IN THE UNITED STATES DISTRICT COURT
FOR THE SOUTHERN DISTRICT OF OHIO
EASTERN DIVISION

TOMMY RAY MAYS II and QUINTON NELSON SR., individually and on behalf of all others similarly situated, Plaintiffs, v. FRANK LaROSE, in his official capacity as Secretary of State, Defendant.	Case No. 2:18-cv-1376 JUDGE MICHAEL H. WATSON Magistrate Judge Chelsey M. Vascura **CLASS ACTION**

PLAINTIFFS' RESPONSE TO DEFENDANT'S MOTION, *IN LIMINE*, TO EXCLUDE THE EXPERT TESTIMONY OF DR. MARK SALLING, PH.D

I. Introduction

Dr. Mark Salling is a respected and experienced data scientist who has worked with Ohio demographic and population data for his entire career. Plaintiffs retained Dr. Salling in this matter to offer opinions on two subjects well within his expertise: (1) an estimate on the number of Ohio voters who are prevented from voting because they were incarcerated after the absentee ballot application deadline and held through Election Day (the "Incarcerated Voters Report," Doc. 30-1) and (2) a mapping of the geographic distances between county Boards of Elections and the (a) hospitals and (b) county jails located within the same county (the "Proximity Report," Doc. 55-2).

Rather than use the data produced by Plaintiffs and publicly available to test Dr. Salling's conclusions, the Secretary recites marginal objections to aspects of Dr. Salling's reports the Secretary thinks should have been analyzed differently. But the Secretary's objections—some at the level of minutiae—are not the proper topic of a motion to *exclude* Dr. Salling's testimony. The

Secretary fails to show that Dr. Salling's data is unreliable, methodology is unsound, or his conclusions inaccurate. His motion to exclude Dr. Salling's testimony should be denied.

II. Factual Background

Dr. Mark Salling is a Senior Research Associate and College Fellow at the Maxine Goodman Levin College of Urban Affairs at Cleveland State University, and the Director of the College's Northern Ohio Data and Information Service ("NODIS"). Doc 30-1 (Plaintiffs' Submission of Expert Testimony) at 15. He has held both positions for more than 30 years. *Id.* Dr. Salling also served for eleven years as the Director of Research for The Center for Community Solutions in Cleveland, Ohio. *Id.* Over his long career, Dr. Salling has also served as Ohio's Liaison to the U.S. Census Bureau's Redistricting Data Program, as the Chairman of the Cleveland Census Statistical Areas Committee, and on the Board of Directors for the Urban & Regional Information Systems Association ("URISA"). *Id.* Dr. Salling's areas of research and academic publications include demographic and urban analysis, analyses of redistricting outcomes using geographic information systems ("GIS"), urban neighborhood economic structure, the use of GIS for social indicators, demographic trends, and urban issues. *Id.* at 16-29. He is an expert in Demography, GIS, and Survey Design and Analysis. *Id.*

In the Incarcerated Voters Report, Dr. Salling "offers estimates of the number of jailed persons in Ohio who were registered to vote in the state, had not voted by absentee ballot, and could not exercise their voting rights because they were detained and confined after the absentee ballot request deadline and held through Election Day of the general elections in November 2012, 2014, 2016, and 2018"—the potential class in Plaintiffs' lawsuit. Doc. 30-1 at 3. Dr. Salling relied on quality, publicly available data to support his analysis. Plaintiffs provided Dr. Salling with a dataset synthesized from two sources: (1) booking data collected from the sheriffs of 13 Ohio

counties pursuant to public records requests issued in 2018; and (2) voter registration data, including the Ohio voter file, compiled and maintained by a reputable data vendor. *Id.* at 3-4. The former was provided to the Secretary and the latter is under the Secretary's control. Ex. A, Diaz Decl. at ¶¶ 4, 7; Ex. B, Brill Decl. at ¶¶ 7-8. The dataset provided to Dr. Salling was created with the assistance of a well-respected organization specializing in data and technology services. Ex. C, Burchard Decl. at ¶¶ 1, 5-7.[1] That firm used a sophisticated algorithm to match the booking data to the Ohio voter file with a high level of accuracy and create a dataset. *Id.* at ¶¶ 6-7. In addition to this dataset (hereinafter "matched dataset"), Dr. Salling used U.S. Census Data on the adult population of each Ohio county, Doc. 30-1 at 3 n.2, and data from the Ohio Department of Rehabilitation and Correction showing the number of prisoners in state institutions, by county. *Id.* at 8 and n.5.

