
State Voter Registration Databases

Impact of Technical Implementation on Policy

Michael P. McDonald

- Associate Professor, George Mason University
- Non-resident Senior Fellow, Brookings Institution

List Integrity Has Policy Implications

- Efficient Election Administration and Vote Caging
 - Allegations of Vote Fraud
 - Cross-Database Matching Procedures
-

What is in a Name?

The 2004 Washington Voter Registration File

| Extract | Names with Apostrophe | McDonalds or Mc Donald with Hyphen | McDonald with Space | McDonald With Space Removed | Total (Duplicates Removed) |
|---------|-----------------------|------------------------------------|---------------------|-----------------------------|----------------------------|
| | Number of Occurrences | 3,772 | 2,234 | 908 | |
| Total | 1,557 | 35,852 | 50,800 | 209,456 | 295,962 |

3.8 million voter registration records

Terminator Voters

stats from “The True Electorate” *Public Opinion Quarterly*, 2007



In New Jersey, “I’ll be back, to vote early and often”
entire Middlesex County missing birth dates!

Registration Date+18 years: 2004 Voters with Missing Birthdate

stats from “The True Electorate” *Public Opinion Quarterly*, 2007

| State | Missing Age | Age - 4 Registration Category | | | | |
|-------|-------------|-------------------------------|-------|-------|-------|---------|
| | | 18-29 | 30-44 | 45-59 | 60+ | Missing |
| CA | 0.9% | 3.7% | 4.5% | 69.4% | 22.4% | 0.0% |
| DC | 4.5% | 1.1% | 7.6% | 91.3% | 0.0% | |
| OH | 3.1% | 1.6% | 23.2% | 36.6% | 33.6% | 5.0% |
| OK | 1.4% | 0.7% | 1.3% | 2.1% | 9.1% | 86.8% |

Common Address Errors

- Misspellings
 - Missing Apartment Numbers (Dorms)
 - Missing Street Directions
 - Missing or Incomplete Zip Codes
-

Policy Implications

- Inefficient contact with voters
 - Allegations of vote fraud
 - Undeliverable mail is not necessarily an indicator of fraud
 - Missing apartment numbers are common, and may disproportionately affect minorities
 - Allegations of double voting
-

Allegations of Double Voting

from “Seeing Double Voters” *Election Law Journal*, forthcoming(?)

- 2004 New Jersey: 4,397 people
 - 2000-02 Missouri: over 300 people
 - 2004 Florida and New York: 400-1,000
-

Method of Formulating Allegation

1. Obtain a state's voter registration file
 2. Find all voters with the same first name, last name, and birth date
 3. Assume that these people are double voters
-



©1997 MAGELLAN GeographixSM
(805) 685-3100 www.maps.com

Double Voters or Mistaken Identity?

- What if two people simply have the same first name, last name, and exact birth date (including year)?



How many of these people
could be expected to have the
same exact birth date?

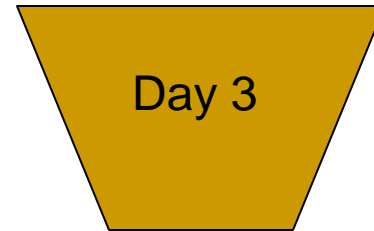
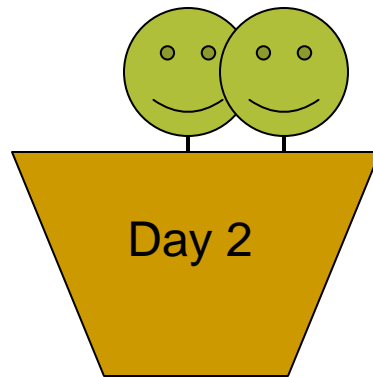
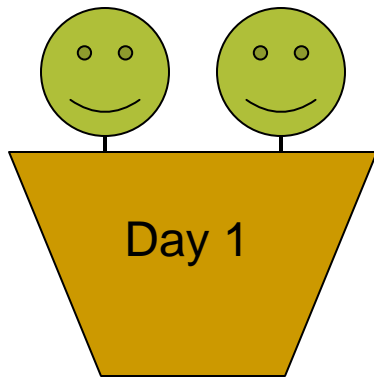
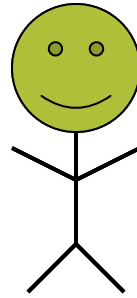
An application of 'The Birthday Problem'

