



New Challenges to Political Privacy:

Lessons from the First U.S. Presidential Race in the Web 2.0 Era

Daniel Kreiss, Philip N. Howard

Presented by: Yang Yu, Zhouxiang Cai, Junli Wu



Overview

- Political campaigns were revolutionized by big data and the internet
- Regulatory on political data is lacking
- Should we be concerned?



1960s - 1970s

Beginning of modern data practices of campaigns


- Direct primaries → voters determine candidates, campaigns became independent and autonomous
- Political consultants → dedicated profession for campaigning
- Advancement in computing → data became a manageable asset



1980s - 1990s

Data began to gain importance

- Prospect direct mail → micro-target discrete individuals
- Use data to get better knowledge of electorate → magazine subscription lists
- Political consultants specialize → marketing, communications, polling
- Republican party started to develop national databases
- Democratic party was a little behind



1990s -2000s

Internet emerged

- First as a broadcast medium
- Tool for mobilization



Internet age

Democrats lead the development

- Online campaigns allow users to participate as content creators → use of personal websites and social media like MeetUp
- Data infrastructure → CRM systems for political purposes



Large Voter Databases

- The Republican Party
 - Developed in mid-1990s called Voter Vault
 - 168 million citizens by 2004
- The Democratic Party
 - VoteBuilder launched in 2007
 - Centralized database shared by all state parties
 - Use of commercial databases
 - Catalist: 450 data points on 250 million people
 - Aristotle: largest bipartisan provider of data



What information is included?

- Public data
 - Voter registration, party identification, turnout records, vehicle registration, ZIP code census data, etc.
- Commercial data
 - Magazine subscription, credit card purchases, etc.
- Dedicated firms to find data correlation and build voter profiles



The Obama Campaign

- E-mail system of 13 million members
- 7 million cell phone numbers
- Cross-platform sync of data between BarackObama.com, MyBarackObama.com (2 million users), Facebook (2.4 million users), and email lists
- 223 million pieces of data given voluntarily by citizens during the campaign
- New media division to analyze the data
- Large centralized database of democratic (potential) supporters

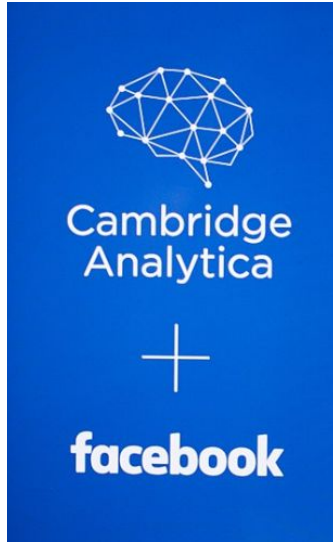


Campaign process was changed

- Dependency on data tools and private consulting firms
- Better data allow campaigns to be more focused
- Data from social media are integrated
 - Obama campaign gathered data from people who friended the campaign's facebook page
- Increase spending on online advertising



Growth of tech company





The U.S. Regulatory Context

- No regulation with respect to the use of data on citizens long time ago
- Non-state actors, candidates and parties' ability to speak to citizens is protected by the First Amendment (Turow, 2006).



The U.S. Regulatory Context

- The regulation for the official state agencies and parties are different
- Executive branch VS State actors
- These regulation do not extend to the actions of candidates
- Candidates for higher office have a general right of free political expression



The U.S. Regulatory Context

- There is a general lack of regulation and broad protection for its political use.
 - California Voter Foundation (Alexander & Mills, 2004).
- The increasing sophistication of database technology is raising significant new concerns for democratic practice.



Campaign Data and Democratic Citizens



Overview

- The Lack of Transparency in Political Data
- Data Security
- Political Privacy and Freedom of Association
- Data and the Democratic Deficit



The Lack of Transparency in Political Data

- Political data is among the most valued commodities in the market for consultant services.
- Political data receives scant public attention



ARISTOTLE™



Alan Westin Surveys

In the 2003 study, to obtain the privacy index, Westin asked the following question:

For each of the following statements, how strongly do you agree or disagree?