Based on this data, Dr. Salling performed a conservative analysis that concludes that during each federal general election from 2012-2018, somewhere between 761 and 1371 registered voters, and likely more, were unable to vote due to their arrest after the absentee ballot application deadline (6 P.M. on the Friday before an election), and detention through Election Day. *See id.* at 4. Dr. Salling used the matched dataset to identify the number of persons in the target population

[1] To be more specific, lawyers for Plaintiffs have a relationship with The Movement Cooperative ("TMC"), a membership organization that provides access to data and technology resources to its members, which are primarily nonprofit organizations. Ex. C, Burchard Decl. at ¶ 3; Ex. A, Diaz Decl. at ¶ 3. TMC contracts with TargetSmart for access to its "national voterbase file," a database of person-level records that includes the publicly available voter file from the Ohio Secretary of State. *See* Ex. C, Burchard Decl. at ¶ 4; Ex. B, Brill Decl. at ¶¶ 7-8. TMC used a licensed automated matching program to match the booking data gathered by Plaintiffs through public records requests against TargetSmart's national voterbase file and returned a matched dataset with numerical values ("match scores") that indicated the strength of the matches between each entry in the two data sets. Ex. C, Burchard Decl. at ¶¶ 5-7, Ex. A. Diaz Decl. at ¶ 4. After performing the matching process, TMC returned the matched dataset to Plaintiffs' attorneys for Dr. Salling to rely upon. Burchard Decl. at ¶ 8.

for the counties from which Plaintiffs had booking data. But any matching between datasets requires the analyst to determine the likelihood that the match accurately identifies the same person in two different datasets. To perform the most comprehensive analysis possible, Dr. Salling looked at the number of impacted individuals at a variety of thresholds of probability provided by TMC. In 2018, for example, the number of impacted individuals in the 13 counties based on a statistically reliable benchmark was calculated to be 639 persons. *Id.* at 7, Figure 2. Using the most stringent threshold, the number of impacted individuals in the 13 counties was calculated to be 342. *Id.*

Dr. Salling then used two methods to extrapolate from the 13 counties data and estimate the number of impacted individuals statewide. He extrapolated based on (1) the proportion of the adult population in the sampled counties versus whole state; and (2) the proportion of the institutional population in the sampled counties versus whole state. *Id.* at 4-8. Because Dr. Salling used more than one method to provide estimates of the impacted voter population, his report provides a range of the likely number of impacted individuals, not a single estimate. *See id.* at 6-10. Dr. Salling's estimates, however, all pointed in a similar direction, providing additional support for his report. *Id.*; *see also* Doc. 50-1 (Salling Dep.) at 165:8-25. Plaintiffs offered the Incarcerated Voters Report primarily to support the numerosity element of class certification—an element the Secretary has conceded. Doc. 35 at 11.

In the Proximity Report, Dr. Salling "maps the geographic distributions of county Boards of Elections (BOEs), hospitals, and jails in the State of Ohio and calculates and compares the distances from each county's jails and hospitals to the county's BOE." Doc. 55-2 (Proximity Report) at 1. The Proximity Report "address[es] the issue of whether travel by BOE staff to jails in order to deliver and retrieve election ballots of voting-eligible inmates is more or less burdensome than travel to hospitals in the county to deliver and retrieve such ballots from voting-

eligible hospital in-patients." *Id.* Dr. Salling was provided the data for this report by counsel. *Id.* at 1 n.1. There are three sets of data supporting this report, coming from three publicly-available sources: (1) a list of addresses of county jail facilities in Ohio, provided by the Ohio Department of Rehabilitation and Correction, Doc. 50-1 at 319; (2) a list of the addresses of Ohio Hospitals Association member hospitals registered with the Ohio Department of Health, as listed online at https://www.ohiohospitals.org/About-OHA/Ohio-Hospitals/Member-Hospitals, *id.* at 325; and (3) a list of the addresses of the county BOEs in Ohio as listed on the Secretary of State's website, *id.* at 323. Dr. Salling concludes that "the average distance [BOEs must travel] to this set of jails is not statistically different than to hospitals at the 90 percent confidence level." Doc. 55-2 at 1. Dr. Salling analyzed the straight-line distance between these locations using GIS software, which is used to calculate distances between locations. Doc. 50-1 at 24:11-15. Dr. Salling would expect no meaningful difference if the analysis was done with driving distances instead, based on previous research he conducted that found no meaningful difference between the two types of analyses. *Id.* at 155:1-23. Plaintiffs offer the Proximity Report to support the ease of an available procedure that the Secretary could use to allow plaintiff's class to vote on Election Day.

Dr. Salling also served as an expert for plaintiffs in the *Fair Elections Ohio v. Husted* litigation. *See, e.g.*, Doc 2-3. In that case, he offered reports on the two subjects for which he offers opinions in the current litigation. *See id*. He used similar, but not precisely the same, methodology in the earlier litigation, and has updated his sources of data. Doc. 50-1 at 58:5-60:16.

III. Argument

A. Standard of Review

Federal Rule of Evidence 702 governs the admissibility of expert testimony. *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993) and *Kumho Tire Co. v. Carmichael*, 526 U.S.

