ERASMUS UNIVERSITY ROTTERDAM

ERASMUS SCHOOL OF ECONOMICS

MASTER THESIS

November 2013

SOCIAL LEARNING AND MOMENTUM EFFECTS

CASE STUDY: 2008 US DEMOCRATIC PRIMARY

Abstract

This paper examines whether there are momentum effects and social learning in the 2008 US Democratic primary. In the econometric model, voters are assumed to engage in sincere voting, but they are uncertain about candidate quality. Thus, voters in late states attempt to benefit from sequential voting by updating their voting intentions with information on voting returns from early states. The empirical application uses voting intentions of late voters and through a two-step estimation process shows whether these voting intentions are updated with the release of voting returns from early states. Our findings indicate that there is no social learning in the 2008 Democratic primary. Although Obama and Edwards outperformed in Iowa, voters did not update appropriately and Obama did not enjoy any momentum effect from this surprising victory relative to expectations, with Clinton winning in New Hampshire. Finally, in an additional empirical specification the paper examines the sources of voting behavior and the results show that demographic characteristics such as gender, age, income, education and region of residence have a significant effect on the probability to choose a candidate.

Student: Despoina Sakoglou

Student Number: 371199

Supervisor: Dr. Dana Sisak

1. INTRODUCTION

Election systems form a very interesting subject of political science, mainly because they vary across different election cases. While in many cases voters can choose simultaneously among several alternatives, in other cases elections take place sequentially, with some voters choosing earlier than others. A well known example of sequential voting is the presidential primary in the United States, during which the Democratic and the Republican Party sequentially elect their nominee for the general election. The primary season starts in the first months of the year that the general election will take place, usually in January, and continues for several months until all states have held their primary, usually until the beginning of June.

An interesting fact governing the presidential primary is that states want to vote as early in the primary season as possible, while the National Parties have set some rules prohibiting some states to vote earlier than a specific date. A widely discussed incident proving this fact is the penalty that the Democratic National Party gave to Michigan and Florida in the 2008 primary, because they violated the party rules by holding their primaries earlier than allowed. Also, taking a look at the primary schedule across election years it is obvious that primary dates change and in many cases move earlier than they were previously. For instance, while the 2004 primary season started with the Iowa caucus taking place on January 19, in 2008 the Iowa caucus moved earlier to be held on January 3, with all consequent primaries also being held earlier compared to 2004.

Many researchers have been intrigued by these facts, because they raise questions regarding the voting order of states and the potential effects this order could have on outcomes. There are several papers in existing literature that support that sequential voting provides late voters with the opportunity to learn from early results and update their preferences, thus causing momentum effects enjoyed by candidates. A very interesting analysis of momentum and social learning in presidential primaries is presented by Knight and Schiff (2010). In their study, the authors provide a theoretical framework for social learning and develop a discrete choice econometric model to test for social learning in the 2004 Democratic primary. Their findings suggest that candidates experience momentum effects when they perform better than expected in early states; indeed in the 2004 primary candidate Kerry enjoyed momentum effects because he outperformed in early states relative to expectations.

Motivated by this study, we decided to use the econometric model Knight and Schiff (2010) developed in order to examine empirically whether the momentum effects in the 2004 primary are persistent and appear in the 2008 primary as well. We consequently need to adopt the theoretical framework the two researchers proposed for the social learning analysis, in order to set the basis for the model. The main goal of our study is to test whether there is social learning and momentum in the 2008 Democratic primary and the results of this analysis will be the main contribution of our paper to existing literature.

The econometric model uses data from daily opinion polls on a rolling cross section design, focusing on the 2008 Democratic primary. The estimation process results in four key social learning parameters that indicate how the voting intentions of late voters evolve after the release of vote shares from early states. Social learning occurs when late voters learn from the release of such information and update their voting intentions appropriately. Our estimates suggest that there is no social learning in the 2008 primary.

Social learning attempts to explain how voting behavior potentially changes after the release of information regarding voting returns from early states. However, it does not provide any information about the possible sources of voting behavior, or in other words those aspects that initially form voting behavior before it is influenced by the potential effects of sequential voting. For that reason, we decided to investigate empirically some potential sources of voting behavior, related to demographic characteristics. The results of this additional specification will be the second contribution of our paper to existing literature.

We develop a discrete choice econometric model that uses daily opinion polls on a rolling cross section design, focusing on the pre-primary season and we test the potential effect of some demographic characteristics on these voting intentions. Our findings indicate that age, gender, some levels of income, some levels of education and region of residence have significant effects on the probability to choose a specific candidate, providing empirical evidence to existing theory that these characteristics can be considered as sources of voting behavior.

The paper is structured as follows. Part 2 gives an overview of the theoretical and empirical works that are related to this study. Part 3 provides an extensive analysis of the theoretical framework the study adopts, originally developed by Knight and Schiff (2010). Part 4 describes the empirical application that derives from the theoretical framework and presents the main findings of our study. Part 5 develops our additional empirical specification regarding the sources of voting behavior and finally Part 6 gives a summary of our key results along with some concluding remarks.

2. LITERATURE REVIEW

There has been a wide range of research papers focusing on the social learning process and looking for potential momentum effects in a sequential game, which in turn can be applied to the presidential nomination system. Many researchers developed theoretical models in an attempt to explain whether social learning occurs and how it affects individual agents' behavior (in this context voter behavior), while others tested these theories empirically, providing evidence for the validity of these theories based on real cases.

The first theoretical works related to social learning belong to Welch (1992) and Bikhchandani et al. (1992), who develop a theoretical analysis of informational cascades as a sort of social learning process. The findings of these works show that under a sequential choice setting with uncertainty about the state of the world, cascades will always arise and, once started, they will potentially last forever. Nevertheless, this fallacy makes them fragile, in the sense that they can easily be broken as soon as even very little information becomes public. Banerjee (1992) follows a similar theoretical structure to prove that the decision rules that individuals choose to follow are characterized by herd behavior- people doing what others do instead of using their private information. Smith and Sorensen (2000) are based on these works to extend the analysis by including a general signal space and allowing for different preferences over voters' actions.

Ali and Kartik (2006, 2010) extend their analysis from the standard theories of informational cascades by introducing the notion of voting as a collective decision problem and construct an equilibrium that generates rational herding in sequential elections. In this framework, voters not only look at voting history, but also have forward-looking incentives to consider the actions of the following voters. Their analysis suggests that momentum can arise from this desire of the voter to be pivotal- to choose the right candidate. In a similar way, Dekel and Piccione (2000) propose that strategic voters condition their actions on being pivotal and they go further to show that symmetric equilibria in simultaneous voting games are also equilibria in sequential games.