- 1. Consumers have lost all control over how personal information is collected and used by companies. (69% agreeing)
- 2. Most businesses handle the personal information they collect about consumers in a proper and confidential way. (54% disagreeing)
- 3. Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today. (53% disagreeing)

Following options were provided to the respondents to choose from:

Strongly Disagree, Somewhat Disagree, Somewhat Agree, Strongly Agree



Alan Westin Surveys

- Privacy Fundamentalists (26%)
- Privacy Unconcerned (10%)
- Privacy Pragmatists (64%)

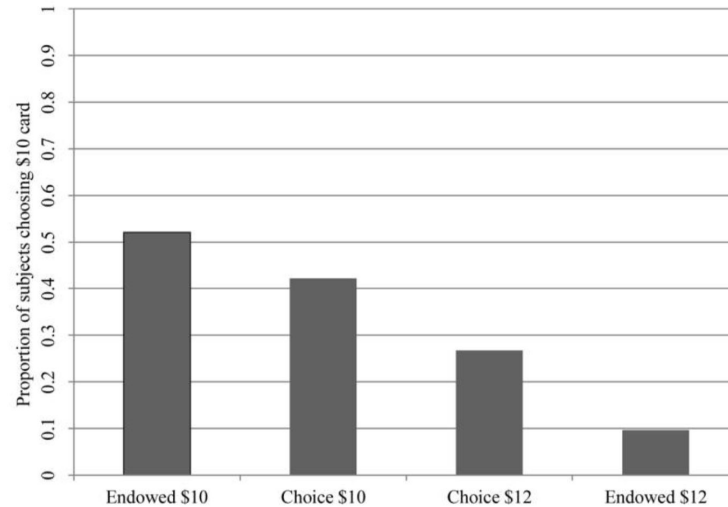


Endowment Effect Experiment

- \$10 endowed: Keep the anonymous \$10 card or exchange it for an identified \$12 card.
- \$12 endowed: Keep the identified \$12 card or exchange it for an anonymous \$10 card.
- \$10 choice: Choose between an anonymous \$10 card and an identified \$12 card.
- \$12 choice: Choose between an identified \$12 card and an anonymous \$10 card.



Endowment Effect Experiment - Result





Data Security

- Political data sold to shore up the finances of failing firms
- Campaign websites hacked
- Questionable data practices exposed
- American voter data has turned up in surprising places overseas.



Data Security

Equifax Data Breach Settlement

January 2020

In September of 2017, Equifax announced a data breach that exposed the personal information of 147 million people. The company has agreed to a global settlement with the Federal Trade Commission, the Consumer Financial Protection Bureau, and 50 U.S. states and territories. The settlement includes up to \$425 million to help people affected by the data breach.

Business

Marriott discloses massive data breach affecting up to 500 million guests

Facebook data breach sees millions of user personal details leaked online

By Mike Moore December 24, 2019

267m Facebook users could have had their contact details exposed

Cambridge Analytica and Facebook: The Scandal and the Fallout So Far

Revelations that digital consultants to the Trump campaign misused the data of millions of Facebook users set off a furor on both sides of the Atlantic. This is how The Times covered it.



Political Data are Special

General Data Protection Regulation (GDPR) identifies special categories consisting of information related to an individual's:

- racial or ethnic origin
- political opinions
- religious beliefs or other beliefs of a similar nature
- ...

The data held by campaigns and parties is explicitly used for political purposes, and thus reveal levels of detail about policy preferences and ideological perspectives that are less relevant to the advertisers of consumer goods.



Political Privacy and Freedom of Association

- Privacy helps ensure robust public debate by providing the opportunity for citizens to form their own viewpoints, craft arguments, and develop political identities free from state surveillance and public pressures to conform to social norms.



Political Privacy and Freedom of Association

- Privacy provides a secure environment for social movements and activists to prepare for engagement with the state.



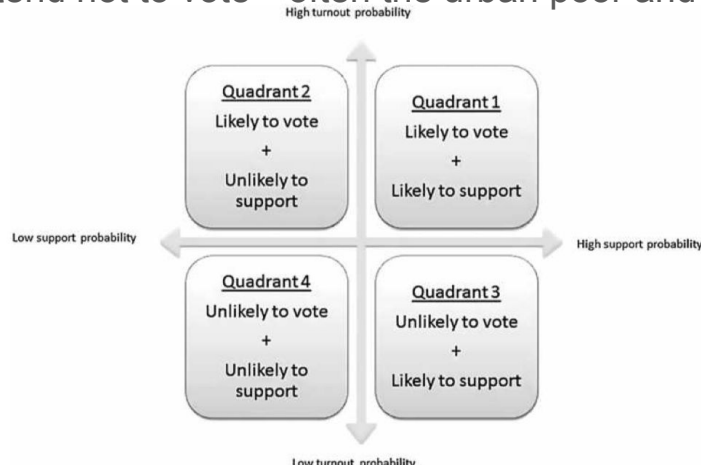
Data and the Democratic Deficit

- Electorate Segmentation
- Narrowed Political Representation
- Gerrymandering

Data and the Democratic Deficit

- Electorate Segmentation

Campaign consultants have demonstrated that they will not spend significant resources engaging citizens who tend not to vote—often the urban poor and ethnic minorities.