137 (1998), counsel that district courts must act as "gatekeepers" by assessing the reliability of the expert's principles and methodologies, although "[t]he 'gatekeeper' doctrine . . . is largely irrelevant in the context of a bench trial." *Deal v. Hamilton Cty. Bd. of Educ.*, 392 F.3d 840, 851-52 (6th Cir. 2004). "Rejection of expert testimony 'is the exception rather than the rule.'" *In re Scrap Metal Antitrust Litigation*, 527 F.3d 517, 531 (6th Cir. 2008) (quoting Fed. R. Evid. 702 Advisory Committee's Notes, 2000 Amend.). The court's gatekeeping function is not "intended to serve as a replacement for the adversary system." *Rose v. Matrixx Initiatives, Inc.*, No. 07–2404–JPM/tmp, 2009 WL 902311, at *7 (W.D. Tenn. March 31, 2009) (quoting Fed. R. Evid. 702 Advisory Committee's Notes); *see also Guglielmo v. Montgomery Cty., Ohio*, No. 3:17-cv-6, 2019 WL 2106103 (S.D. Ohio May 14, 2019). Further, *Daubert* does not require a district court to look at whether a proposed expert could have performed her analysis better. *U.S. v. Phung*, 127 F. App'x 594, 598 (3d Cir. 2005) (*citing In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717, 744 (3d Cir. 1994) (observing that "good grounds" for an expert's opinion may exist "even if the judge thinks that a scientist's methodology has some flaws such that if they had been corrected, the scientist would have reached a different result")). Rather, the "key question to be answered in determining whether a theory or technique is scientific knowledge [is] . . . whether it can be (and has been) tested." *Daubert*, 509 U.S. at 593 (citing K. Popper, Conjectures and Refutations: The Growth of Scientific Knowledge 37 (5th ed. 1989) ("The criterion of the scientific status of a theory is its falsifiability, or refutability, or testability")).

"[T]he district court has wide latitude in determining what factors to consider in establishing the reliability of an expert witness." *Bureau v. State Farm Fire & Cas. Co.*, 129 F. App'x 972, 976 (6th Cir. 2005). Thus, while the four *Daubert* factors set forth in the Secretary's Motion, Doc. 53 at 5-6, may be pertinent in certain cases, in cases involving nonscientific expert

testimony, the relevant reliability concerns may instead focus upon the expert's personal knowledge or experience. *Ohio Organizing Collaborative v. Husted*, No. 2:15-cv-1802, 2016 WL 8201848, at *2 (S.D. Ohio May 24, 2016) (Watson, J.). In fact, courts "take a liberal view of what 'knowledge, skill, experience, training, or education' is sufficient to satisfy the requirement." *Id.* at *6 (quoting *Bradley v. Ameristep, Inc.*, 800 F.3d 205, 208-98 (6th Cir. 2015)).

B. Dr. Salling is an Experienced, Non-Biased Data Scientist Offering Opinions Within His Realm of Expertise.

The Secretary attempts to paint Dr. Salling as a biased source who is a "quintessential" litigation-for-hire expert. *See, e.g.*, Doc. 53 at 17-18. That is far from the case. While Dr. Salling has served as an expert witness in several litigation matters, his primary employment has been and remains as a professor teaching, conducting research, and publishing that research. Prior to this case, Dr. Salling last gave expert testimony in 2014. Doc. 30-1 at 30-31. Since that time, Dr. Salling has continued to publish articles, and given seven presentations at professional conferences. *Id.* at 16-17, 22. He has also continued to teach courses and supervise graduate students at Cleveland State University. *Id.* at 15. He has served as an expert in fourteen cases over the past twenty years. *Id.* at 30-31. Dr. Salling has been engaged by nonprofit organizations in many of these cases, but he has also testified on behalf of the government. *See, e.g.*, Doc. 50-1 at 31:11-21. Dr. Salling has never turned down a client based on their aims in the litigation. *Id.* at 41:21-24.

The Secretary argues that Dr. Salling's prior experience serving as an expert on the matter *Fair Elections Ohio v. Husted*, No. 12-cv-0797 (S.D. Ohio), somehow decreases his credibility and objectivity in this case. *See* Doc. 53 at 1, 4. To the contrary, Dr. Salling's prior work on a similar topic only bolsters his credentials. Dr. Salling did prepare reports on two similar subjects, using similar methodologies, in *Fair Elections Ohio*. Indeed, the District Court judge in that case cited to evidence from Dr. Salling's report as part of granting Plaintiffs' Motion for Summary

Judgment on their equal protection and due process claims. *Fair Elections Ohio*, 47 F. Supp. 3d at 611, 617 (S.D. Ohio 2014) (reversed on standing grounds, 770 F.3d 456 (6th Cir. 2014)). The foregoing only makes Dr. Salling more qualified to testify in this case—his experience allowed him to reliably apply and refine the methodologies based on what he learned in *Fair Elections Ohio*. Doc. 50-1 at 58:5-60:16.

The Incarcerated Voter Report and Proximity Report are consistent with the type of research, analysis, and publications Dr. Salling has produced throughout his thirty years in academia, and with the type of analysis ordinarily relied upon by data experts.[2] Dr. Salling has worked with census and voter registration data throughout his career, in both academic and litigation contexts. Doc. 50-1 at 21:12-22:1. As director of NODIS, a position which he has held for more than 30 years, Dr. Salling's responsibilities include tasks such as demographic and urban analyses, urban neighborhood economic structure, and the use of GIS for social indicators and related areas. Doc. 30-1 at 16. Plaintiffs' engagement letter with Dr. Salling allows him to use the analyses conducted and conclusions reached in his expert reports as the subject of future academic papers and presentations. Doc. 50-1 at 266, Ex. 3, Sponsored Research Agreement Sec. 3.3.

