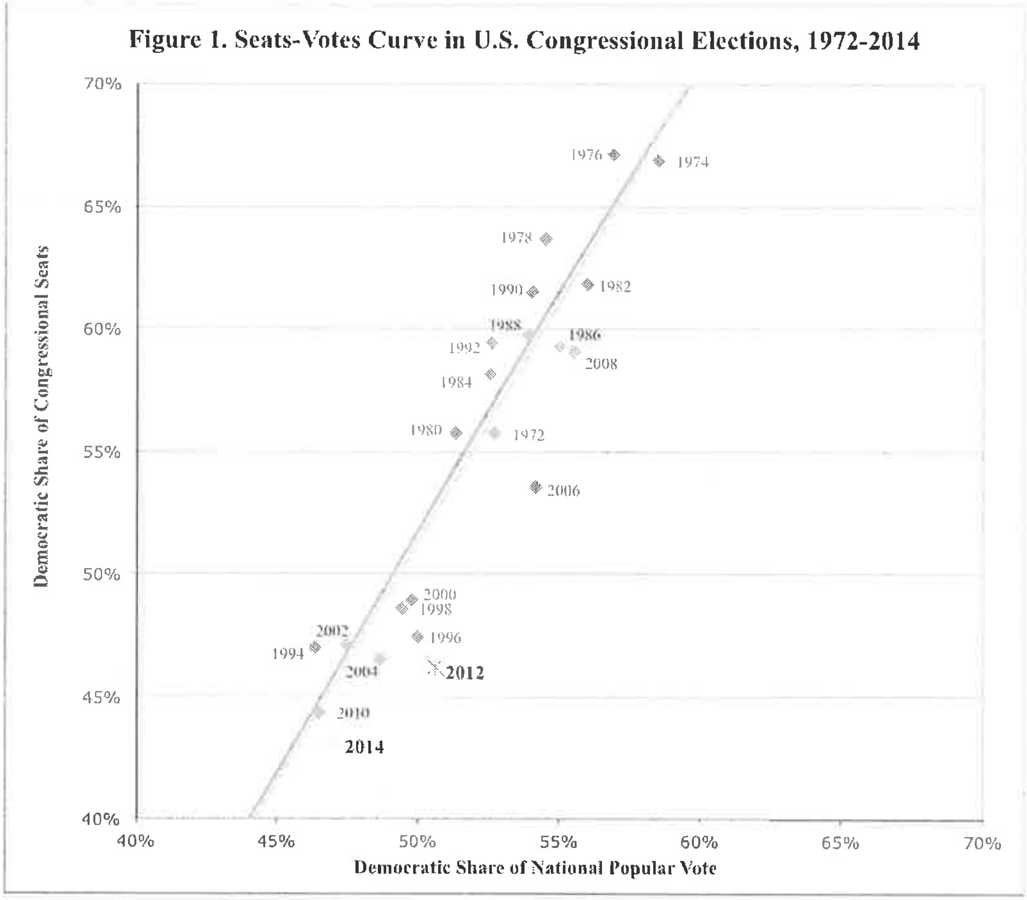

Figure 1. Seats-Votes Curve in Congressional Elections 1972-2014

Table 1. Seats Won vs. Mean Vote Share By Gerrymandering Party: 2014 Congressional Elections

Republican Gerrymanders

State	CDs	Dem. Vote Share	Dem. Seats Won	Dem. Seats Expected	Won-Exp. Difference
Indiana	9	40%	22%	29%	-7%
Michigan	14	52%	36%	54%	-18%
Missouri	8	39%	25%	28%	-3%
North Carolina	13	44%	23%	37%	-14%
Ohio	16	41%	25%	31%	-6%
Pennsylvania	18	45%	28%	39%	-11%
Tennessee	9	35%	22%	22%	0%
Virginia	11	42%	27%	33%	-6%
Wisconsin	8	48%	38%	45%	-8%
Weighted Average	**106**	**43%**	**27%**	**36%**	**-9%**
2012 Average	***106***	***48%***	***28%***	***47%***	***-19%***

Democratic Gerrymanders

State	CDs	Dem. Vote Share	Dem. Seats Won	Dem. Seats Expected	Won-Exp. Difference
Illinois	18	53%	56%	57%	-1%
Massachusetts	9	86%	100%	97%	3%
Maryland	8	59%	88%	68%	20%
Weighted Average	**35**	**63%**	**74%**	**70%**	**5%**
2012 Average	***35***	***63%***	***80%***	***75%***	***5%***

Bipartisan or Court Gerrymanders

State	CDs	Dem. Vote Share	Dem. Seats Won	Dem. Seats Expected	Won-Exp. Difference
Colorado	7	48%	43%	46%	-3%
Florida	27	43%	37%	35%	2%
Kentucky	6	36%	17%	23%	-7%
Minnesota	8	52%	63%	55%	8%
New Jersey	12	55%	50%	59%	-9%
New York	27	63%	67%	75%	-8%
Washington	10	50%	60%	50%	10%
Weighted Average	**97**	**51%**	**51%**	**53%**	**-2%**
2012 Average	***97***	***55%***	***54%***	***61%***	***-7%***

High Hispanic Population States

State	CDs	Dem. Vote Share	Dem. Seats Won	Dem. Seats Expected	Won-Exp. Difference
Arizona	9	45%	44%	40%	5%

California	53	58%	74%	66%	7%
New Mexico	3	52%	67%	54%	13%
Nevada	4	44%	25%	38%	-13%
Texas	36	39%	31%	29%	2%
Weighted Average	**105**	**50%**	**54%**	**50%**	**5%**
2012 Average	***105***	***53%***	***56%***	***56%***	***0%***