Apart from the aforementioned social learning literature, there are also some papers investigating alternative reasons for momentum effects, or even potential counterfactual results to conventional findings. Klumpp and Polborn (2005) suggest a model that includes campaign spending and they propose that campaigning is more intensive in the first stage, which creates an asymmetry in the candidates' incentives to continue campaigning in later stages, thus increasing the probability of the first stage winner to win in later stages. Selman (2010) focuses on optimal sequencing of the presidential primaries and, in contrast to the results of Banenjee (1992) and Bikhchandani et al. (1992) his findings suggest that there can be beneficial informational cascades. He proposes that a party can benefit from voter herding when one candidate has more expected loyal support than the other, despite the loss of valuable information caused by herding. Strumpf (2002) also studies a sequential election contest and proposes that potentially there can be an opposing force to momentum. If a candidate is expected to win several elections at the end of the contest, he will have an incentive to stay in the race, despite possible negative results of early elections. Thus, he can benefit from late elections, in contrast to momentum theory favoring early winners.

There is an increasing empirical literature that attempts to examine whether these theoretical models, developed in the previously mentioned papers, are valid and still hold in real cases. Several empirical works are based on the social learning framework proposed by Welch (1992), Bikhchandani et. al. (1992) and Banerjee (1992). Specifically, Cai et al. (2008) conduct a randomized natural field experiment in order to test for social learning effects. They use a restaurant dining setting and they find that when customers are presented with rankings of the top selling dishes, they tend to order those dishes more frequently, especially if they belong to infrequent customers. Glaeser and Sacerdote (2007) also use the social learning framework to explain aggregation reversals, which occur when a relationship at the individual level is reversed at the group level. Using data from the National Annenberg Election Survey 2000, they examine the relationship between income and Republicanism as one of their aggregation reversal examples. These papers provide empirical support for the theoretical models of information cascades and herding, on which our theoretical model is based.

Other empirical papers mainly focus on the extensions of the basic social learning framework, such as the ones that introduce the notion of the pivotal voter, proposed by Ali and Kartik (2006, 2010) and Dekel and Piccione (2000). In this framework, Goeree et al. (2012) conduct some experiments to test the pivotal voter model, by focusing on the link between voters' beliefs and participation decisions. They find that voting propensity increases when voters predict that their preferred alternative has an advantage. Another empirical paper related to Dekel and Piccione (2000) is the one by Deltas et al. (2010). They develop a model that allows them to analyze the advantages and disadvantages of the US presidential primary system, in a setting where candidates have different policy positions and qualities and there is uncertainty about the state of the world. In this setting, a sequential voting system minimizes vote-splitting in late stages, as late voters update their information about the state of the world; however, they might be misled to choose the wrong candidate. Using the 2008 Democratic presidential primaries they find that the current sequential system is preferable to a simultaneous one, as there is a substantial probability that the wrong candidate drop out in the early stages. These papers are based on the assumption that the voter is pivotal; however, our framework assumes sincere voting, which means that early voters do not account for how their vote will affect the collective decision. Nevertheless, when Knight and Schiff (2010) proposed this framework that we adopted in the current study, they showed that even with strategic voting the results are robust, thus our framework is valid.

Regarding empirical applications on momentum effects, Bartels (1987, 1988) examines the dynamics of candidate choice in the 1984 Democratic primary. He focuses on how prospective voters respond to several campaign events such as media coverage and primary outcomes, by using data taken from the NES 1984 rolling cross-section survey. Our empirical application similarly uses a rolling cross-section survey to obtain the data. As Bartels (1987) suggests, these data provide a continuous monitoring of voter reactions to campaign events during the primary season. In his work he shows that thermometer ratings of candidates alone cannot explain candidate preferences, as they might be related to other preference sources. To avoid this omitted variable bias, he develops a model that includes two measures of candidate viability interacting with each other (predispositions based on demographic and political characteristics and perceptions based on temporal variations in primary outcomes). The results of this analysis are used to address questions about the role of momentum in the presidential nominating process. In contrast to this model that is tied to the 1984 campaign, the model used in this paper and developed by Knight and Schiff (2010) is an improvement in the sense that it directly examines momentum effects and it can be adjusted to different electoral years.

In a related research to the one presented here, Adkins and Dowdle (2001) examine the effects of voting returns from Iowa and New Hampshire on the overall primary outcomes and they find that New Hampshire plays a key role in determining the ordinal ranking of candidate finishes; however, the role of momentum is not straightforward, and should be interpreted with caution, as it is only confirmed by New Hampshire and not by Iowa returns. Similarly, Steger et al. (2004) find that New Hampshire has an important effect on the final outcome of the primaries.

Although our study is based on the previously mentioned theoretical models and can be associated with previous empirical applications, it is most closely related to the research of Knight and Schiff (2007, 2010). These researchers examine social learning and momentum effects in the 2004 Democratic primaries using daily polling data. They are the first researchers to develop a discrete choice econometric model of voting and social learning. Their model assumes uncertainty over candidate quality, where voters update their information from voting returns in early states. Their findings show that candidates can benefit from momentum effects when they perform better than expected in early states. Our analysis adopts this econometric model in order to examine whether these momentum effects can be confirmed in the 2008 Democratic primaries.

3. THEORETICAL FRAMEWORK

In this part we provide the theoretical framework, developed by Knight and Schiff (2010), which forms the basis for the econometric model we adopt as the baseline model of our study. Our main research question is whether the momentum effects found by Knight and Schiff (2010) in the 2004 presidential primaries are persistent and occur in the 2008 presidential primaries. As our main aim is to test the same research question for a different elections year, we use the same setup and the same assumptions in the model. At this point it is indicative to mention that all the notations and formulas we use in this section follow the ones proposed in the theoretical framework of Knight and Schiff (2010).

In order to develop a model for measuring momentum effects in sequential elections, we need to understand that we face a problem of a discrete nature, as voters have to choose among a specific amount of candidates. Econometric literature

(Verbeek, 2012) suggests that in order to put some structure on the different probabilities, we have to use a random utility framework, in which the utility of each alternative (in our case of each candidate) is a linear function of observed characteristics and an additive error term. According to this framework, Knight and Schiff (2010) develop the setup of the model, which is explained as follows.

Suppose that there is a choice between C candidates, indexed $c=0,1,\dots,C$ in an arbitrary order, and a set of states s voting in a sequential election, where the voting order is taken as given. Given that there are primaries taking place in more than one state on the same date, this setup provides the option for a set of states Ω_t, with a size $N_t \geq 1$ to vote on date t.

Also, assume that the utility level that voter i residing in state s receives from candidate c winning the election is given by:

$$u_{cis} = q_c + \eta_{cs} + \nu_{cis} \quad (1)$$

where q_c is the quality of candidate c, valued equally by all voters and considered a positive characteristic, η_{cs} is a state-specific preference for candidate c, and ν_{cis} is an unobservable error term, which in our context represents an individual preference for candidate c and is assumed to be independent across states and candidates with a log Weibull distribution (type I extreme value distribution). With this distribution, it is common to normalize one of the deterministic utility levels to be zero (Verbeek, 2012), which in our case will be the utility from the baseline candidate 0 ($u_{0is}=0$ for all voters).