Defendant makes much of Dr. Salling's "code of ethics," suggesting that it makes him "personally invested" in the case and implying that such a code biases his opinions offered in this litigation. Doc. 53 at 2, 4. But that is not the case. Dr. Salling testified that he belongs to an

[2] *See, e.g.*, Doc. 30-1 at 16-28 (listing among Dr. Salling's papers: Kaufman, Miron, Sanda Kaufman, and Mark Salling, "Dynamic firm location network model: empirical validation", forthcoming in Journal on Policy and Complex Systems; Salling, Mark, "Ohio's Use of Geographic Information Systems to Demonstrate Public Participation in the Redistricting Process," Duke Journal of Constitutional Law & Public Policy, Vol. 5, 2010, pp.112-123; Salling, M., and E. Cyran. "Estimates of the Number of Voters Whose Driver's License Address May Differ from Their Voting Address," Cleveland State University, Center for Election Integrity, Research Series, August 2, 2006).

association, URISA, that has a code of ethics that "says that there are certain responsibilities that such a professional has," including taking "opportunities . . . to work . . . on projects and issues which serve the larger public." Doc. 50-1 at 40:15-41:20. As the publicly-available code states, its primary goal is to "preserve and enhance public trust in the discipline." URISA Code of Ethics, available online at https://www.urisa.org/about-us/gis-code-of-ethics/. Among the "Obligations to Society" detailed in the code are to "Do the Best Work Possible" including by being "objective," "us[ing] due care, "practic[ing] integrity," and "provid[ing] . . . accurate information." *Id.* And, consistent with that code, there have been instances when Dr. Salling concluded that the data did not support the arguments made against an existing policy. *See* Doc. 50-1 at 42:16-43:6. It is surprising that the Secretary would target this code of ethics as somehow disqualifying.

C. Dr. Salling's Reports Are Based on Relevant, Appropriate Facts and Data.

As the Sixth Circuit has explained, "[w]here an expert's testimony amounts to 'mere guess or speculation,' the court should exclude his testimony, but where the opinion has a reasonable factual basis, it should not be excluded." *United States v. LE Cooke Co.*, 991 F.2d 336, 342 (6th Cir. 1993). Dr. Salling has adequately set forth in his reports and deposition testimony the reliable sources of the data integral to his analysis, and the Secretary's quibbles around the margins are insufficient to warrant the exclusion of his expert testimony.

i. Dr. Salling Is Reasonably Familiar with the Source and Content of the Data He Used.

The Secretary suggests that Dr. Salling does not know or understand the data underlying both of his analyses, and that his "mere cursory explanations of the data" mean "there is no way to independently test and verify the accuracy and reliability of the data." *See* Doc. 53 at 8-9 (citing *Buck v. Ford Motor Co.*, 810 F. Supp. 2d 815 (N.D. Ohio 2011)). The Secretary is not only misrepresenting Dr. Salling's knowledge but also using the wrong legal test. As stated above, the

test is that "where the opinion has a reasonable factual basis, it should not be excluded." *LE Cooke Co*., 991 F.2d at 342. Dr. Salling's report undeniably relies upon reasonable factual bases, including public data provided by state actors and matching analysis provided by a recognized expert in the field of matching voter registration data with other datasets.

The Secretary complains that Dr. Salling "does not fully remember which data fields he actually used" or "what the data refers to" for his Incarcerated Voters Report. Doc. 53 at 8 (citing Salling Dep. 111-112; Exs. 11, 12; Salling Dep. 120, 127; Exs. 12, 14). But the Secretary misrepresents the deposition testimony he relies upon. That testimony, in context, merely shows Dr. Salling using a printout of the data to refresh his recollection. When presented with an exhibit and asked which fields he relied on, Dr. Salling stated "I can't remember where they are. Let's see if I can find them." Following his review of the printout, Dr. Salling identified the specific columns of data he relied on. Doc. 50-1 at 119:25-120:21.[3]

In fact, the data Dr. Salling relied upon is not reasonably in doubt and all of it was provided to the Secretary to permit his independent verification, which he chose not to conduct. *See* Doc. 51-2 (Mockabee Dep.) at 25:5-26:21. With respect to the Incarcerated Voters Report, Dr. Salling relied upon (1) booking reports provided by county officials; (2) matching analysis between those reports and a national voter database, which was conducted through a reputable data matching service provider; and (3) other public demographic data. Doc. 30-1 at 3 and n.1&2, 4 and n.4, 8 and n.5; Ex. A, Diaz Decl. at ¶¶ 4-7; Ex. B, Brill Decl. at ¶¶ 7-8; Ex. C, Burchard Decl. at ¶ 5-7; *see also, e.g.*, *Gussack Realty Co. v. Xerox Corp.*, 224 F.3d 85, 94 (2d Cir. 2000) ("[A]n expert

