Documentation for Momentum and Social Learning in Presidential Primaries

Brian Knight[*] Nathan Schiff[†]

December 2010

1 Introduction

This note documents the empirical analysis for our paper. We first describe the methods used to construct the tables and figures. We then list the specific programs and corresponding datasets that can be used to replicate our results.

2 Construction of Tables

2.1 Table 1: Estimation of State Preferences

In table 1 we estimate the state preferences η_{cs} for each candidate by running a multinomial logit specification where the dependent variable is the candidate for whom the survey respondent intended to vote and the independent variables are binary variables for each state. These preferences capture voting intentions before any elections have taken place and thus the sample period is the 103 day interval between October 7, 2003 and January 19, 2004 ($t = 1$). In our model the state preferences are distributed $\eta_{cs} \sim N(0, \sigma_\eta^2)$ therefore we demean the coefficients

[*]Department of Economics, Brown University, Providence RI 02912 *Brian_Knight@brown.edu*

[†]Sauder School of Business, University of British Columbia, Vancouver BC V6T 1Z2 *Nathan.Schiff@sauder.ubc.ca*

on the state fixed effects from our multinomial logit, estimated without a constant, to get the η_{cs}:

$$\eta_{cs} = \beta_{cs} - \bar{\beta}_c \tag{1}$$

The average of these coefficients ($\bar{\beta}_c$) is the initial prior for the candidate, μ_c. We normalize the utility from John Kerry (the base case) to 0 so that all η_{cs} are interpreted with respect to Kerry, for whom all η_{cs} equal zero.

In order to calculate standard errors for our overall estimation, which is a two-step procedure, we have to bootstrap the results in this first step. We do that by estimating results for 100 bootsamples, which are samples randomly drawn from our original sample with replacement. If the randomly drawn sample did not include at least one voter in favor of each candidate, making estimation of state preferences impossible, then all observations for that state were dropped. The reported point estimates for η_{cs} are the median values from estimating the parameters on 100 bootsamples and the confidence intervals represent the 2nd lowest and 98th highest value.

In our base case we simply regress voting intentions against state fixed effects. Letting y_i be the candidate choice of voter i and s_c be a vector of state fixed effects *for each candidate* relative to Kerry we estimate:

$$Pr(y_i = c) = \frac{exp(\beta_c s_c)}{\sum_{j=1}^{C} exp(\beta_j s_j)} \tag{2}$$

In our distance specification we relax the constraint that voters do not know the preferences in other states and allow voters to partially infer other states' preferences by the distance of that state to the candidate's home state. To do this we first estimate the specification in 2 and then run a Seemingly Unrelated Regression of this η_{cs} on the relative distance of the state to the candidate's home state. This relative distance is defined as $rd_{cs} = \frac{d_{sc}}{d_{sk}}$, which is the distance from the voting state to the candidate's home state divided by the distance from the voting state to John Kerry's home state (MA). We constrain the coefficient on this relative distance to be the same for all candidates, meaning distance affects preferences in the same way for all candidates.

$$\eta_{cs} = \alpha_c + \beta rd_{cs} + \varepsilon_{cs}, c = 1, ..., C \tag{3}$$

The values for η_{cs} that we report in Table 1 are the residuals from this regression, or the ε_{cs} in the above equation.

Finally, our third specification relaxes the assumption that the average of voter preferences across all states would not have changed in the absence of social learning. In this specification we allow for candidate specific national trends by estimating a time trend over the first 103 days in period 1 in the multinomial logit model:

$$Pr(y_i = c) = \frac{exp(\beta_c s_c + \gamma t_c)}{\sum_{j=1}^{C} exp(\beta_j s_j + \gamma t_j)} \tag{4}$$

