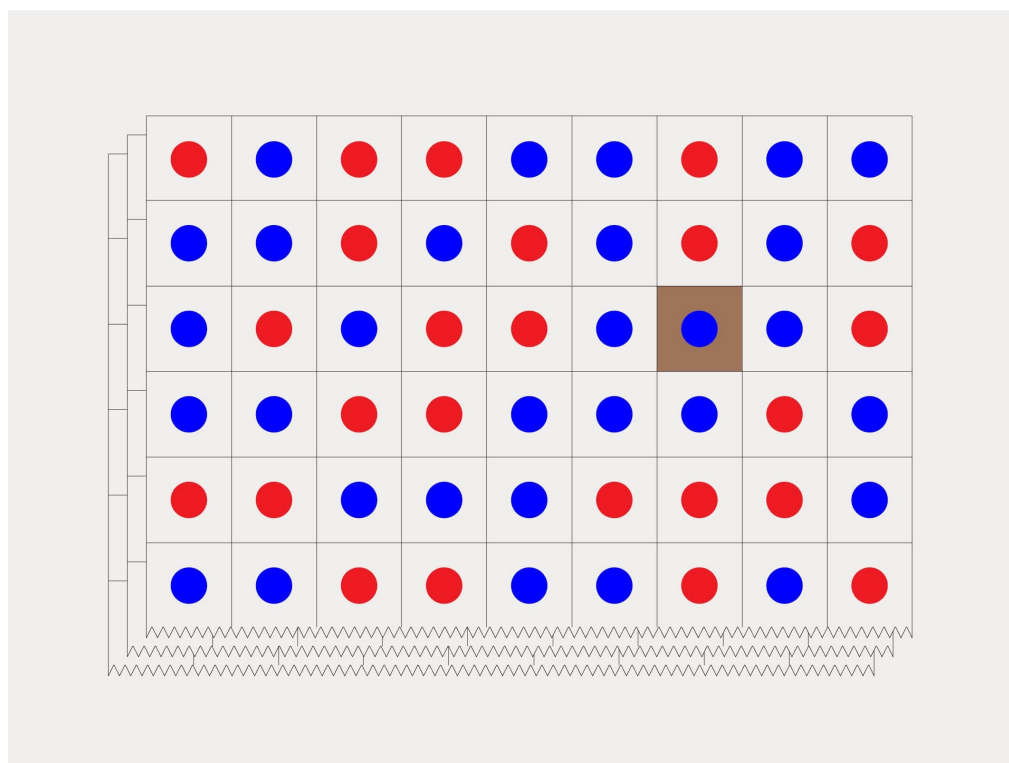


Voter ID Laws Really Do Discriminate, a New Study Shows

Author: Issie Lapowsky



Being able to match voters with their records in ID databases using just a few basic details might help dispel some myths about whom laws do and don't hurt.

Frank Augugliaro

After announcing the closure of his short-lived commission to end voter fraud, President Trump made it clear Thursday that he wants more states to require identification at the ballot box to prevent what he believes is rampant—but still unproven—election rigging.

Ever since the Supreme Court struck down a key part of the Voting Rights Act in 2013, laws requiring voters to show identification when they vote have speckled the nation, popping up in states from Rhode Island to Arizona. Almost as quickly, voting rights advocates have taken states like Texas and Alabama to court, arguing that these laws intentionally discriminate against minority voters. Just last summer, a federal judge tossed out Texas's voter ID law, in a case that's now being

[revisited](#) by an appeals court. But proving exactly how discriminatory these laws are requires far more complexity than it might seem.

Sure, there are endless anecdotes of well-meaning, well-prepared citizens being turned away on election day, but anecdotes are not data. There are ample surveys asking voters whether these laws came between them and the ballot box, but people can always misrepresent themselves on surveys, and courts tend to dismiss them, anyway. Seven states include Social Security numbers in voter files and driver's license records, but across the rest of the country, determining whether a single individual voter is also listed in any number of identification databases has become a complex and nettlesome problem for voting access advocates and statistics researchers alike.

Recently, however, researchers at Tufts University and Harvard University demonstrated that it's possible to match individuals across government databases with nearly perfect accuracy, using just a few basic identifiers like a person's name, date of birth, and address. They developed the algorithm while working as expert witnesses in the Department of Justice's case against Texas. Now, in a newly [published paper](#), researchers Stephen Ansolabehere of Harvard and Eitan Hersh of Tufts have explained the underlying methodology. Their goal, according to Hersh, is to create a system courts can easily understand, which can not only be used in future voter ID law cases, but can also help dispel some myths about who those laws do and don't hurt.

"The more we can agree on methods that are easy to explain, the better off we are," says Hersh.