State preferences are time independent, but there is uncertainty because we want to test for social learning and this might raise some expectations during the primary season. State level preferences are also assumed to be mutually independent, in order to allow for checking how a particular state preference affects the probability of a candidate to win[1]. For that reason, voters only know the preference of their own state without observing the other state preferences and only knowing that they are normally and independently distributed [$\eta_{cs} \sim N(0,\sigma_\eta^2)$]. Moreover, voters are assumed to be Bayesian, in order to capture the possibility of obtaining private information[2], which in this case is a noisy signal over candidate quality, explained below. Uncertainty is assumed for candidate quality as well, because of potential social learning. Initial priors (t=1) over candidate quality are normally distributed, with mean μ_{c1} and variance σ_1^2, an assumption that makes sense in our case as our sample size is quite

[1] This assumption is crucial for the econometric model, in order to avoid multicollinearity among the state level preferences. Later we explain why this assumption is important in the theoretical context of our model.
[2] According to Belleflame and Peitz (2010), pp. 684.

large[3]. Consequently, posterior probabilities will also be normally distributed, following the prior distribution[4].

With regard to the noisy signal over candidate quality that voters receive before going to the polls as their private information, it is assumed to be a linear function of candidate quality and a noise in the signal as follows:

$$\theta_{cs} = q_c + \varepsilon_{cs} \tag{2}$$

where θ_{cs} is the state-specific signal[5], common for voters residing in the same state and unobserved by voters of other states[6], and ε_{cs} is the noise in the signal, assumed to be normally and independently distributed across states [$\varepsilon_{cs} \sim N(0,\sigma_\varepsilon^2)$].

Thus, the expected utility given the state level signal, the state level preferences and the individual preferences is given as follows:

$$E(u_{cis}|\theta_{cs}, \eta_{cs}, \nu_{cis}) = E(q_c|\theta_{cs}) + \eta_{cs} + \nu_{cis} \tag{3}$$

Voting behavior is based on sincere voting, as voters are assumed to support the candidate that provides the highest expected utility level[7]. As shown above, this expected utility level depends on expected candidate quality that is privately updated given the state level signal (θ_{cs}) and the prior (μ_{ct}, σ_t^2):

$$E(q_c|\theta_{cs}) = \alpha_t\theta_{cs} + (1-\alpha_t)\mu_{ct} \tag{4}$$

where α_t is the weight voters put on their signal. Knight and Schiff (2010), based on the Bayesian learning literature, suggest that this weight increases with the variance in the prior over quality (σ_t^2) and decreases with the degree of noise in the signal (σ_ε^2), given by:

$$\alpha_t = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_\varepsilon^2} \tag{5}$$

This suggests that voters put more weight on their private signal when the prior over quality deviates a lot from its expected value and thus is not very precise in informing

[3] Mordkoff (2000, 2011) explains that "according to the Central Limit Theorem, given random and independent samples of size N, the sampling distribution of the mean converges to normal the larger the size N, regardless the distribution of the population".

[4] According to Bayesian probability theory, as the priors and posteriors belong to the same family, they are conjugate distributions, meaning that the posterior distribution will also be normal, following the prior conjugate distribution.

[5] Kinght and Schiff (2010) explain that "signals could be town hall meetings with candidates, media coverage of candidate debates, political advertising etc".

[6] Knight and Schiff (2010) give a quite convincing explanation with regard to this assumption, as they suggest that individually independent signals instead of state-level ones would make the model unattractive and unrealistic (because then voters would perfectly know quality, revealed from voting returns of large states, and thus would only learn from the voting returns of the first state and not of the consequent ones).

[7] This approach proposed by Knight and Schiff (2010) is similar to Ali and Kartik (2010) regarding sincere voting, while it differs from Ali and Kartik (2006, 2010), and Dekel and Piccione (2000) in the sense that voters are not pivotal (they do not have forward looking incentives).

about quality (higher variance (σ_t^2)) and when the signal is not noisy (smaller variance (σ_ε^2)).

Given the expected candidate quality update, the expected utility is given by:

$$E(u_{cis}|\theta_{cs}, \eta_{cs}, \nu_{cis}) = \alpha_t\theta_{cs} + (1\text{-}\alpha_t)\mu_{ct} + \eta_{cs} + \nu_{cis} \qquad (6)$$

After defining the expected utility level of each candidate, the next step is to find the probability that a candidate will be chosen. Specifically, candidate *c* will be chosen by individual *i* if he gives the highest expected utility, that is, if $u_{cis}=\max\{u_{0is},\ldots,u_{Cis}\}$. To evaluate this probability, we have to say something about the maximum of a number of random variables, which becomes easy under the assumption that ν_{cis} follows the type I extreme value distribution, since its distribution function does not involve any unknown parameters (Verbeek, 2012). Thus, it can be shown that:

$$\Pr(E(u_{cis} \mid \theta_{cs}, \eta_{cs}, \nu_{cis}) > E(u_{dis} \mid \theta_{ds}, \eta_{ds}, \nu_{dis}); \forall d \neq c) = \frac{\exp(\alpha_t\theta_{cs} + (1-\alpha_t)\mu_{ct} + \eta_{cs})}{\sum_{c=0}^{C}\exp(\alpha_t\theta_{cs} + (1-\alpha_t)\mu_{ct} + \eta_{cs})} \qquad (7)$$

These probabilities are the voting probabilities for each candidate and since voting behavior is based on sincere voting, they are assumed to be the vote shares for each candidate[8]. As our model is a multinomial logit model and the outcome is a log-odds ratio instead of a probability, we transform the probabilities to odds, and then log-odds ratio. Thus, the vote share (that is the probability ratio, or odds ratio) for candidate *c* relative to the baseline candidate 0 is the following:

$$u_{cst} / u_{0st} = \frac{\exp(\alpha_t\theta_{cs} + (1-\alpha_t)\mu_{ct} + \eta_{cs})}{\exp(\alpha_t\theta_{0s} + (1-\alpha_t)\mu_{0s} + \eta_{0s})} \qquad (8)$$

Recalling that the utility level of the baseline candidate 0 is normalized to be zero, we then take the log of equation (8), which results in the following log-odds ratio:

$$\ln(u_{cst} / u_{0st}) = \eta_{cs} + \alpha_t\theta_{cs} + (1-\alpha_t)\mu_{ct} \qquad (9)$$

It is known from the econometric literature that the multinomial logit model gives the logit-transformed probability (log-odds ratio) as a linear function of the predictor variables. In our case these variables are the state-level preferences (η_{cs}), the state-level signal (θ_{cs}) weighted by the parameter α_t and the mean (μ_{ct}) of the prior over candidate quality, which is the quality distribution prior to the realization of the signal, weighted by the complement of this parameter ($1\text{-}\alpha_t$).