2.2 Table 2: Estimation of Model Parameters

In Table 2 we report the key parameters of the model for our base case, distance case, and trend case: the variance of state preferences (σ_η^2), the variance of the initial prior (σ_1^2), and the variance of candidate quality (σ_ε^2). The σ_η^2 is estimated in the first-stage multinomial regressions described earlier and is simply the variance of the estimated etas[1]. The remaining parameters, σ_1^2 and σ_ε^2, are estimated using maximum likelihood as outlined in the paper. We implement the distance case by simply subtracting out the predicted state preference from the surprise, which is the difference between the observed vote share for a state and the current mean quality estimate:

$$\ln(v_{cst}/v_{0st}) - \gamma X_{cs} - \mu_{ct} = \ln(v_{cst}/v_{0st}) - \alpha_c - \beta r d_{cs} - \mu_{ct} \tag{5}$$

The trend specification is implemented in a similar way by subtracting out the trend from this surprise term

$$\ln(v_{cst}/v_{0st}) - \gamma X_{cs} - \mu_{ct} = \ln(v_{cst}/v_{0st}) - \gamma t_c - \mu_{ct} \tag{6}$$

However, an important difference between the distance implementation and the trend implementation is the method of estimation. In the distance implementation we first estimate state preferences with the standard multinomial logit of the base case and then use a second regression to estimate the effect of distance on these preferences while in the trend specification we put a trend term directly into the multinomial logit.

2.3 Table 3: Alternative Quality Regressions

In Table 3a we look at the assumption of sincere voting by investigating how other measures of perceived candidate quality—survey questions on favorability—change

[1] As before, this is calculated with bootstrapping and we report the median value from 100 estimations.

over time. To do this we pool the perceived quality measures over Dean and Edwards (again, everything is relative to Kerry) so that the unit of observation is the respondent-candidate and each respondent has two observations for an opinion on each candidate. We then regress each alternative measure of quality (favorability, trustworthy, knowledgeable, etc...) on the estimated mean quality level per period (μ_{ct}), allowing for candidate specific constants. If voters are strategic and our estimated mean quality level (μ_{ct}) simply increases due to electability, viability, or bandwagon concerns then there should be no relationship between these alternative measures of quality and μ_{ct}. In the following regression repeated from the paper, we test whether $\kappa = 0$:

$$\text{quality}_{itc} = \delta_c + \kappa \times \mu_{ct} + \xi_{itc} \tag{7}$$

In Table 3b we use a different sample from the NAES to see whether respondents who have already voted answered three questions about candidate characteristics more accurately than those had not yet voted. Here we use a linear probability model where the dependent variable is an indicator variable for whether a given question is answered correctly and the independent variable is whether the respondent had voted. We also include date and state fixed effects, η_c and η_t. The specification is for each of the three questions is:

$$\text{AnswerCorrect}_{itc} = \beta \text{Voted} + \eta_c + \eta_t + \xi_{itc} \tag{8}$$

2.4 Table 4: Counterfactual Primaries

In Table 4 we report the results of a counterfactual primary in which all states vote simultaneously for two scenarios: one with all three candidates (Dean, Edwards, Kerry) and one in which only Edwards and Kerry run. We run both scenarios because Dean dropped out in period 8 of our sample (2/18/2004), making it impossible to estimate his vote share in states voting after this period since we do not observe the private signal about Dean's quality that they would have received. Similarly, because we back out the signal received by using the vote share and the estimated state preference, we can only run simulations on states for which we had enough observations to estimate the state preferences. In the primaries we covered in our sample period there were 30 states but we can only estimate preferences for 25 of them. We exclude the following states: ND (t=3), DC (t=7), HI and ID (t=9), and VT (t=10). In the "Sequential Primary" columns of Table 4 we list these 25 states and the actual vote share received by each candidate. In the

”Simultaneous” columns we estimate the vote share had every state voted at the same time as Iowa using the following equation from the paper:

$$\ln(v_{cs1}/v_{0s1}) = \eta_{cs} + \alpha_1\theta_{cs} + (1-\alpha_1)\mu_{c1} \tag{9}$$