A Better Model

If all data were clean and complete, it wouldn't be so hard to figure out if a voter named, say, John Smith, was the same John Smith listed in the federal driver's license database. According to Hersh and Ansolabehere's research, only 1 in 2.7 billion individuals have the same zip code, gender, date of birth, and last name, making those four details combined a fairly spot-on indicator of a person's identity. But often, government records contain typos, incomplete fields, nicknames, or outdated addresses. To link databases, researchers often need to make do with less information.

They also need to be able to show their work in a way that lawyers and judges can understand. So many algorithms that purport to match people across databases run up against the so-called black box problem. They may be able to make statistically sound decisions, but they can't easily explain how they made them. In a recent [Supreme Court hearing over partisan gerrymandering](#) in Wisconsin, Chief Justice John Roberts dismissed research-backed methods to measure gerrymandering as "sociological gobbledygook." Hersh and Ansolabehere wanted to develop a tool that could be easily understood.

'The more we can agree on methods that are easy to explain, the better off we are.'

Eitan Hersh, Tufts University

So, working with the Department of Justice, the researchers set out to determine whether they could match voters on the voter roll with their corresponding records in ID databases using just a few basic details. To do that, they developed an algorithm that scanned the state of Texas's voter rolls and compared it to the federal list of driver licenses, state IDs, and concealed handgun permits, among other forms of acceptable identification. It scanned each record by address, date of birth, gender, and name, to see if, for instance, a combination of address, gender, and name would be as accurate a predictor as a combination of date of birth, gender, and name.

To check their results, the researchers relied on a subset of the voter data that contained Social Security numbers. Those records effectively served as the algorithm's answer key. They ultimately found that 98 percent of the records that could be matched using Social Security numbers could also be matched using any three of the four key identifiers—address, date of birth, gender, and name.

"This combination is as good as a Social Security number," Hersh says.

That high accuracy rate is essential in court, says Charles Stewart III, a political scientist at Massachusetts Institute of Technology, who has served as an expert witness in a case against South Carolina's voter ID law. Most database-matching algorithms are used in low-risk scenarios, he explains, like advertising, where companies want to target customers across a range of platforms. If they target the wrong customer, at worst,

they've lost a marginal amount of money. In court, it could mean losing the case altogether. "There is really no room for error," Stewart says, citing the risk of having to demonstrate an algorithm's chops before the bench. "If the judge doesn't get matched properly, for instance, you might as well have not done anything."

Deep Impact

Once Hersh and Ansolabehere were confident they had properly matched registered voters to their ID records, they used a commercial tool called Catalist to predict each voter's race. That tool analyzes names to determine how likely a given name is to be associated with one race or another. It also accounts for the demographics of the Census block where a given voter lives. Using this tool, the researchers confirmed what voting rights advocates already know to be true—that black voters are more likely to lack adequate identification under voter ID laws. According to the study, 3.6 percent of registered white voters had no match in any state or federal ID database. By contrast, 7.5 percent of black registered voters were missing from those databases.

The algorithm shows a clear and disturbing racial disparity on voting rights. But Hersh says that it also shows that voter ID laws affect a relatively small percentage of the population. Across all registered voters in Texas, the researchers found 4.5 percent lack proper identification. For registered voters who actually showed up at the polls in 2012, it's 1.5 percent.

"You're down to a small percentage of the population that doesn't have an ID," says Hersh. That's one reason why, despite Alabama's restrictive voter ID law, [black turnout in the recent Senate election](#) still exceeded expectations. Still, while the percentages may sound small, that 4.5 percent still represents 608,470 Texas citizens who could potentially be disenfranchised.

Hersh says he agrees the public ought to be outraged by racially motivated attempts to suppress the vote, and that courts ought to crack down on the practice. But he cautions against Democrats artificially inflating the impact of voter ID laws in their messaging. Shortly after the 2016 election, Democratic Senator Tammy Baldwin attributed the entire dropoff in voter turnout in Wisconsin to the state's new voter ID

law. Hersh points to plenty of other reasons why voter turnout may have been down in Wisconsin in 2016, including the simple fact that Hillary Clinton didn't campaign in Wisconsin.

"It ... can be dangerous to sow doubt in the integrity of the election system with these big claims," Hersh says. Besides, the raw data is bad enough.

And Hersh's findings are limited to Texas. In other states with less of a car culture and more densely populated cities, the number of people who lack ID may be as high as 10 percent, MIT's Stewart explains. And because this study only looks at voters who are currently registered, Stewart says, it doesn't take into account the discouraging effects that voter ID laws may have on people attempting to register in the first place.

"Whether it's intended to harm 600,000 African American and Latino voters or 2 million, our concern is people are passing these laws with the intent to discriminate or the effect of discriminating," says Deuel Ross, a civil rights attorney for the Legal Defense Fund.

Now, science is bringing people like Ross that much closer to proving it.

Update: 8:43 am ET 01/04/17 This story has been updated to include the President's tweet and news about the voter fraud commission's closure.