Now recall that we have assumed uncertainty over quality, in order to be able to test for potential momentum effects. This assumption is crucial because it implies that

[8] Knight and Schiff (2010) explain that the state-level vote shares are equal to the voting probabilities under the assumption of a continuum of voters, because in a model with such an assumption sincere voting is an equilibrium (individual voter behavior does not affect the overall vote shares not the behavior of consequent voters).

voters in late states observe the vote shares in early states and with the signal they receive, they can update their beliefs over quality, as it is shown by equation (2) given by $\theta_{cs} = q_c + \varepsilon_{cs}$. However, updating voter beliefs over quality is not possible by directly using the vote shares of equation (9), since the state-level preferences (η_{cs}) are not observed by voters in other states. To overcome this problem, Knight and Schiff (2010) proposed a transformation of vote shares that would provide a public noisy voting signal of quality that combines the state level private noisy signal (θ_{cs}) and a belief update from voting returns, by rearranging equation (9) as follows:

$$\frac{\ln(u_{cst}/u_{0st}) - (1-\alpha_t)\mu_{ct}}{\alpha_t} = q_c + \frac{\eta_{cs}}{\alpha_t} + \varepsilon_{cs} \tag{10}$$

Equation (10) gives the noisy voting signal of quality, which includes the noise in the state level private signal over quality (ε_{cs}) and the noise caused by unobserved state preferences ($\eta_{cs}/\ \alpha_t$). The noise in the voting signal [($\eta_{cs}/\ \alpha_t$) + (ε_{cs})] has a combined variance, equal to $(\sigma_\eta^2/\alpha_t^2) + \sigma_\varepsilon^2$. At time *t*, voters in later states receive a number of voting signals equal to the number of states voting at time *t*, that is $N_t \geq 1$ signals. Thus, the posterior distribution over quality that refers to time *t+1*, is also normal following the prior distribution, so it is characterized by its first two moments, which are the prior and the voting signals from the states that voted at time *t*. The posterior distribution is (μ_{ct+1}, σ_{t+1}^2), where the mean is the sum of the mean of the prior (μ_{ct}) and the average (mean) of the voting signals at time *t* and is given by:

$$\mu_{ct+1} = \beta_t \left[\frac{1}{N_t} \sum_{s \in \Omega_t} \frac{\ln(u_{cst}/u_{0st}) - (1-\alpha_t)\mu_{ct}}{\alpha_t} \right] + (1-\beta_t)\mu_{ct} \tag{11}$$

The inverse of the variance (the posterior's precision) increases with the number of states (N_t) voting at time *t* and the precision in the voting signals (that is the inverse of the variance of the noise in the voting signal $[(\sigma_\eta^2/\alpha_t^2) + \sigma_\varepsilon^2]^{-1}$), as follows:

$$\frac{1}{\sigma_{t+1}^2} = \frac{1}{\sigma_t^2} + \frac{N_t}{(\sigma_\eta^2/\alpha_t^2) + \sigma_\varepsilon^2} \tag{12}$$

The weight on the average of the voting signals is as follows:

$$\beta_t = \frac{N_t \sigma_t^2}{N_t \sigma_t^2 + (\sigma_\eta^2/\alpha_t^2) + \sigma_\varepsilon^2} \tag{13}$$

This posterior distribution over candidate quality (μ_{ct+1}, σ_{t+1}^2) can be interpreted as the update in the voter beliefs over quality. This update, or in other words this social learning can be seen by transforming equation (11) in the following way:

$$\mu_{ct+1} - \mu_{ct} = \frac{\beta_t / N_t}{\alpha_t} \sum_{s \in \Omega_t} [\ln(u_{cst}/u_{0st}) - \mu_{ct}] \tag{14}$$

where the difference between the mean of the posterior and the mean of the prior (μ_{ct+1} - μ_{ct}) is the social learning over candidate quality from the voting returns at time *t*. Equation (14) suggests that social learning does not depend on the primary wins of candidates, but on the surprises in voting returns (Knight and Schiff, 2010). This social learning rule implies that momentum effects can occur for candidates that perform better that expected, even if they do not win the primary election (positive deviation of vote shares ($\ln(u_{cst}/u_{0st})$) from expected performance (μ_{ct})). The reverse also holds, implying that candidates might win the primary election, but do not necessarily benefit from momentum effects and they can even be negatively affected if their voting shares are smaller than expected (negative deviation of vote shares ($\ln(u_{cst}/u_{0st})$) from expectation (μ_{ct})).

At this point, it is indicative to highlight that expectations over candidate performance, given by μ_{ct}, are not state specific, but national, because we have assumed that state level preferences are unobserved by other states. Knight and Schiff (2010) explain that this assumption is very important, because if state level preferences were perfectly observed by voters in all states, then voters would update their beliefs over quality directly from the voting returns, since they would infer the private state level signals directly from equation (9). This would imply that preferences in early states do not have a disproportionate impact on final results of the primary election, excluding one of the most important features of the primary system[9]. Another reason why state preferences should be assumed unobserved is because it is practically impossible that they are observed by all voters in all states, since polls are not frequently reported in several states. Finally, with perfectly observed state preferences, the weight on the private signal (α_t) would be equal to the weight on the public signal (β_t) for dates when only one state votes (N_t=1). Then, momentum effects would be assumed from the updating rule, rather than estimated. Thus, it is important for our analysis to assume that state level preferences are unobserved by voters of other states, in order to be able to estimate momentum effects of voting returns[10].

As a final remark of the theoretical framework governing social learning, it is interesting to explain how an increase in the vote shares of one state affects the mean of the posterior distribution. This effect can be shown as follows:

$$\frac{\partial \mu_{ct+1}}{\partial \ln(u_{cst}/u_{0st})} = \frac{\beta_t / N_t}{\alpha_t} = \frac{\sigma_t^2 + \sigma_\varepsilon^2}{N_t \sigma_t^2 + (\sigma_\eta^2 / \alpha_t^2) + \sigma_\varepsilon^2} \tag{15}$$

The above formulation indicates that the degree of social learning from an increase in the vote shares of one state is given by the ratio of the weight on the public signal (β_t) to the weight on the private signal (α_t). As we have assumed that state preferences are unobserved by voters in other states, this ratio ranges between zero and one. As we

[9] Hummel and Knight (2012) indeed find that sequential elections (like the primary system) place too much weight on the preferences of early states.

[10] Knight and Schiff (2010) include also a specification where state preferences are partially observed.

explained previously, in the case of observed state preferences where there is no heterogeneity (σ_η^2=0) and on a date where only one state votes (N_t=1), we have that (α_t= β_t), so the ratio is one. When the degree of heterogeneity in state preferences increases (in the case of unobserved preferences), the ratio becomes smaller and moves towards zero. Regarding the effect of an update from the private signal, the ratio increases in the variance of the prior (σ_t^2) and decreases in the degree of noise in the signal (σ_ε^2), implying that social learning increases when the variance of the prior is large as it results in an update of the posterior, and it decreases when the noise of the private signal is large and thus makes updating difficult.

The analysis presented in this part aimed to illustrate how social learning can be derived from voting returns of early states. Following the instructions of Knight and Schiff (2010), and relying on well known econometric and Bayesian literature, we showed that social learning depends on the deviations of the actual vote shares from the expected electoral outcomes. In the following part, we continue with the empirical application of our study.

4. EMPIRICAL APPLICATION

Our empirical application focuses on the 2008 US Democratic primary. In this part we provide an analysis of our data collection as well as the empirical specifications with the main results.