[3] There are a couple of instances during his deposition where Dr. Salling was not able to explain the meaning of a particular column of data, but in each of those instances that is because it was not a type of data used in the analyses, or because he was asked about nuances in the data that did not affect his analysis. *See, e.g.*, Doc. 50-1 at 108:5-109:1; 127:22-130:2.

may rely upon data that she did not personally collect. The Federal Rules of Evidence specifically provide that an expert may rely on facts or data' perceived by or made known to the expert at or before the hearing.'"). Likewise, with respect to the Proximity Report, Dr. Salling reasonably relied upon the lists of jails and hospitals assembled from State of Ohio data sources and provided to him by counsel. Doc. 55-2 at 1 and n.1. Dr. Salling was aware of the sources of that data, which provided a reliable base of facts from which he could conduct his analysis. *See, e.g.*, *Nelson v. Walnut Inv. Part'rs*, No. 1:09-CV-00750, 2011 WL 2711318, at *4-5 (S.D. Ohio July 13, 2011) (rejecting *Daubert* challenge where expert "clearly made judgment calls [on which data source to use] that Defendant disagrees with, but that does not mean his methods were unreliable").

ii. The 13-County Sample of Booking Data Provided a Reasonable Factual Basis for a Statewide Estimate of Impacted Individuals.

The Secretary first quibbles with Dr. Salling's use of booking data for thirteen Ohio counties to extrapolate a statewide estimate of affected individuals. *See* Doc. 53 at 7-8. But the Secretary provides no evidence that this sample is biased or skewed. The *Daubert* standard does not require perfect data, nor could it. Plaintiffs are entitled to rely upon the data reasonably available to them. The Secretary had every opportunity to counter Dr. Salling's analysis with analysis of his own. He did not. As such, Dr. Salling's report represents the only evidence available to the Court of the number of affected individuals. That evidence is unrefuted, and therefore the Court should not only not exclude it; it can affirmatively rely upon it. *See, e.g.*, *Directors of Ohio Conference of Plasterers v. Meerkat Constr., Inc.*, No. 3:18-CV-00122, 2019 WL 3358598, at *3 (S.D. Ohio July 19, 2019) ("[T]his Court can properly rely on the unrefuted evidence before it.").

The Incarcerated Voters Report explains that Plaintiffs' counsel gathered this data to "include the 11 largest counties in Ohio by population," with two additional counties selected to "increase the geographic diversity of the sample to include medium-sized counties in west-central

and north-central Ohio." Doc. 30-1 at 3 n.3. The counties produced the booking data in a variety of different formats, and did not include uniform variables in the data sets. Ex. D, Pasternak Decl. at ¶¶ 4, 7. Plaintiffs' counsel consolidated the booking data produced by the counties into one spreadsheet and standardized the formatting of the information provided; the substance of the data was not altered in any way. Ex. A, Diaz Decl. at ¶ 5; Ex. D, Pasternak Decl. at ¶¶ 4-9.

Due to time and resource constraints, Dr. Salling did not participate in the selection of the counties in the sample; Plaintiffs' counsel worked to gather this data for months, including prior to the filing of this lawsuit. But when he was informed of the counties from which Plaintiffs had received data, Dr. Salling agreed that those counties' data was adequate and addressed his primary concerns for achieving a representative sample—population size, urban versus rural counties, and geographic diversity. Doc. 50-1 at 91:19-96:2; 157:7-18. Dr. Salling stated that he would have liked to have information from additional counties, because "a researcher always wants more data," but understood that the timeline of the case,[4] along with the cost and time of gathering actual data from all counties in Ohio, were "prohibitive," and agreed that a solid statistical analysis could be done with the sample that had been collected. Doc. 50-1 at 91:5-96:2.

Finally, any concerns—however unsupported—about the representative nature of the county sample do not impact the number of affected persons identified in the actual data sample. Even prior to calculating the statewide estimate, the data in the sample alone established between 409-492 people deprived of their right to vote, per general election, due to incarceration after the absentee ballot application deadline between 2012 and 2018. (Doc. 30-1 at 10, Figure 6).

[4] Expert reports in this case were completed and submitted to Defendant on April 12, 2019—approximately five months after the case was filed, *see, e.g.*, Doc. 30, as required by the preliminary pretrial order, Doc. 28 at 3.

iii. The Jail and Hospital Data Provided a Reasonable Factual Basis for Dr. Salling's Proximity Report.

As described above, Dr. Salling relied upon data publicly available to analyze the location of jail and hospital facilities. Doc. 50-1 at 319, 325. Yet, the Secretary contends that the jail and hospital data Dr. Salling used was unreliable because he would have included three regional jails Dr. Salling did not include. Doc. 53 at 8. He also questions Dr. Salling's inquiry into which medical facilities constitute "hospitals" but does not dispute the accuracy of the underlying data. *Id.*

The Secretary provides no reason to believe the data sources Dr. Salling relied upon are more generally inaccurate or unrepresentative of jail or hospital facilities in Ohio overall. Nor could he. *See supra*. If the Secretary believes that a different or more comprehensive set of jails or hospitals should have been used, he was free to provide that data to Plaintiffs or conduct that analysis with his own expert. He did not do so and provides no evidence that the three additional facilities he cites would alter Dr. Salling's conclusions. Indeed, the Secretary's expert did not even opine on Dr. Salling's proximity analysis. Doc. 51-2 at 26:12-21. In any event, the Secretary's contentions do not warrant exclusion of Dr. Salling's testimony.

