

**PAST + PRESENT = FUTURE**

*A New Approach to  
Predicting Voter Turnout*

**ANDREW THERRIault, PHD**

*SENIOR DATA SCIENTIST*

*GREENBERG QUINLAN ROSNER RESEARCH*

*Presented at the  
2014 Annual  
Meeting of the  
Midwest Political  
Science  
Association*

*April 4, 2014*

*Chicago, Illinois*

# INTRODUCTION

- Academic research often focuses on modeling turnout retrospectively
- Future turnout is more relevant in applied research
- Turnout prediction is important for:
  - Likely voter models for survey screening and weighting
  - Election forecasting models
  - Persuasion and GOTV targeting in campaigns
- The dilemma: modeling an outcome that hasn't happened yet

# PROBLEMS WITH PREDICTING TURNOUT

- We can model stated turnout intention, but:
  - Responses are plagued by measurement error
  - Survey response propensity is highly correlated with turnout propensity, creating the potential for selection bias
  - Because of both these things, models based solely on stated intentions are generally not very effective
- Turnout history in recent elections generally produces more accurate results, but:
  - Doesn't work well for youth and others with weak history
  - Not good at accounting for changing patterns over time

# A HYBRID APPROACH TO MODELING TURNOUT

- Conceptually, turnout propensity in a given year can be broken down into two distinct components:
  1. General turnout propensity
  2. Election-specific motivations
- The best reflection of general turnout propensity is turnout history over a series of elections
- Election-specific factors can be seen in survey questions about turnout intent and election interest
- Use data from previous elections to find weights

# EXAMPLE: PREDICTING TURNOUT IN 2010

- Survey data from Democracy Corps National polls
  - 5 polls, RDD samples, live landline and cell interviews
  - Names and phones matched to voter records post-election
  - 2,193 matched respondents
  - Includes both turnout intention and voter interest questions
- Turnout data from 5 previous federal elections
  - 2010 primary, primaries and generals from 2008 and 2006
  - Combined through principal-components factor analysis
- Outcome is recorded turnout in 2010 general

# PREDICTING TURNOUT WITH ACTUAL SURVEY RESPONSES

	<b>All Respondents</b>				
<b>Turnout Intention</b>	2.07 (0.22)		1.78 (0.24)		0.90 (0.27)
<b>Voter Enthusiasm</b>		1.51 (0.23)	0.89 (0.25)		0.94 (0.30)
<b>Turnout History</b>				1.67 (0.08)	1.64 (0.08)
<b><i>n</i></b>	2193	2186	2186	2193	2186
<b>Correctly Predicted</b>	82.0%	81.8%	82.0%	85.7%	86.8%

*Coefficient estimates and robust standard errors from binary logit models*

# PREDICTING TURNOUT WITH MODELED VOTER ENGAGEMENT SCORES

	<b>All Respondents</b>		
<b>Engagement Score</b>	0.47		0.25
	(0.05)		(0.06)
<b>Turnout History</b>		1.67	1.63
		(0.08)	(0.08)
<b><i>n</i></b>	2193	2193	2193
<b>Correctly Predicted</b>	82.0%	85.7%	86.6%

*Coefficient estimates and robust standard errors from binary logit models*

# PREDICTIONS FOR YOUNGER VOTERS

	<b>Under 40</b>		
<b>Engagement Score</b>	0.38		0.32
	(0.10)		(0.10)
<b>Turnout History</b>		1.48	1.46
		(0.21)	(0.22)
<b><i>n</i></b>	269	269	269
<b>Correctly Predicted</b>	68.4%	71.8%	74.4%

*Coefficient estimates and robust standard errors from binary logit models*



# PREDICTIONS FOR LOW-PROPENSITY VOTERS

	<b>Below-Average Turnout History</b>		
<b>Engagement Score</b>	0.24		0.25
	(0.06)		(0.07)
<b>Turnout History</b>		1.77	1.78
		(0.16)	(0.16)
<b><i>n</i></b>	784	784	784
<b>Correctly Predicted</b>	60.0%	69.4%	72.1%

*Coefficient estimates and robust standard errors from binary logit models*

# CONCLUSION

- By combining survey responses and turnout history, we can predict turnout better than with either alone
- This hybrid approach works especially well for younger voters and those with sparse histories
- The biggest challenge is in figuring out how to combine these two types of predictions into one
- Solution: historical data
- With wider use, we can develop better surveys, forecasts, and campaigns