4.1 Data

In order to test for social learning and potential momentum effects in the 2008 primaries, we use daily opinion polls to examine how voters react to candidate performance. The analysis is focused on the campaigns of Clinton, Obama and Edwards, where Clinton is considered to be the baseline candidate. The database we use is the National Annenberg Election Survey 2008, which provides both an online version with questionnaires filled in online and a phone version with phone interviews. To be more specific, the online version was conducted in a panel structure divided in five period waves covering the entire election process. However, each wave was conducted in a national rolling cross section, on a daily basis, and separated the periods to pre-primary, primary and post-election. The phone version of the database was conducted by daily interviews in a national rolling cross-section design, starting from December 17, 2007 and covering the whole primary season through November 3, 2008, the day before the general election. Post-election re-interviews were also conducted between November 5 and 10, 2008, the week following the presidential election; nevertheless these data are not included in this study as we are interested in the primary season. Our data is a combination of the phone and the online versions. The main reason why we decided to use both versions for our social learning analysis was that the data on the phone version covered a too small pre-primary period, starting from December 17, 2007, thus we would not be able to draw safe conclusions on potential momentum effects. Nor could we compare our results to the ones of Knight and Schiff (2010), as their sample covered a wider pre-primary period starting from October 7, 2003. Adding the data from just the first wave of the online version, that cover a three month pre-primary period from October 2, 2007, to the data of the phone version, extends our pre-primary period and thus provides a complete sample for our analysis and makes our results comparable to the ones of Knight and Schiff (2010)[11].

In order to use the same empirical model with Knight and Schiff (2010) that will allow us to test the same empirical question about social learning for the 2008 primaries, we constructed the final data in the same way, according to the supplement of their study[12]. To be more specific, we use a sample of 12,923 respondents from both the NAES Rolling Cross-Section (RCS) survey 2008 and the first wave of the NAES Online Survey 2008 that are likely Democratic primary voters in a period between October 2, 2007 and January 29, 2008, which was the day before Edwards suspended his nomination. This period covers the entire pre-primary season and the primary season just before Super Tuesday. The respondents of the survey are voters who live in states that have not held their primaries yet, which means that residents of a state that has already held the primary are not asked their voting intentions in the survey, as we are interested in checking how early results affect the voting intentions of later voters.

As it was described in the theoretical framework above and will be further analyzed in the empirical model below, we are interested in estimating the state-level preferences (η_{cs}) as part of our social learning estimation process. For that reason, we can only include in our analysis those states that have enough respondents who would vote for Obama, Clinton or Edwards, as these are the three candidates we focus on. Thus, we have to exclude nine states from our analysis, which are: Delaware (DE), Montana (MT), Vermont (VT), North Dakota (ND), Rhode Island (RI), South Dakota (SD), Wyoming (WY), Hawaii (HI) and Alaska (AK)[13].

Furthermore, as it was explained previously in the theoretical framework, in order to estimate any momentum effects and social learning, we need the actual vote shares of each candidate relative to the baseline candidate. The data for the vote shares of the 2004 democratic primary are taken from the website of CNN (www.cnn.com). These data, which are state aggregate vote shares, are merged with the individual-level data taken from the NAES 2008.

---

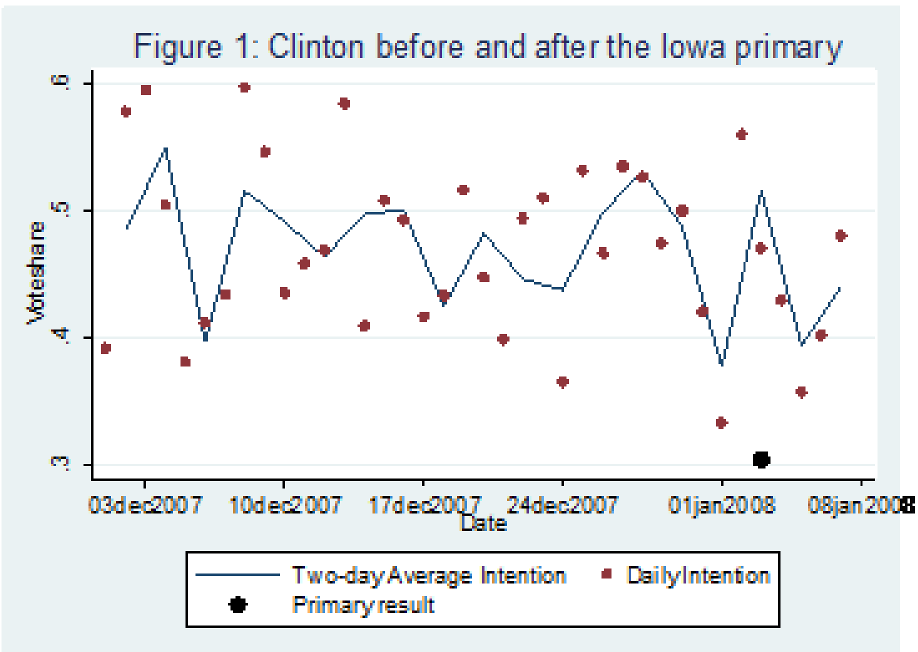
[11] Both versions of the survey used basically the same questions, adding some questions or rephrasing some others in some cases. Our combined data are taken from the responses of the same questions, so there is no difference on the data that could cause problems in combining them.

[12] Professor Nathan Schiff provided us with the supplement of the study that allowed us to construct the data in a similar way as in their work.

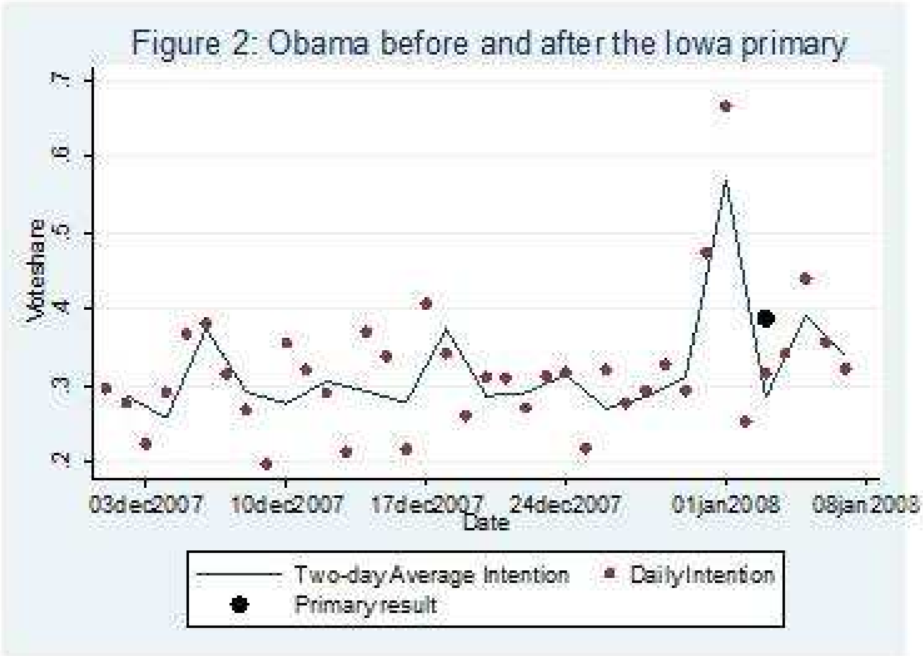
[13] The selection of the states to be excluded was based on the number of respondents per state who would vote either for Obama or for Clinton or for Edwards. If the total number of these respondents of a specific state was less than 50, the state was excluded from the analysis because this number of observations would not be sufficient to estimate the state level preferences.