IV. The Underlying Data Was Appropriately Checked For Accuracy.

The data check Dr. Salling performed was adequate to ensure the reliability of the data he used. Defendant complains that Dr. Salling did not do enough to verify the data he was given but points to nothing to suggest that more was required by law, academic practice, or any other standard. Doc. 53 at 9. For the TargetSmart dataset underlying the Incarcerated Voters Report, Dr. Salling testified that he did a "visual spot check" of the data he received "to see if there's anything that didn't seem to fit," and that he looked to confirm that the population included appeared to be eligible for inclusion in the analysis under the parameters he had been given and that he did not "recall finding any problems or errors." Doc. 50-1 at 120:20-121:12; 149:19-24. Such a review by

a trained expert is permissible. *See, e.g.*, *Little Hocking Water Ass'n v. E.I. du Pont de Nemours & Co.*, 90 F. Supp. 3d 746, 468 (S.D. Ohio 2015) (rejecting *Daubert* challenge where expert, with adequate "qualifications and background, . . . reviewed [the data] . . . and concluded, based on her general understanding of [the source's] methodologies and reasoning, that his data was reliable"). As explained herein, Dr. Salling received this dataset from a reputable third-party company and had no reason to believe that a more extensive check of the data was required. The visual check that Dr. Salling chose to do returned no issues.

Dr. Mockabee's report challenged the reliability of the matched dataset Dr. Salling used but importantly, he identified no errors in that dataset. *See generally* Doc 51-2. As Dr. Mockabee admitted in his deposition, publicly available data—including the public booking reports Plaintiffs provided and the public Ohio voter file under the Secretary's control—would allow the Secretary's expert to easily test the reliability of the matched dataset and scrutinize its accuracy. *Id.* 85:15-88:2.The Secretary chose not to do so and therefore his challenge to Dr. Salling's choice of data rings hollow. For the data used in the Proximity Report, Dr. Salling testified that he compared the geo-coded locations in his GIS program to the locations in the data list he was provided and used that to ensure the accuracy of the data. Doc. 50-1 at 68:1-12, 85:12-25. These were reasonable steps to take to ensure the accuracy of the data in the final reports.

V. The Data-Based "Assumptions" Dr. Salling Made Are Not Reason To Exclude His Testimony

Dr. Salling made appropriate fact-based assumptions in conducting his analysis for the Incarcerated Voters Report. As an initial matter, some of the purported "assumptions" identified by the Secretary are actually *facts*, others are irrelevant to Dr. Salling's analysis. First, the Secretary contends that "data derived from a national voter database contained assumptions that all of the voters identified were registered to vote in Ohio." Doc. 53 at 10. That is not an

assumption, that is a fact. The database maintained by TargetSmart that TMC used to create the matched dataset contains voter registration information for the State of Ohio obtained from the Secretary of State's voter rolls. Ex. B, Brill Decl. at ¶ 8; Ex. C, Burchard Decl. at ¶¶ 4, 6. A person would not be listed in the database as registered in Ohio unless that information appeared in the Secretary of State's own records. *Id.* Second, the Secretary challenges Dr. Salling's use of prison instead of jail population data. Doc. 53 at 10. But as Dr. Salling testified, the prison data was the best institutional population data available, and he has no basis to believe that the different data would lead to a different result—for it to make a difference, the comparative incarceration rate between counties would have to differ between jails and prisons, and there is no evidence for that. Doc. 50-1 at 143:22-145:5. And even if there were a difference, that would only affect Dr. Salling's estimations in one of two methods, and would not affect the actual data at all. Third, the Secretary argues that it was improper to assume "all voters who were jailed during the weekends before election days intended to vote and that their arrests were unforeseeable." Doc. 53 at 10-11. But Dr. Salling did not assert that all voters identified in his analysis "intended to vote" or that "their arrests were unforeseeable." Doc. 53 at 10-11. Instead, he concluded that these registered eligible voters were incarcerated after the absentee ballot deadline and thus would not have the opportunity to cast a ballot. Doc. 30-1 at 10. Plaintiffs submit Dr. Salling's expert report and testimony for that proposition and thus the alleged "assumptions" Defendant raises are irrelevant to this *Daubert* motion, nor do they rise to the level of warranting exclusion of an expert's testimony.

VI. Dr. Salling Correctly Applied Well-Accepted Methodologies From His Field

The Secretary's objections to Dr. Salling's methodology are largely recycled from his objections to the underlying data Dr. Salling relied upon. For similar reasons, those objections fail.