We back out the θ_{cs} for each state using the fact that $\theta_{cs} = \frac{\ln(v_{cs}/v_{0s}) - \eta_{cs} - (1-\alpha_t)\mu_{ct}}{\alpha_t}$ and our estimated values for η_{cs}, α_t, and μ_{ct}. To get the vote share from the log of the vote share ratio we exponentiate the log vote shares using the normal logit equation where the denominator is all three candidates for ”Simultaneous Primary (3 way)” and just Edwards and Kerry for ”Simultaneous (2 way).” The logit equations for Edwards (candidate 2) in the three-way and two-way simulations are given below:

$$v_{2s}^{3way} = \frac{\exp(ln(v_{2s}/v_{0s}))}{1+\exp(ln(v_{1s}/v_{0s}))+\exp(ln(v_{2s}/v_{0s}))} \tag{10}$$

$$v_{2s}^{2way} = \frac{\exp(ln(v_{2s}/v_{0s}))}{1+\exp(ln(v_{2s}/v_{0s}))} \tag{11}$$

The delegates for this table are simply assigned by vote share so that if state s has dl_s delegates than candidate c gets $v_{cs} * dl_s$ delegates.

2.5 Table 5

In Table 5 we run a different set of simulations evaluating the scenario in which states vote in a different order than observed but the same number of states still vote in each period. To do this we randomly assign states to each period (again, keeping N_t as before) and then recalculate every state’s voting results using our previously estimated private state signals (θ_{cs}) and weights on the private (α_t) and public (β_t) signals. Again, this is only done for the 25 states for which we can estimate state preferences.

3 Construction of Figures

3.1 Brief Notes on Figures 1-9

Note: all figures use the base specification (not the distance or trend specifications).

- Figure 1: We report the number of Democratic primaries on each date in 2004 and 2008, including D.C., using data from cnn.com.
- Figure 2-4: We plot the single day and two day average intended vote share from the NAES survey for voters in all states and then also plot the observed vote share in the Iowa primary.
- Figure 5: In panel *a* we plot the estimated μ_t for Dean and Edwards against time and where μ_t for Kerry is always zero. In panel *b* we plot the estimated σ_t^2 against time.
- Figure 6: We plot the estimated weight on the private (α_t) and public (β_t) signals against time, along with the ratio $\frac{\beta_t}{\alpha_t}$.
- Figure 7: all details in paper and code (”unknown_break_struct_FINAL.do”)
- Figure 8: see ”campaign_contributions.do”
- Figure 9: see ”create_figures_JPE.do” and ”Advertising Data” section

3.2 Figure 10 and Calculation of Implied Voting Weights

In Figure 10a we quantify the voting weights of states in each position in the sequential primary. As before, we first restrict the sample to the 25 states for which we have η_{cs} estimates since we will be simulating voting results under various changes to η_{cs} and θ_{cs}. We then calculate the vote share for every state based upon our estimated parameters (σ_ε^2, σ_η^2, and σ_1^2) but restricting to the 25 state sample and changing the number of states (N_t) accordingly. This can be interpreted as asking what would be the outcome in the primaries if the parameters were the same as before but only 25 states voted, the average vote share in each period was still the same as for the larger sample, but the weight on the signal reflected the number of states in the smaller sample. We use this scenario as our base case and then ask what would be the effect on total vote share for John Edwards if one state in a given period had a larger preference for Edwards. There are two effects of shocking a state's preference on total vote share (across all 25 states) : 1) vote share increases because a larger percentage of voters in that state vote for the candidate (”direct effect”) and 2) vote shares increase because the voters in subsequent periods update favorably over the candidate's quality (”indirect effect”). To calculate these effects we use the following procedure:

1. Calculate the vote share for Edwards using the values of the parameters calculated in the base model but restricting to the 25 states for which we have eta values (”base case”).

2. Shock each state in a period and calculate the change (compared to the base case) to total vote share for Edwards, allowing social learning (subsequent μ values change). This yields 25 values–the ”total effect” for each shocked state.