The Birthday Problem

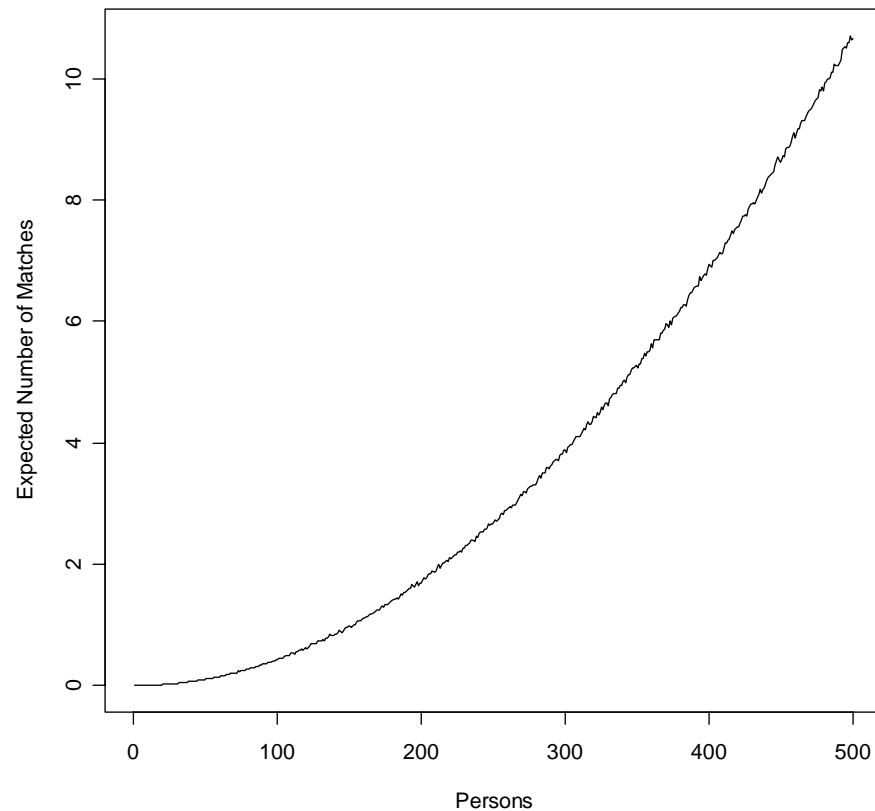
- Question: How many people do you need in order to have a fifty percent probability that two people out of the group share the same birthday (disregarding year)?
 - Answer: 23
-

Extending the Birthday Problem to the Birthdate Problem

- Need to account for years
 - Need to know the expected number of matches (not the probability of a single match)
 - This is difficult! But it is easy to figure out using simulation
-