A. Dr. Salling Properly Relied Upon the Matched Dataset.

First, the Secretary objects to the use of the matched dataset because Dr. Salling could not fully explain the matching algorithm used by TMC. But relying on a dataset from a reliable third-party for some of the underlying facts of the analysis is an appropriate method for an expert to use. The Federal Rules allow an expert to "base an opinion on facts or data in the case that the expert has been made aware of or personally observed." *Ohio Organizing Collaborative*, 2016 WL 8201848, at *2 (quoting Fed. R. Evid. 703). "If experts in the particular field would reasonably rely on those kinds of facts or data in forming an opinion on the subject, they need not be admissible for the opinion to be admitted." *Id.* For example, this Court has previously found that it was appropriate for an expert to form the opinions in his report by relying on data collected by an experienced surveyor employed by a surveying company specifically for that case. *See Thomas & Marker Constr. v. Wal-Mart Stores*, 2008 WL 5119587, at *8-11 (S.D. Ohio Nov. 30, 2008). This was true even where the expert testified "that he never verified the methods used to determine the [data points] and never inquired as to whether the people who [created the data points] were qualified to do so." *Id.* at *8. Similarly, here it is appropriate for Dr. Salling to rely on the matched dataset that matched public booking report data against voter registration records, even though he was not provided with the underlying matching algorithm TMC utilized.[5] That dataset was prepared by a company that specializes in matching data to voter registration records and was retained specifically for this task. Ex. C, Burchard Decl. at ¶¶ 1, 3-7.

[5] TMC utilized an automated matching system licensed from a data vendor to match the booking report data against a national voterbase file maintained by TargetSmart. Ex. B, Brill Decl. at ¶¶ 7-8; Ex. C, Burchard Decl., at ¶¶ 6-7. The matching system utilizes a proprietary Bayesian Belief Net matching inference algorithm, which matches records based on several variables, such as first name, last name, and birthdate, and returns a match score indicating the probability that the two records match. Ex. C, Burchard Decl. at ¶ 7. TMC recommends use of a match score threshold of .4 and greater. *Id*.

Even though Dr. Salling did not need to know all of the details of the matched dataset, he both understood and explained the content of the dataset and its general method of creation. Both Dr. Salling's report and his deposition testimony make clear that the dataset he received from TMC was composed from two sources of underlying data, and a match score identifying the probability that the identities matched between individuals in the two sets of data. Doc. 30-1 at 4 and n.4; Doc. 50-1 at 111:20-112:18. The two sources of underlying data were (1) booking data received by Plaintiffs' counsel from the sheriffs of thirteen Ohio counties in response to public records requests, and (2) a national voter registration database, which was compiled from the voter registration databases maintained by each state. Ex. A, Diaz Decl. at ¶¶ 4-7; Ex. B, Brill Decl. at ¶¶ 7-8; Ex. C, Burchard Decl. at ¶¶ 4-8. The Secretary insists Dr. Salling should have used state voter data instead of national voter data, Doc. 53 at 10-11, but Dr. Salling has already made clear his understanding that the data in the national voter database that he used in his analysis in fact *included* the state voter data that Defendant contends should be used. Doc. 50-1 at 158:12-22. That understanding is confirmed by TargetSmart, the third party that maintains the national voter database that was used. Ex. B, Brill Decl. at ¶¶ 7-8.

Dr. Salling testified that he has performed similar analyses himself to determine the strength of a match between an individual's entries in two different datasets, one of which is a dataset of registered voters—i.e., determining whether the same person is included in both sets of data. Doc. 50-1 at 112:13-113:16. He testified that "one uses information that's available such as the name, the address, date of birth, et cetera, and anything else that might be useful for matching the two records." *Id*. 113:8-11. As set forth in Dr. Salling's report and the declarations attached herewith, that is in fact what TMC did in preparing the matched dataset. Doc. 30-1 at 4 and n.4; Ex. C, Burchard Decl. at ¶¶ 6-7.

Dr. Salling also adjusted his methodology to account for any potential uncertainty about the appropriate threshold to determine whether records in the two datasets matched. Dr. Salling understood from TMC that a .4 match score met the data group's industry standard of reliability. Doc. 50-1 at 135:24-136:5; *see also* Ex. C, Burchard Decl. at ¶ 7. Nonetheless, to account for any potential risk that such a match score might be insufficiently stringent, he performed his analysis of the estimated impacted voters at both more and less stringent matching thresholds (i.e., requiring both more and fewer indicia that a person in both datasets was the same) and comparing the differences in the resulting conclusions.[6] The final conclusions of total voter estimates in the Incarcerated Voters Report reflect this additional control.

Dr. Salling's use of a dataset from a reputable third-party as part of the underlying facts of his analysis is proper and provides no grounds to exclude Dr. Salling's Incarcerated Voters Report.