In Knight and Schiff (2010), the authors summarize their identification strategy using only the case of Iowa in three graphs that show the daily voting intentions of the respondents during a month before the Iowa caucus compared to the actual result of the caucus. These graphs illustrate the performance of the candidates before and after the Iowa caucus as well as any updates from the respondents after the result was published. We decided to include similar graphs to our analysis, as they are very informative regarding momentum and social learning.

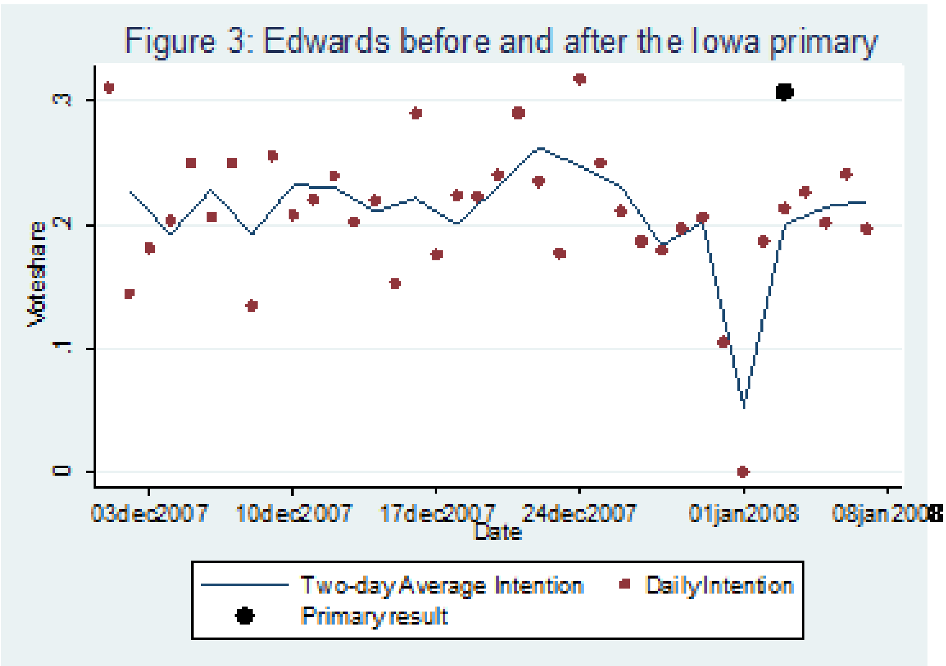
Specifically, Figures 1, 2 and 3, referring to Clinton, Obama and Edwards respectively, show the daily intention as well as the two-day average intention of the respondents, covering the month before the Iowa caucus took place and including the few days after the Iowa caucus and before the New Hampshire primary (1/12/2008-7/1/2008). The actual result of the primary is shown by a black dot. By looking at the three figures below, we can see that Clinton led Obama and Edwards during the month preceding the Iowa caucus, with the exception that just a few days before the primary date, Obama passed Clinton in intended vote shares. Nevertheless, the general pattern of voter intentions shows that Clinton holds the lead followed by Obama and Edwards. Taking into account the primary results from Iowa, we can see that Clinton underperformed in Iowa relative to expectations, and respondents of the survey updated their intentions in the next few days with the pattern going slightly down; however Clinton’s underperformance did not cost her much at this stage. On the contrary, Obama and Edwards outperformed relative to expectations with Obama winning the primary and Edwards finishing second. However, respondents of the survey did not update appropriately during the next few days until the New Hampshire primary, keeping voting intentions for both Obama and Edwards to the pre-Iowa levels.



Source: STATA



Source: STATA



Source: STATA

4.2 Empirical Model

4.2.1 Baseline Specification

Our interest in testing for potential social learning in the 2008 primaries and providing results comparable to the ones of Knight and Schiff (2010), led us to adopt the theoretical framework they developed in their paper. For that reason, we will also use the same empirical specification they used, as it derives directly from their theoretical framework, an analysis of which we presented in Part 3. In this part, we explain the econometric model of our case.

As it was mentioned previously, in order to examine the potential social learning in the primary season, we have to examine how voter support for candidates evolves when the first primaries take place and the voting returns of early states are released. The assumption that voters engage to sincere voting and thus will support the candidate that maximizes their utility led our theoretical analysis to equation (7). This equation summarizes the voting probabilities, which are assumed to be the voting intentions for each candidate, since we assume sincere voting. Equation (7) shows that voting intentions depend on the state level preferences and the private updating over quality by the state specific private signal (θ_{cs}) and the prior (μ_{ct}, σ_t^2):

$$\Pr(E(u_{cis} \mid \theta_{cs}, \eta_{cs}, \nu_{cis}) > E(u_{dis} \mid \theta_{ds}, \eta_{ds}, \nu_{dis}); \forall d \neq c) = \frac{\exp(\alpha_t \theta_{cs} + (1-\alpha_t)\mu_{ct} + \eta_{cs})}{\sum_{c=0}^{C} \exp(\alpha_t \theta_{cs} + (1-\alpha_t)\mu_{ct} + \eta_{cs})} \quad (7)$$

In the theoretical framework we have already assumed that the state specific signal that affects candidate quality is common for voters living in the same state and unobserved by voters residing in other states. In order to proceed with our econometric model, we further assume that voters from a state receive their signals just before the date of their primary, which means that voters from late states have not observed their private signals in the stage of the primary season our analysis focuses on. This assumption makes the estimation of the social learning parameters much easier as quality is not weighed by the private signal and depends only on the prior. Thus, voting intentions from equation (7) can be modified to equation (16) as follows:

$$\Pr(E(u_{cis} \mid \eta_{cs}, \nu_{cis}) > E(u_{dis} \mid \eta_{ds}, \nu_{dis}); \forall d \neq c) = \frac{\exp(\eta_{cs} + \mu_{ct})}{\sum_{c=0}^{C} \exp(\eta_{cs} + \mu_{ct})} \quad (16)$$

In the theoretical framework we also explained that social learning derives from the update in the voter beliefs over candidate quality, given by the posterior distribution over candidate quality (μ_{ct+1}, σ_{t+1}^2), as shown in equation (14):

$$\mu_{ct+1} - \mu_{ct} = \frac{\beta_t / N_t}{\alpha_t} \sum_{s \in \Omega_t} [\ln(u_{cst} / u_{0st}) - \mu_{ct}] \quad (14)$$