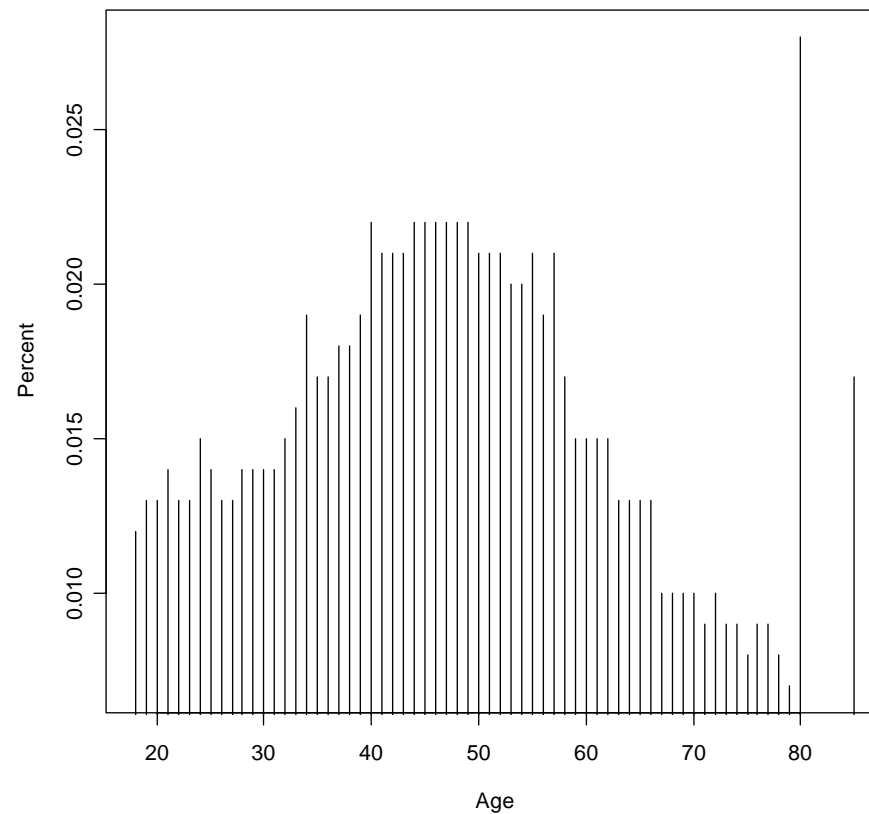
Expected Number of Matches (Uniform Birthdate Distribution)



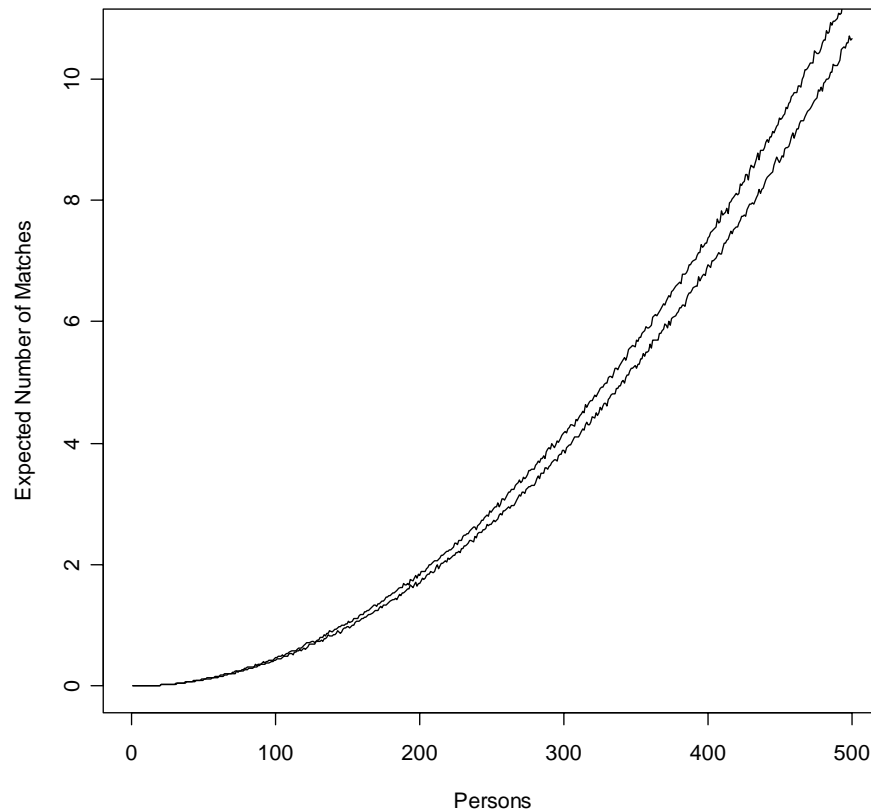
Complications

- Fewer births on weekends
 - Name periodicity effects (e.g., naming children after famous people, migrant patterns)
 - Voters tend to be older
 - Deviations from a uniform distribution of birthdates increases the expected number of matches. Our estimate of matches is therefore smaller than it should be.
-

