Evaluating Estimates of the Number of Incarcerated Registered Voters in Ohio Who Could Not Vote in the November 2012 General Election

Prepared by:

Stephen T. Mockabee, Ph.D.
Associate Professor and Graduate Program Director
Department of Political Science
University of Cincinnati

January 13, 2014

Introduction and Overview

I was asked by the Ohio Attorney General’s office to provide my professional opinion on the data analysis conducted and conclusions reached in a report entitled, “Analysis of the Number of Incarcerated Registered Voters who could not Vote During Their Incarceration on the Weekend and Monday before the November 2012 General Election,” authored by Dr. Mark Salling. This report was conducted by Dr. Salling at the request of the Ohio Justice and Policy Center (hereafter OJPC) in regard to the Fair Elections Ohio v. Husted case. The Ohio Attorney General’s office provided me with electronic copies of Dr. Salling’s final report, the data files used by Dr. Salling to develop estimates of the numbers of incarcerated individuals, and email correspondence between Dr. Salling and the OJPC staff regarding the data collection and analysis process.

Based on the information given to me, my professional assessment is that the OJPC report does not provide the analysis necessary to make an accurate estimate of the number of Ohio registered voters incarcerated the weekend before the election who were unable to vote in the 2012 general election. I find no evidence in the OJPC report demonstrating that the nine Ohio counties analyzed—Adams, Athens, Butler, Cuyahoga, Franklin, Hamilton, Lawrence, Montgomery, and Summit—are representative of the state as a whole with respect to the incarceration of registered voters in the three days prior to the election. The OJPC report asserts that the nine counties provide a basis for generalizing, but offers no justification for this claim. The OJPC report does not include any information about the criteria used to select the nine counties studied, and provides no assessment of the extent to which using these particular counties might distort the overall statewide estimate. Therefore, estimating a statewide total by extrapolating from these nine counties is problematic. In sum, although the OJPC report identifies some registered voters in the nine counties analyzed who were unable to vote due to incarceration, it does not present sufficient information to estimate a statewide total.

In the remainder of the report I provide a more detailed assessment of the data and methods used in the OJPC report and further explanation of my conclusions.
Assessment of Data and Analysis

The type of analysis undertaken by Dr. Salling for the OJPC report is fairly labor intensive in the sense that it involves the use of multiple data sets and has to confront problems of data availability and compatibility that often come up in these sorts of projects. In other words, because different counties keep their records differently, the analyst has to engage in various matching techniques to come up with a workable set of data. My assessment is that Dr. Salling followed a reasonable set of steps to identify incarcerated individuals who were registered but did not vote. Because I did not find significant methodological problems with the matching exercise, I did not undertake a replication of Dr. Salling’s findings for the nine counties studied. It is possible that, as Dr. Salling rightly acknowledges in the report, errors could have been made in the matching process. However, my judgment based on prior experience working with voter files is that errors in matching are more likely to be caused by missing or incomplete information in the records provided to Dr. Salling than by a systematic problem with the approach Dr. Salling took in identifying matches. Having not replicated the analysis, I have no opinion to offer regarding the number of errors in matching that may have occurred.

In my professional opinion the significant methodological problem with the OJPC report exists in the phase of research design that social scientists refer to as case selection—that is, the process of choosing the units (cases) from which data are collected. Ideally data would be collected from all units in a population of interest (in this case all 88 counties in Ohio), but as a practical matter it is at times difficult to obtain data from all units due to constraints such as time or cost. If it is not feasible to collect data from all units in a population, a process of selecting a subset of cases needs to be undertaken so as to yield a group of cases for analysis that will be representative of the population as a whole. In other words, the cases selected should be able to “stand in” for those cases not included in the analysis. If the cases selected for analysis differ significantly on key variables of interest compared to the rest of the cases not selected, then the estimates drawn from the data analysis will be inaccurate. Therefore, a critical component of research design is developing clear and compelling criteria for the selection of cases that are to be included in a study. Only when the selection criteria are reasonable can a subset of cases be assessed as being potentially representative of the larger population of interest.

The OJPC report does not provide the criteria that were used in selecting the nine counties used in the report, leaving the reader to speculate as to the possible reasons why these counties were chosen. I do not know whether the OJPC attempted to obtain data from all Ohio counties. Having data from all counties would be the most desirable way to proceed with the analysis because there would be no distortion in the estimates due to taking a subset of cases rather than examining the entire state.1 If obtaining data from all counties is deemed to be infeasible, then a process of case selection should be developed that accounts for differences across counties on variables (characteristics) that are expected to be related to the outcome variable of interest (the number of registered voters who could not vote in the 2012 general election because they were incarcerated). For example, it would be desirable for the counties selected for analysis to be representative of the state as a whole with respect to the rate of incarceration on the weekend before Election Day. If a particular county incarcerates people at a high rate, this makes it more likely that registered voters in that county will be unable to vote due

---

1 Errors could still emerge due to factors outside the researcher’s control such as inaccuracies in county records. But any inferential errors due to case selection would be eliminated if data were obtained from all 88 counties.
to incarceration. If the counties selected for analysis in the OJPC report have higher than average incarceration rates, then the estimate of the number of registered voters who are unable to vote due to incarceration will be inflated (biased upward). Similarly, if the selected counties had a lower than average incarceration rate for the weekend before the election the estimate of incarcerated voters would be too low.

In seeking to extrapolate from the nine selected counties based solely on their share of Ohio’s total adult population, the OJPC report implicitly makes the assumption that the counties selected for analysis are representative of the state as a whole with regard to incarceration rates on the weekend before Election Day. However, no information is offered to support this assumption.

There is reason to suspect that the nine counties used in the OJPC report may not be typical of Ohio counties in terms of their incarceration rates. These nine counties include the five most populous counties and the three largest urban centers in the state. If urbanicity is related to incarceration rates, then the OJPC sample of counties would overstate the extent of incarceration and therefore would over-estimate the number of registered voters unable to vote due to incarceration. To be clear, I have not reviewed data that demonstrate a statistical association between urbanicity and incarceration rates, and whether or not such a relationship exists is an empirical question. I mention the possible relationship here to underscore the point that the OJPC report makes an unsupported assumption about the representativeness of the nine counties examined. More information is needed before one can confidently assess the representativeness of the counties studied in the OJPC report.

Given the concerns mentioned above, I do not place a high degree of confidence in the OJPC estimate of 479 registered voters statewide who were unable to vote in the 2012 general election because they were incarcerated the weekend before the election. However, it is worth noting that even if one accepts this estimate as reliable, the number is quite small relative to the statewide electorate. According to the official election results posted at the Ohio Secretary of State’s web site, there were 7,987,203 registered voters and 5,633,246 votes cast in the 2012 general election. Thus the OJPC estimate of 479 voters would be equal to 0.0006% of registered voters and 0.0085% of voters who cast ballots.

---

2 http://www.sos.state.oh.us/SOS/elections/Research/electResultsMain/2012Results.aspx, last accessed January 7, 2014