Taking a closer look at equation (14), it can be seen that social learning is actually the difference between the mean of the posterior and the mean of the prior and thus it can be summarized by the weight on the private signal α_t (equation 5), the weight on the public signal β_t (equation 13), the update over the mean (μ_{ct+1} - μ_{ct}) (equation 14) and the update over the variance (equation 12) as follows:

$$\alpha_t = \frac{\sigma_t^2}{\sigma_t^2 + \sigma_\varepsilon^2} \tag{5}$$

$$\beta_t = \frac{N_t \sigma_t^2}{N_t \sigma_t^2 + (\sigma_\eta^2 / \alpha_t^2) + \sigma_\varepsilon^2} \tag{13}$$

$$\mu_{ct+1} - \mu_{ct} = \frac{\beta_t / N_t}{\alpha_t} \sum_{s \in \Omega_t} [\ln(u_{cst} / u_{0st}) - \mu_{ct}] \tag{14}$$

$$\frac{1}{\sigma_{t+1}^2} = \frac{1}{\sigma_t^2} + \frac{N_t}{(\sigma_\eta^2 / \alpha_t^2) + \sigma_\varepsilon^2} \tag{12}$$

Thus, in order to be able to examine whether there is any potential social learning or momentum in the 2008 primary we need to estimate the above parameters for every period of our case. Specifically, in order to calculate the weight on the private signal in the first period (α_1) we need information about the initial prior (μ_{c1}, σ_1^2) and the degree of noise in the private signal (σ_ε^2). Then, in order to calculate the weight on the public signal (β_1) we need only information about the variance of the state level preferences (σ_η^2), provided that we already calculated (α_1). Thus, when we obtain (σ_η^2) and (σ_ε^2) and all the first period values (μ_{c1}, σ_1^2, α_1, β_1), we only need information about the voting returns of the first period to be able to calculate equations (14) and (12), which in turn will give us all the values for the second period (μ_{c2}, σ_2^2, α_2, β_2). If we continue this estimation process, along with information about voting returns for every period, we can compute these values for every period.

The date when a primary takes place is the start of the new period, with the whole pre-primary season up to the date of the Iowa caucus on January 3, 2008 being the first period. As described in the data section above, our sample covers the primary season until January 29, so we include five periods in our model. In our empirical case, only one state per period held its primary, although our empirical model allows for more states holding their primaries at the same period. This does not affect our estimations by any means; on the contrary it allows for the possibility to include more periods to the model, in case we want to extend the sample period.

At this point we should explain why we have five periods (with one state voting at every period respectively), while there were six democratic primaries held in the primary period of our sample between January 3 and January 29. The reason for excluding one period and thus not taking into account the voting returns from state Michigan in our analysis is that the Democratic National Committee penalized

Michigan by not counting the results of the Michigan Democratic Primary. This was because the Michigan Legislature passed a bill to move the date of the primary to January 15, although Federal Democratic Party rules only allow Iowa, New Hampshire, Nevada and South Carolina for holding primaries before February 5. Obama and Edwards among other candidates withdrew from the Michigan ballot and only Clinton and a few other candidates remained. As there would not be any voting returns from this state, and thus the results of the primary could not affect voter responses, we decided to exclude it from the analysis.

The Democratic National Committee did not penalize only Michigan in the 2008 elections, but Florida as well, which held its primary on January 29, the last period in our model. All Democratic candidates decided to withdraw from the Florida ballots; however a legislation rule in Florida suggests that any candidate who withdraws from the Florida primary will automatically withdraw from the overall Democratic nomination race. For that reason, the main three candidates that our analysis focuses on stayed in the ballot, so there are voting returns from the Florida primary. This is the main reason why we decided to include Florida in our analysis. Nevertheless, as probably there would not be any social learning from that state either, we also estimated the whole model with only four periods, reducing our sample to January 26[14]. Table 5 (see Appendix) shows the dates and the time periods of the primaries we used in our analysis. We report the results from both models in the next section.

The aforementioned analysis indicates that we need to estimate four key parameters in order to test whether there is social learning: the mean and the variance in the initial prior (μ_{c1}, σ_1^2), the variance in the state level preferences (σ_η^2) and the degree of noise in the private signal (σ_ε^2).[15] We adopt the two-step estimation process that was developed by Knight and Schiff (2010) to estimate these four parameters. The mean of the initial prior (μ_{c1}) along with the variance of the state level preferences (σ_η^2) are estimated in the first step, while the variance of the initial prior (σ_1^2) and the degree of noise in the private signal (σ_ε^2) are estimated in the second step. Below we explain the two-step approach more analytically.

In the first step, we run a multinomial logit model only for the first period (t=1) to estimate the first two parameters (μ_{c1}) and (σ_η^2). We use voting intentions from the pre-Iowa opinion polls that cover the pre-primary season between October 2, 2007 and January 3, 2008 (period t=1) in order to estimate the state level preferences (η_{cs}). Specifically, our multinomial logit model is given by equation (16) for t=1 as follows:

$$\Pr(E(u_{cis} \mid \eta_{cs}, \nu_{cis}) > E(u_{dis} \mid \eta_{ds}, \nu_{dis}); \forall d \neq c) = \frac{\exp(\eta_{cs} + \mu_{c1})}{\sum_{c=0}^{C} \exp(\eta_{cs} + \mu_{c1})} \tag{17}$$

[14] The reduced sample is 12,521 respondents.

[15] Knight and Schiff (2010) explicitly mention in their analysis that the degree of noise in the signal is assumed to be time independent, and even in the case of a release of more information at an early or later part of the campaign that would make the variance (σ_ε^2) time dependent [$\sigma_\varepsilon^2(t)$], our estimate can be considered as the average variance (σ_ε^2) of the primary season.

The dependent variable in this model is the candidate for whom the respondents of the survey intended to vote (Clinton is the baseline candidate and Obama and Edwards the remaining two alternatives), while the independent variables are dummy variables for each state[16]. At this point we should recall the assumption that state level preferences are mutually independent, meaning that voters do not observe the state level preferences of the other states. This assumption is linked to the fact that we use state dummy variables that have to be independent in order to avoid multicollinearity. The estimated coefficients are $\beta_{cs}=\eta_{cs}+\mu_{c1}$ and *c* refers to Obama and Edwards, because the utility from Clinton is normalized to zero (state level preferences for Clinton are zero). We have also assumed in the theoretical framework that the state level preferences are distributed $\eta_{cs} \sim N(0,\sigma_{\eta}^{2})$, therefore we normalize the state dummy variables to sum to zero and we demean the coefficients on the state dummies (state fixed effects) from our multinomial logit, estimated without a constant, in order to get the η_{cs}. The standard deviation of the state coefficients for each candidate is used to estimate candidate specific variances, and the average of these two variances is the variance (σ_{η}^{2}) of the state level preferences. Also, the average of these coefficients is the mean of the initial prior (μ_{c1}).