2004 Current Population Survey Age Distribution of Voters



Expected Number of Matches (CPS Birthdate Distribution)

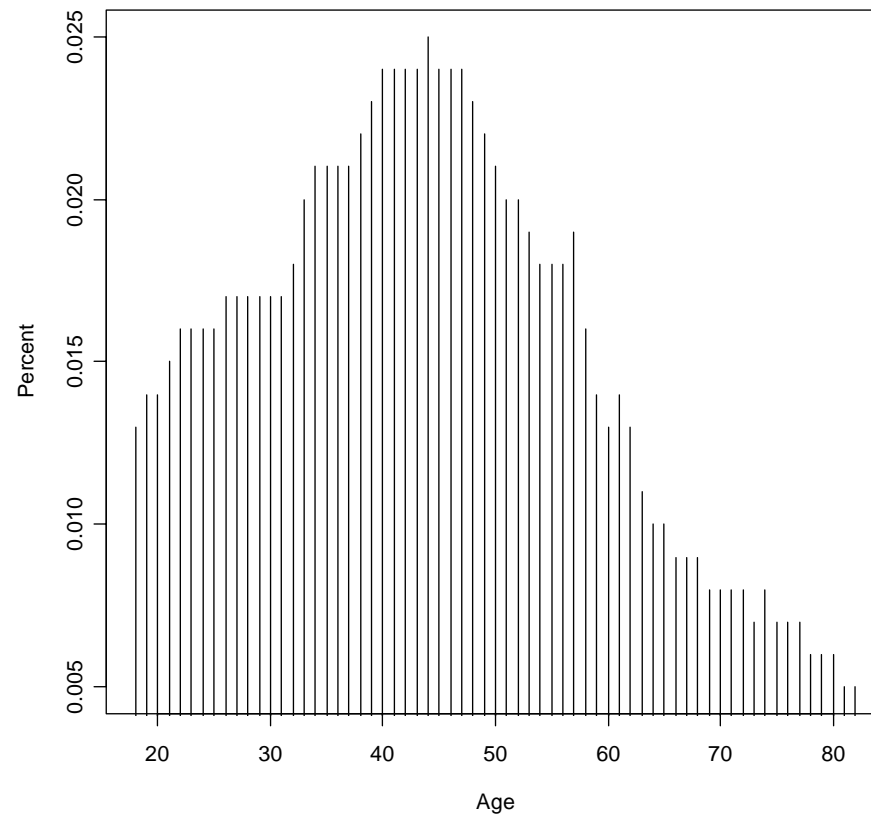


Case Study:

2004 New Jersey General Election

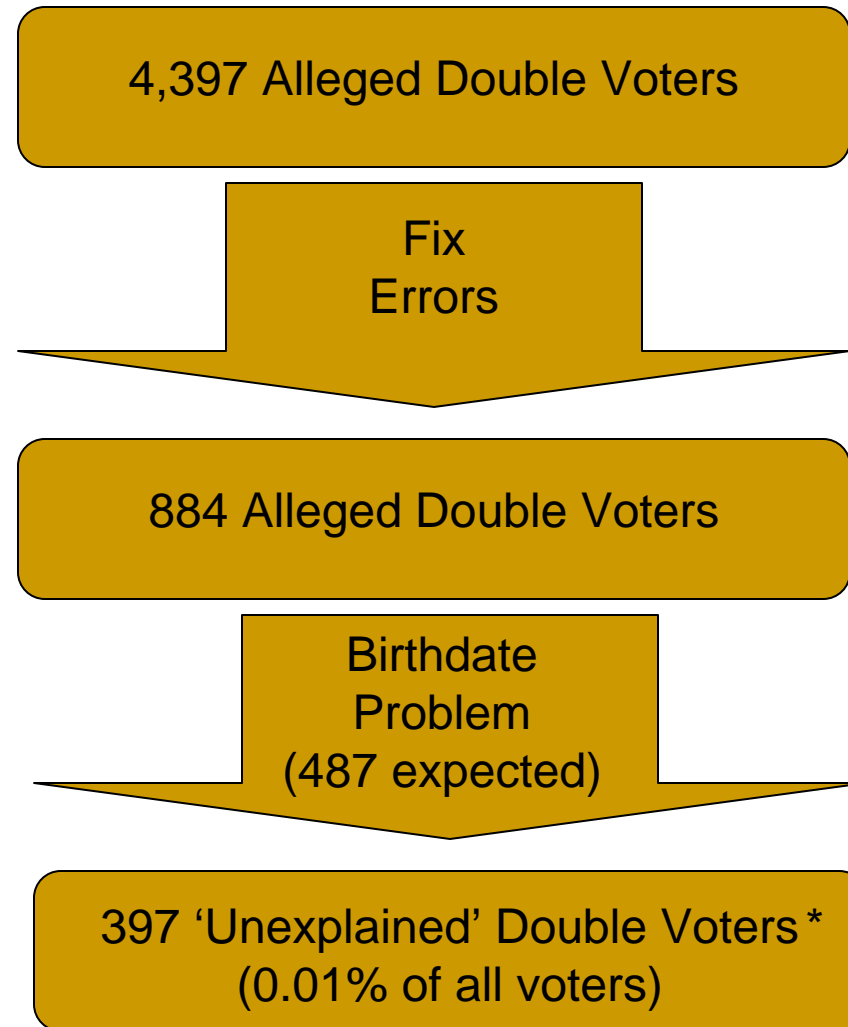
- Obtained allegations from New Jersey Attorney General.
 - Obtained 2004 New Jersey voter file from SEIU
 - Identified Obvious Errors:
 - ❑ Missing middle names, missing suffixes
 - ❑ Missing birth dates (all of Middlesex County)
-

2004 New Jersey Voters Age Distribution



| All Registered Voters | | | Registered Voters with 2004 Vote History and Valid Birthdate | | |
|-----------------------|-----------|-----------|---|-----------|-----------|
| First Name | Last Name | Frequency | First Name | Last Name | Frequency |
| ROBERT | SMITH | 424 | ROBERT | SMITH | 325 |
| JAMES | SMITH | 374 | WILLIAM | SMITH | 282 |
| WILLIAM | SMITH | 371 | JAMES | SMITH | 280 |
| JOHN | SMITH | 326 | JOHN | SMITH | 243 |
| JOSE | RODRIGUEZ | 263 | MICHAEL | SMITH | 220 |
| MARIA | RODRIGUEZ | 262 | MARIA | RODRIGUEZ | 209 |
| MICHAEL | SMITH | 259 | JOSE | RODRIGUEZ | 206 |
| ROBERT | BROWN | 257 | ROBERT | BROWN | 193 |
| ROBERT | JOHNSON | 257 | ROBERT | JOHNSON | 193 |
| MARY | SMITH | 255 | THOMAS | SMITH | 192 |

New Jersey's 3,580,074 Voters...



* Further investigation may reveal other reasonable explanations

Policy Implications

- It is possible that some people voted more than once, and these cases should be treated with all due diligence
 - Allegations of double voting are exaggerated, both by data errors and mistaken identity
-

False Negative and False Positive Matches

- False Negatives: A person may be in both databases, but will not match because identifying data is not an exact match
 - False Positives: By dumb luck (and poor matching procedures) a matched record may be two different people
-

Policy Implications: Voting Process

- Need to implement *convenient* provisional ballot procedures to resolve registration problems that arise from no fault by voters, such as an incorrect automatic change of address through database matching.
 - ❑ An alternative to Election Day Registration: Extend NVRA required registration portability within local jurisdictions to an entire state (CO, CT, DE, DC, FL, MD, OH, SD, WA, WV).
 - ❑ An alternative to precinct-based voting: vote centers (e.g., CO, TX).
-

Policy Implications: Purging, Vote Fraud, and At-Risk Photo ID Populations

- Need to provide adequate safeguards and verification procedures when purging voter registration records...
 - ...and more generally claiming that a match (or lack of a match) is evidence of fraud or a voter has a given identification
-