3. Repeat previous step but eliminate social learning (subsequent μ values do not change). This yields 25 more values–the ”direct effect” for each shocked state.

4. Calculate average total effect for a period by averaging total effect for each state in the period.

5. Calculate average direct effect for a period by averaging direct effect for each state in the period.

6. Calculate average indirect effect for a period by subtracting average direct effect from average total effect.

Finally, to turn these calculate effects into weights we normalize by the period 10 values. For example, we divide the average total effect for each period by the average total effect for period 10. Thus the values for periods 1-9 are in relation to the period 10 value and the value for period 10 is 1.

In Figure 10b we plot the average advertising coverage (see ”Advertising Data” section of this note) of each state voting in a period (again restricting to the 25 states).

4 Data Files and Programs

4.1 Advertising Data

The advertising data comes from the Wisconsin Advertising Project where the unit of observation is the Designated Market Area (DMA), a geographic area comprised of contiguous counties considered to be in the same media market. There are 100 DMAs in the Wisconsin data, covering a large percentage of the urbanized area in the US. We matched county populations to DMAs to find the population size of each DMA using data from the 2004 Census population estimates. We wanted a measure of the number of minutes that a representative viewer in a given state is exposed to. If coverage in the Wisconsin Advertising data were comprehensive, total advertising coverage across all residents of state s could be approximated by: $A_s = \sum_m N_{sm} A_m$ where m indexes media markets, N_{sm} represents the number of individuals from state s living in market m, and A_m represents total advertising in market m. Note that N_{sm} will often be zero (e.g., there are no residents of Maine in the Los Angeles media market). Coverage received by an average viewer in state s would be given by $a_s = A_s/N_s$ where N_s is the population of state s. Thus, we have that:

$$a_s = \sum_m \frac{N_{sm}}{N_s} A_m \tag{12}$$

where $\frac{N_{sm}}{N_s}$ can be interpreted as the weights placed on the different markets within state s.

Given that the Wisconsin Advertising data do not cover every market, we adjust our measures as follows:

$$\begin{aligned} A_s &= \frac{N_s}{N_s^{wisc}} \sum_m N_{sm} A_m \\ a_s &= \sum_m \frac{N_{sm}}{N_s^{wisc}} A_m \end{aligned}$$

where N_s^{wisc} is the number of residents from state s who live in the 100 markets covered in the Wisconsin Advertising data.

4.2 Included Data and Programs

We have included all the programs used to do the analysis in our paper. Several of the datasets we used must be purchased and so we have included only the final data used to run our analysis, and not the intermediate datasets[2]. However, we have included all intermediate programs used to build the final data from the raw datasets. We have also not included all 100 bootsamples due to their large size but have included the code to generate these. The next page lists describes all the Stata DO files and how to use them to replicate our results.

[2]The Wisconsin Advertising Project (http://wiscadproject.wisc.edu/), The National Annenberg Election Survey (http://www.annenbergpublicpolicycenter.org/Default.aspx), Nielsen Media Research defines the counties in the DMAs (http://www.nielsen.com/us/en/industries/media-entertainment/television.html)