Data and the Democratic Deficit

- Narrowed Political Representation

Activity from Jan. 1, 2019 through June 30, 2019 (figures in millions)

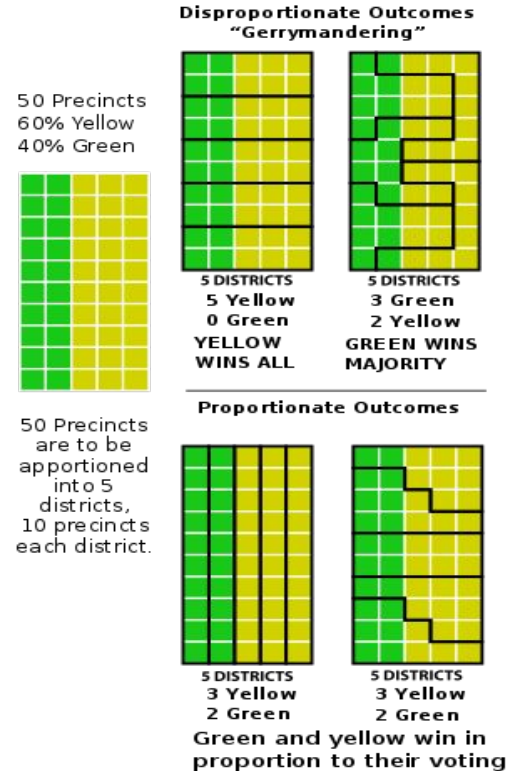
Filers	Receipts	Disbursements
2020 Presidential Candidates	\$337.1	\$159.9
2020 Congressional Candidates	\$389.0	\$172.2
Party Committees	\$353.7	\$279.9
PACs	\$958.2	\$818.7

Data and the Democratic Deficit

- **Gerrymandering**

Gerrymandering is a practice intended to establish an unfair political advantage for a particular party or group by manipulating district boundaries, which is most commonly used in first-past-the-post electoral systems.

Gerrymandering; drawing different maps for electoral districts produces different outcomes



Data and the Democratic Deficit

The Supreme Court on June 27, 2019 ruled that federal courts are powerless to hear challenges to partisan gerrymandering.

Supreme Court Bars Challenges to Partisan Gerrymandering



Assistants run outside the Supreme Court after the court rejected efforts to rein in partisan gerrymandering. Samuel Corum for The New York Times

By Adam Liptak

June 27, 2019





Conclusion

- The professionals who work behind the scenes with data to identify voters and coordinate collective action, remain largely unanalyzed.
- The lack of transparency and security in political data raises significant concerns for citizens.



Conclusion

- As a start, data practices and storehouses of information should be made transparent and accessible.
- Citizens should have the means to find out what data is collected about them and have the ability to opt out of political databases.
- Parties and candidates should be required to develop privacy statements that cover the full range of personal data.
- Political data should be subject to the breach reporting requirements that many states currently mandate of commercial providers.



Implications for computer science

- Meanwhile, much more scholarly work is needed into the democratic effects of the data-driven polity.
- **Security:** would help to ensure that databases of citizens are more accurate, secure, private, and reliable.
- **Data Mining:** What political data is collected about citizens? What is it used for? Who owns it? How is it stored



References

- Kumaraguru, Ponnurangam, and Lorrie Faith Cranor. Privacy indexes: a survey of Westin's studies. Carnegie Mellon University, School of Computer Science, Institute for Software Research International, 2005.
- Recode. “Full interview: Hillary Clinton, former U.S. Secretary of State | Code 2017.” Online video clip. Youtube, 31 May 2017. Web. 15 Apr 2020.
- Acquisti, Alessandro, Leslie K. John, and George Loewenstein. "What Is Privacy Worth?" *The Journal of Legal Studies* 42, no. 2 (2013): 249-74. Accessed April 13, 2020. doi:10.1086/671754.
- Rusch, Thomas, Ilro Lee, Kurt Hornik, Wolfgang Jank, and Achim Zeileis. "Influencing elections with statistics: Targeting voters with logistic regression trees." *The Annals of Applied Statistics* 7, no. 3 (2013): 1612-1639.
- Liptak, Adam. “Supreme Court Bars Challenges to Partisan Gerrymandering.” *The New York Times*, *The New York Times*, 27 June 2019, www.nytimes.com/2019/06/27/us/politics/supreme-court-gerrymandering.html.
- Turow, J. (2006). *Niche envy: Marketing discrimination in the digital age*. Cambridge, Mass: MIT Press.
- Alexander, K., & Mills, S. (2004). *Voter privacy in the digital age*. California Voter Foundation. Available at <http://www.calvoter.org>