

New Statistical Method Linking Databases Shows Negative Effect of Texas Voter Identification Law

ALEXANDRIA, Va. (December 11, 2017) – As state voter identification (ID) laws across the country are being contested amid questions about the integrity of the voting process, researchers have developed a new statistical method that not only matches multiple records with precision, but can also identify the scope of discrimination when applied to voter ID laws. Recently featured in the American Statistical Association’s journal *Statistics and Public Policy*, the research titled “[ADGN: An Algorithm for Record Linkage Using Address, Date of Birth, Gender and Name](#)” was applied to a 2011 Texas voter ID law (S.B. 14), which the United States Department of Justice investigated as possibly discriminatory.

“Our evidence suggests a smaller number of people lack ID than recent survey evidence suggests, and it also suggests a discriminatory effect of the law, in line with concerns of those who believe these laws disproportionately affect minorities,” noted Eitan Hersh, associate professor of political science at Tufts University and co-author of the paper. “Specifically, we found that white registered voters are significantly more likely to possess a voter ID than African-American or Hispanic voters.”

Hersh and fellow researcher Stephen Ansolabehere, professor of government at Harvard University, designed a way to link individual records across databases. Such linking is challenging because databases tend to have errors and missing values and often lack unique identifiers such as social security numbers. Their efforts using combinations of address (A), date of birth (D), gender (G), and name (N) match records about as accurately as if they had an individual’s social security number.

“Voter identification laws have become a source of debate and controversy, and we think this is the strongest evidence to date about the magnitude and the discriminatory effect of laws on protected minorities versus white voters,” continued Hersh.

Typically, voter identification laws are studied by conducting and analyzing data from surveys, although surveying people to find out if they have valid IDs—and whether an invalid ID stopped them from voting—remains difficult. In the case of the Texas voter identification law, findings were the result of matching 13 million registered voters to all forms of identification that could be used by voters showing an ID.

Application of the ADGN model is not limited to voter registration and identification litigation. In fact, Hersh and Ansolabehere believe the ability to link records with a high degree of

accuracy is informative to research, in general, as well as to applications in a range of areas such as public health, criminology, marketing and government censuses.

###

About the American Statistical Association

The ASA is the world's largest community of statisticians and the oldest continuously operating professional science society in the United States. Its members serve in industry, government and academia in more than 90 countries, advancing research and promoting sound statistical practice to inform public policy and improve human welfare. For additional information, please visit the ASA website at www.amstat.org.

For more information:

Jill Talley

Public Relations Manager

(703) 684-1221, ext. 1865

jill@amstat.org