#	Do files	Input	Output	Description
1	**adver_analysis**	1) adver_by_state.dta 2) equityEta.dta 3) equityTheta.dta	1) adver_analysis.dta	Merges advertising data with implied voting weights. Runs regression of advertising coverage on both measures of voting weights, creates figures showing average advertising coverage against voting weights for each period.
2	**adver_by_state.do**	1) wisc_dma_to_dma_codes.dta 2) county_pop_estimates_2004.dta 3) demAds_subsets--all 4 versions **(not included)** 4) dma.dta **(not included)**	1) adver_by_state.dta	Calculates advertising by state and coverage
3	**campaign_contributions.do**	1) contributions_dean.txt 2) contributions_kerry.txt 3) contributions_edwards.txt	1) contributions.dta	Merges campaign contributions data for creation of figure 9 by create_figures_JPE.do
4	**clean_wisc_adver.do**	1) wisc_dma_to_dma_codes.dta 2) wiscads_2004_presidential.dta (**not included**)	1) demAds_subset.dta 2) demAds_subset_bIowa.dta 3) demAds_subset_allDems.dta 4) demAds_threeDems_allDates.dta	Takes original Wisconsin advertising file and creates subsets
5	**counter**	1) after1_nodist.dta, complete_data1	1) counter1.dta 2) counter2.dta 3) after3.dta 4) delegates.dta	Runs counterfactual simultaneous election simulation for **Table 4**
6	**create_complete_data1**	1) 2004_dem_aggregate_cnn2v9.dta 2) state_dist_miles_wide.dta	1) complete_data1 2) complete_data2	Creates dataset of aggregate voting returns and distance by state (complete_data1) and one obs (complete_data2)
7	**create_complete_data2**	1) complete_data2 2) data2004nationalrolling4 3) data2004nhcross3	1) complete_data 2) bootsample 3) bootsample1-bootsample100	Creates individual dataset, merges with aggregate returns, creates bootstrap samples
8	**create_figures_JPE.do**	1) many (all included)		Creates all 10 figures for paper
9	**equity_eta_allStates**	1) after3.dta	1) equityEta_allstates.dta 2) equityEta_weights.dta	Used to calculate implied voting weights
10	**favorability**	1) after2_nodist.dta 2) complete_data		Merges results with favorability data and runs favorabilty regressions for **Table 3a**.
11	**figures_after_nodist**	1) bresults_nodist.dta	1) after1_nodist.dta 2) after2_nodist.dta	Calculates weights and creates files to be used by create_figures_JPE
12	**Iowa_graphs**	1) complete_data1.dta	1) iowa_data.dta	Formats complete_data1 to create figures 2-4
13	**knowledge2.do**	1) data2004nationalrolling2.dta (**not included**) 2) data2004nhcross.dta (**not included**)	1) table3b.dta	Merges NH data and rolling data to create large sample of voters who answered factual questions, including many voters who had already voted at time question was asked. Runs regressions for **Table 3b**.
14	**MNL_bstrap1_dist**	1) bootsample1-bootsample100 (**not included**)	1) bresults_dist.dta	1) Allows observers to expect different preferences from previously voting states based on distance to candidate's home state 2) calls MNL_bstrap2_dist, runs regressions, creates set of results
15	**MNL_bstrap1_nodist**	1) bootsample1-bootsample100 (**not included**)	1) bresults_nodist.dta	calls MNL_bstrap2_nodist, runs regressions, creates set of results
16	**MNL_bstrap1_trend**	1) bootsample1-bootsample100 (**not included**)	1) bresults_trend.dta	1) Allows opinions of candidates to follow trend observed from beginning to right before IA vote 2) calls MNL_bstrap2_trend, runs regressions, creates set of results
17	**MNL_bstrap2_dist**	1) bootsample1-bootsample100 (**not included**)		Program to run model for distance spec
18	**MNL_bstrap2_nodist**	1) bootsample1-bootsample100 (**not included**)		Program to run model for base spec
19	**MNL_bstrap2_trend**	1) bootsample1-bootsample100 (**not included**)		Program to run model for trend spec
20	**random_schedules_2way**	1) after3.dta 2) delegates.dta	1) rand2way.dta	Does counterfactual election simulation with random order--two way election between Edwards and Kerry for **Table 5**
21	**random_schedules_3way**	1) after3.dta 2) delegates.dta	1) rand3way.dta	Does counterfactual election simulation with random order--all three candidates for **Table 5**
22	**tables_1_2**	1) bresults_nodist.dta 2) bresults_dist.dta 3) bresults_trend.dta		This program collapses the results from bresults files for formatting (done in Excel) to create the exhibits for the paper
23	**unknown_break_struct_FINAL**	1) bootsample	1) structural_break_logit.dta	Runs structural break analysis