Issues, Media and Momentum in the 2008 Primary: Automated Retrieval of Dynamic Content

Dino P. Christenson

Department of Political Science

THE OHIO STATE UNIVERSITY

RESEARCH QUESTIONS

- How does media coverage of issues change over the primary campaign?
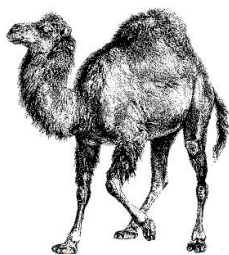- To what extent is the coverage related to key campaign events?

SUMMARY

- Explore issue salience in the 2008 presidential primary campaign
- Expose dynamic nature of the media's issue coverage as it pertains to both the invisible primary period and the early primary
- Generate smooth summary measures of the dynamics with Bayesian state space models
- Find that particular issues dominate the campaigns at the start and finish, but that there are also changes in the relative salience of issues throughout the campaign
- Showcasing data collection and potential use
- Just the first few of several steps toward understanding the dynamic relationship of media issue coverage to indicators of candidate status in the 2008 primary

LARGER PROJECT

THE VISIBLE PRIMARY 2008

- Original longitudinal dataset culled from newspaper articles
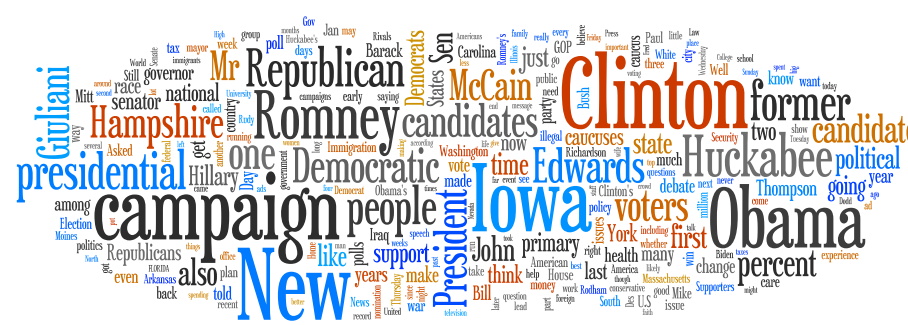- Use Perl to automatically retrieve and code articles
- Sample national and local newspapers
- Test the dynamic mechanisms of primaries
- Evaluate roles of media, polls, contributions and candidate visits
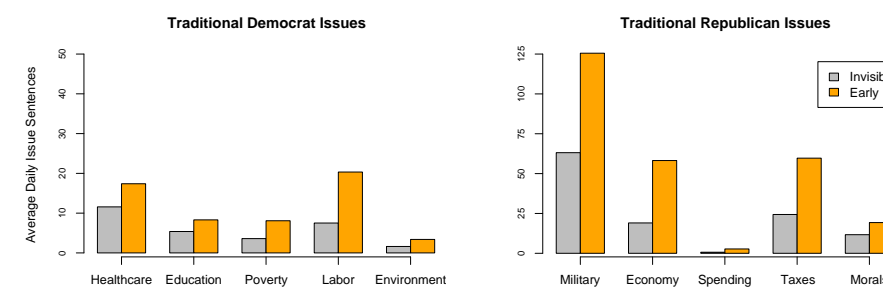- Project with Corwin D. Smidt: http://VisiblePrimary.com

DYNAMIC ISSUE CONTENT

- Counting number of sentences per day with a salient issue in national newspapers
- Timeline: July 9, 2007 to February 5, 2008
- 4,184 articles mentioning an issue
- 151,702 issue sentences

DATA WORD CLOUD

RAW ISSUES DATA

MOTIVATING KALMAN FILTER

Dynamic process
- Daily censoring at zero
- Sampling error from large daily sample sizes
- Large volatility in issue mentions
- Days of week and holiday media reporting
- FEC reporting deadlines

BAYESIAN STATE SPACE MODEL

State equation:

$$Issues_t = \theta_t + Day_t + FEC_t + \epsilon_t \ \ where \ \epsilon \sim N(0, \sigma_t^2)$$