B. Dr. Salling Made Appropriate, Conservative Conclusions in His Incarcerated Voters Report

Dr. Salling's conclusions in his Incarcerated Voters Report are presented as a range of estimates for the number of incarcerated people affected by Ohio's absentee voting restrictions. This type of presentation properly accounts for uncertainties inherent in any estimation process and does not suggest that they are "technically imprecise, layman's conclusions." Doc. 53 at 16. Dr. Salling used two different methods to reach his estimate—one based on general population and the other based on institutional population; he also used different reliability cutoffs for inclusion in the data sample to account for factors that could bias an estimate. (*See supra*, Factual Background at 4.) Calculating an estimate using these two methods resulted in quite similar

[6] Match thresholds ranged "from 0 to 1.0, with higher scores indicating a higher likelihood that the person in the national database of registered voters is the same one in the database of incarcerated persons." Doc. 30-1 at 4. The most stringent match threshold went up to 0.9 in Dr. Salling's analysis and the least stringent was set at 0.3. *Id.* at 6.

numbers, all clustered around 1,000 affected people per election. The consistency of results across these methodologies suggests an overall level of reliability in the estimates. Doc. 50-1 at 165:8-25. The type of conclusion that is appropriate for an expert varies depending on his field of expertise—in the field of statistical analysis, a range of estimates is appropriate. In any event, the Rule 702 inquiry is not focused on an expert's conclusions, but rather only the data and methodology. *Buck v. Ford Motor Co.*, 810 F. Supp. 2d 815, 831 (N.D. Ohio 2011) ("[t]he important thing is not that experts reach the right conclusion, but that they reach it via a sound methodology") (citing *Daubert*, 509 U.S. at 595).

The Secretary makes much of an offhand, joking comment by Dr. Salling that he did not include some additional statistical analyses in his report because he was "lazy" and "running out of time," Doc. 53 at 14, but those additional numbers were in no way necessary or important to the report's final conclusions. Dr. Salling testified that he calculated a range and standard deviation for the numbers he reached for the 2018 election to test the reliability of his conclusions, and when that showed reliability of the data, for the other election years he did not perform those additional analyses, but "look[ed] at the data and they seemed to be very similar." Doc. 50-1 at 140:23-141:16. While the additional statistical measures are useful to see, they are not necessary and do not change the numbers Dr. Salling reached.

The Secretary contends that Dr. Salling admitted "that his numbers might be inaccurate." Doc. 53 at 9. But the Secretary overstates the import of the deposition testimony he relies on for this point. In the cited text, Dr. Salling observes that a number in the text of his report meant to calculate the average of his estimates is smaller than the data in the corresponding figure, and states that the average number may be a mistake. Doc. 50-1 at 138:2-14 (referencing Doc. 30-1 at 7,

Figure 2). This typographical error in a summary paragraph does not call into question the numerical conclusions Dr. Salling reached, which are clearly set forth in the Figure.

C. Dr. Salling's Proximity Geocoding Was Reliable.

Dr. Salling used an appropriate methodology by calculating the distances between county Boards of Election and jails and hospitals, respectively, using a straight-line calculation rather than a driving distance calculation. The Secretary argues Dr. Salling gave no rationale for analyzing distance in this way and did not conduct an analysis to see whether changing that method would result in any "appreciable difference." Doc. 53 at 8-9. But during his deposition, Dr. Salling explained that he used the tools available in GIS software and SAS programming to calculate distance (Proximity Report at 1), and would not expect any meaningful difference between straight line and driving distances for these types of measurements. Doc. 50-1 at 155:1-10. He explained that he had "done previous network analyses . . . and the results are really rarely any different than as the crow flies, especially if you look at within a county. *Id.* at155:15-19. At most, he said, there might be "a difference of a minute or two." *Id.* at 155:21-22.

VII. Conclusion

For the foregoing reasons, the Secretary's motion should be denied.

Dated: August 16, 2019

Respectfully submitted,

/s/ Naila S. Awan
Naila S. Awan, Trial Attorney (0088147)
Kathryn C. Sadasivan*
DĒMOS
80 Broad Street, 4th Floor
New York, NY 10004
(212) 485-6065
nawan@demos.org
ksadasivan@demos.org

Chiraag Bains*†
DĒMOS
740 6th Street NW, 2nd Floor
Washington, DC 20001
(202) 864-2746
cbains@demos.org

Attorneys for Plaintiffs

Mark P. Gaber*
Danielle M. Lang*
Jonathan M. Diaz*
CAMPAIGN LEGAL CENTER
1411 K St. NW, Ste. 1400
Washington, DC, 20005
(202) 736-2200
mgaber@campaignlegalcenter.org
dlang@campaignlegalcenter.org
jdiaz@campaignlegalcenter.org

Locke E. Bowman*
Alexa Van Brunt*
Laura C. Bishop*
RODERICK AND SOLANGE MACARTHUR JUSTICE CENTER
NORTHWESTERN PRITZKER SCHOOL OF LAW
375 East Chicago Avenue
Chicago, Illinois 60611
(312) 503-1271
l-bowman@law.northwestern.edu
a-vanbrunt@law.northwestern.edu
laura.bishop@law.northwestern.edu

**admitted* pro hac vice
†admitted to practice only in Massachusetts; practice limited pursuant to D.C. App. R. 49(c)(3)

CERTIFICATE OF SERVICE

The foregoing Plaintiffs' Response to Defendant's Motion, *In Limine*, to Exclude the Expert Testimony of Dr. Mark Salling Ph.D was filed this 16th day of August, 2019 through the Court's Electronic Filing System. Parties will be served, and may obtain copies electronically, through the operation of the Electronic Filing System.

/s/ Naila S. Awan
Naila S. Awan (0088147)