Deep South States

State	CDs	Dem. Vote Share	Dem. Seats Won	Dem. Seats Expected	Won-Exp. Difference
Alabama	7	35%	14%	22%	-8%
Georgia	14	40%	29%	31%	-2%
Louisiana	6	28%	17%	13%	4%
Mississippi	4	38%	25%	27%	-2%
South Carolina	7	31%	14%	17%	-2%
Weighted Average	**38**	**36%**	**21%**	**23%**	**-2%**
2012 Average	***38***	***37%***	***24%***	***26%***	***-2%***

Table 2. Summary of Bias in 2012 vs. 2014

Districting	Seats	2012 Bias	2014 Bias
Republican	106	GOP +19%	GOP +9%
Non/Bipartisan	97	GOP +7%	GOP +2%
Democratic	35	Dem +5%	Dem +5%

Table 3. Regression Results

Democrat % Seats Won Minus % Seats Expected		**>6 CDs 2012**		**>6 CDs 2014**		
Democratic Gerrymander		16.6***		11.3*		
		(4.75)		(5.86)		
Republican Gerrymander		-13.6**		-12.6*		
		(4.86)		(6.31)		
Percent Black	13.2	-0.29	6.6	-0.32		
		(0.24)		(0.31)		
Percent Hispanic	17.4	0.77***	6.5	0.26		
		(0.24)		(0.28)		
Urbanization	80.7	-0.72**	70.2	-0.35		
		(0.34)		(0.43)		
Democratic Vote	51.0	0.11	50.8	-0.33	47.2	47%
		(0.24)		(0.24)		
Number of Seats	9	-0.16	8	0.12	8	9
		(0.18)		(0.21)		
Constant		45.0		44.0		
		(29.2)		(35.6)		
Observations		21		21		
R-squared		0.829		0.570		

Notes: Standard errors in parentheses. Data points weighted by state size. *** $p<0.01$, ** $p<0.05$, * $p<0.10$

1.85% 4.39%

References

Abramowitz, Alan I. 2014. "Forecasting the 2014 Midterm Elections with the Generic Ballot Model". *PS: Political Science & Politics* 47(2): 772-74.

Campbell, James E. 2014. "The 2014 Midterm Election Forecasts". *PS: Political Science & Politics* 47(2): 769-71.

Chen, Jowei and Jonathan Rodden. 2013. "Unintentional Gerrymandering: Political Geography and Electoral Bias in Legislatures". *Quarterly Journal of Political Science* 8: 239-69.

Goedert, Nicholas. 2014. "Gerrymandering or Geography?: How Democrats Won the Popular Vote but Lost the Congress in 2012". *Research & Politics* 1(1): 2053168014528683.

Highton, Benjamin, Eric McGhee and John Sides. "Election Fundamentals and Polls Favoring the Republicans". *PS: Political Science & Politics* 47(2): 786-88.

Lewis-Beck, Michael S. and Charles Tien. 2014. "Congressional Election Forecasting: Structure-X Models for 2014". *PS: Political Science & Politics* 47(2): 782-85.

McGhee, Eric. 2014. "Measuring Partisan Bias in Single-Member District Electoral Systems". *Legislative Studies Quarterly* 39(1): 55-85.

Stephanopoulos, Nicholas O. and Eric M McGhee. 2015. "Partisan Gerrymandering and the Efficiency Gap." Forthcoming in University of Chicago Law Review 82.

Tufte, Edward R. 1973. "The Relationship between Seats and Votes in Two-Party Systems." *The American Political Science Review* 67(2): 540-54.

---

[1] In this table, Democratic Vote share is the mean popular vote share across the state by Congressional district, Democratic Seats Expected is the number of seats we estimate Democrats should have won in a fair map given their vote share according to historical average, using a probit curve with a slope of 0.026 and an intercept of 0.

[2] The average expected seats in these states declines despite the very little change in mean vote share because vote share increased in MA, where further increase has little effect on expected seats because they were already expected to win almost every seat, but decreased in IL, where expected seats was much more sensitive to the change. Note that Democrats also lost all seven seats in Arkansas and West Virginia, two smaller states where they controlled the gerrymander.

[3] This breakdown is calculated by assuming the average bias observed in the bipartisan states (7%/2% in 2012/2014) is the overall bias due to geography, and then subtracting this from the total bias in the partisan states to yield the portion of bias in partisan maps due to deliberately gerrymandering. (E.g. the total Republican bias in 2014 GOP maps is 9%, so this is 7% due to gerrymandering if it is 2% due to geography.) In both the case of 2012 and 2014, this turns out to be the same absolutely bias for Democrats and Republicans.

[4] All three were swing districts at the national level; the Texas seat was Hispanic majority, while the Nevada and Arizona seats had approximate Hispanic populations of 30% and 20% respectively.

The measure of partisanship should exist to establish the change in the partisan balance of the district. We are not in court this time; we do not need to show that we have created a fair, balanced, or even a reactive map. But, we do need to show to lawmakers the political potential of the district.

I have gone through the electoral data for state office and built a partisan score for the assembly districts. It is based on a regression analysis of the Assembly vote from 2006, 2008, and 2010, and it is based on prior election indicators of future election performance.

I am also building a series of visual aides to demonstrate the partisan structure of Wisconsin politics. The graphs will communicate the top-to-bottom party basis of the state politics. It is evident, from the recent Supreme Court race and also the Milwaukee County executive contest, that the partisanship of Wisconsin is invading the ostensibly non-partisan races on the ballot this year.

EXHIBIT 25
WIT: Goedert
DATE: 12-15-15
Halma-Jilek Reporting, Inc.