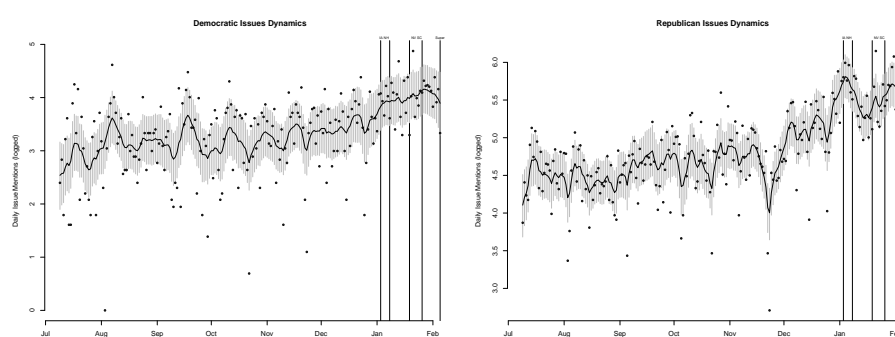Transition equation:

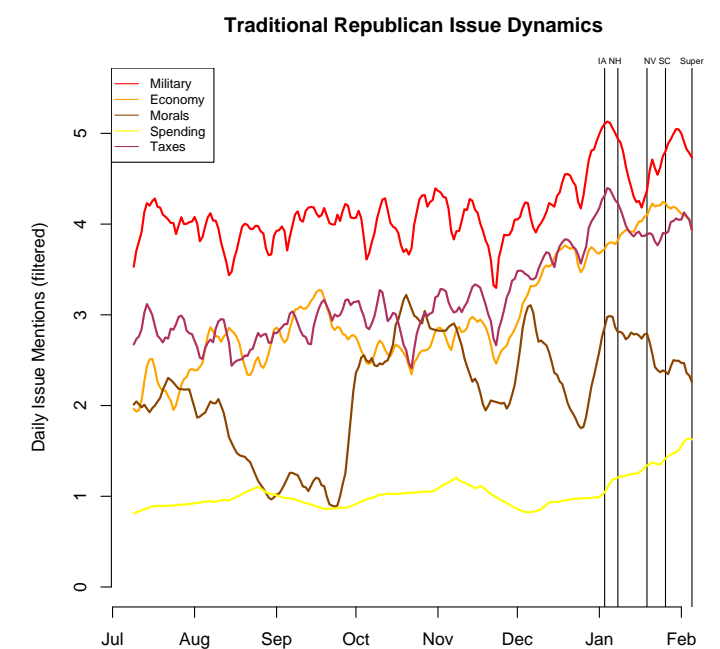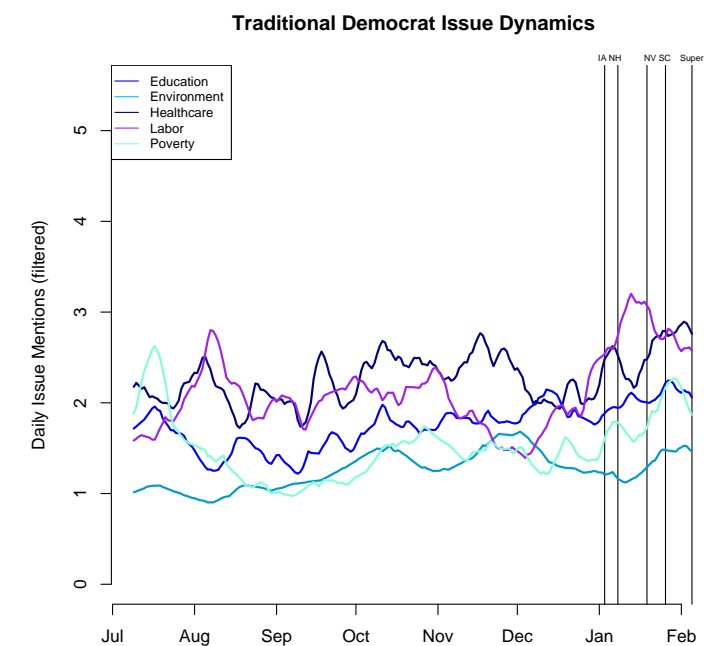$$\theta_t = \theta_{t-1} + \upsilon_t \ \ where \ \upsilon \sim N(0, \sigma_\upsilon^2)$$

Bayesian State Space Model:
- Forward filtering backward sampling algorithm
- Inverse-Wishart priors for residuals
- Normal diffuse priors for regressors

AGGREGATE FILTERED SERIES

FILTERED ISSUES

CONCLUSION

- Both horse-race & issue coverage increase over the 2008 Primary
- Defense dominates traditional Rep issues
- Healthcare strongest among Dem issues
- Economy important but not the leading concern
- Moral values & spending with relatively little attention

NEXT STEPS

- Further filter the series by taking into account candidate visits & other cycles
- Extend timeline to include period leading up to conventions
- Explicit hypothesis tests of role of issue coverage (dynamic ownership & status)
- Address theories of bottom-up vs top-down issue salience in campaigns
- Multivariate perspectives: BVAR