In the second step, we run a maximum likelihood model for all the remaining periods to estimate the last two parameters (σ_{1}^{2}) and (σ_{ε}^{2}). The idea is that we use the reactions of voters, captured by voting intentions in post-Iowa opinion polls, to the release of the voting returns from the states that held their primary in the remaining four periods of our sample, that are states voting prior to January 29. The log likelihood that is maximized is the sum of the probabilities for each of the three candidates that his or her utility will have the highest estimate. These utilities that are included in the probabilities are functions of the estimated state level preferences η_{cs} from the first step and are updated with the mean of the posterior of every period. As shown in equation (14), these means μ_{c2} to μ_{c5} are a function of (σ_{1}^{2}) and (σ_{ε}^{2}), which are the only unknown parameters, thus they are estimated by maximizing the likelihood.

Like Knight and Schiff (2010) did in their model, we decided not to use conventional confidence intervals for our standard errors, because we follow a two-step procedure and we use the estimated regressors from the first step to the second stage of estimations, which causes an uncertainty that cannot be reflected by conventional confidence intervals. Therefore, in order to calculate standard errors for the overall estimation, we compute bootstrap confidence intervals by bootstrapping the results of the two step estimation. To compute these intervals we have to estimate results for 100 bootsamples, drawn randomly from our original sample with replacement. We adopted the way that Knight and Schiff (2010) used to create the bootsamples for their case, where they controlled for the fact that if a randomly drawn sample did not include at least one voter supporting each candidate in a state, thus

[16] Except from the nine states we excluded from the analysis because of lack of data, as explained in the Data Section.

making estimation of the state level preferences impossible, then all observations from this state were dropped. For that reason, in Tables 1-3 we report the median values from the estimated state specific coefficients of 100 bootsamples, while the confidence intervals account for the second and the ninety-eighth percentiles (2nd lowest and 98th highest value), representing the bootstrap percentile confidence interval for 95% confidence[17].

Before we move forward to the analysis of our results, it is informative at this point to highlight the interpretation Knight and Schiff (2010) gave for the key social learning parameters that reflect the responses of voters to the revelation of voting returns in early states. Specifically, the authors explain that there is no social learning when the voters do not respond to the release voting returns, so the variance in the initial prior (σ_1^2) will have a small estimate and the degree of noise in the signal (σ_ε^2) will have a large estimate. On the contrary, they suggest that there is social learning when the voters respond to voting returns, so the variance in the initial prior (σ_1^2) will have a large estimate and the degree of noise in the signal (σ_ε^2) will have a small estimate.

4.2.2 Baseline Results

In this section, we present the results of our baseline specification. Table 1 shows the results of the multinomial logit model from the first step of our estimation, while Table 2, columns (1) and (2) provide the results of the maximum likelihood estimation from the second step of our analysis. Both tables report the results of the model when we use five periods (excluding Michigan only) and of the model when we use four periods (excluding Michigan and Florida). When we compare these two models, it is obvious that the results are quite similar, with almost no changes in the significance levels of the coefficients and with very slight changes in their magnitudes. This can probably be explained by the fact that Florida got penalized by the Democratic National Committee for holding its primary earlier than allowed, and this influenced voters not to take into account the results at this point. Thus, the Florida primary was not informative and it could not lead to any social learning. Here we will focus our analysis on the four-period model.

Looking at Table 1 in columns (3) and (4), we see that the coefficients of the constant term, which can be interpreted as the mean of the initial prior, are both negative and statistically significant at the 5% level. This indicates that Clinton has a substantial lead over Obama and Edwards in the first period covering the whole pre-primary season. This goes in accordance with Figures 1-3 shown in the data section above, that indeed demonstrate that Clinton was leading the voter preferences during the month preceding the Iowa caucus. Looking at the coefficients of the state dummy variables, we can conclude that there is a relative variation in candidate preferences across states. At this point it is interesting to specifically look at the coefficients of the

[17] "The Practice of Business Statistics", by Moore et al, Chapter 18, page 18-40.

home states of the three candidates to see whether they enjoy any regional advantages. Indeed, Obama has a substantial advantage over Clinton in his home state of Illinois as the Illinois coefficient for Obama is positive and statistically significant at the 5% level. Similarly, Edwards enjoys statistically significant advantages over Clinton in his home state of North Carolina, while Clinton holds an advantage in her home state of New York, with both coefficients for Obama and Edwards being negative and statistically significant at the 5% level.

The results of the key social learning parameters are shown in column (2) of Table 2. The estimated variance (σ_η^2) that shows the degree of heterogeneity in the state level preferences is equal to 0.219 and statistically significant at the 5% level, while its magnitude suggests that there is not so much heterogeneity in state level preferences. The variance in the initial prior (σ_1^2) is also significant at the 5% level and equals 0.147. This is a relatively small estimate, which probably suggests that voters are unresponsive to the release of information on voting returns and thus there is no social learning in the 2008 primaries. The estimated variance (σ_ε^2) that shows the degree of noise in the signal is even smaller than the variance in the initial prior and almost approaches zero, but it is insignificant, thus we cannot draw safe results regarding social learning from this estimate.

If we focus on the intuition Knight and Schiff (2010) gave regarding these two variances, explained in the end of the previous section, our conclusion that there is no social learning can be based only on the magnitude of the variance in the initial prior. This conclusion can be also supported by Figures 1-3, because they indicate that although Clinton underperformed while Obama and Edwards outperformed in Iowa, the survey respondents did not update their voting intentions appropriately during the days following the primary, thus candidates did not seem to enjoy any momentum effects. More specifically, prior to the Iowa caucus, voters ranked Clinton on top of their preferences followed by Obama and Edwards, as it can be seen both from the coefficients of the constant term in Table 1 and from the pattern in Figures 1-3. With the start of the primary season, Obama won in Iowa; however he did not enjoy any momentum effects as he ranked second in New Hampshire, which suggests that although he outperformed expectations in the first primary, voters did not update properly. The New Hampshire primary results are very crucial for the social learning analysis as they contradict with the initial pre-Iowa preferences of the state. Before the Iowa primary, New Hampshire preferences favored Obama and then Edwards over Clinton, which ex-post was in accordance with the Iowa results. Based on that, we would expect that the Iowa results would strengthen the state preferences and New Hampshire would vote for Obama and then Edwards, causing momentum to Obama and giving a second place finisher’s advantage to Edwards. However, the New Hampshire results offered victory to Clinton, suggesting that NH voters were not influenced by the Iowa results, but on the contrary they put weight on the the initial prior. Clinton sustained her ranking of the first place finisher in Nevada, but she lost to Obama in South Carolina.

Nevertheless, as the magnitudes of our results cannot be associated with the ones of Knight and Schiff, we can use our own approach to explain them, based on our theoretical framework. For that reason, it is instructive at this point to give some intuition regarding the magnitude of the degree of noise in the signal (σ_ε^2). As it can be seen in Table 2, this variance is very small and approaches zero, which intuitively means that there is no noise in the private signal over quality and consequently voters put a large weight on their private signal that they receive just before their primary, while they put almost no weight on the prior. This suggests that expected quality is mostly updated by the private signal instead of the initial prior. Now let's take a look at how this large weight on the private signal affects the public noisy voting signal of quality that updates voter beliefs. As we mentioned in the theoretical framework, the noise in the voting signal has a combined variance equal to (σ_η^2/α_t^2) + σ_ε^2, shown by equation (10). Our results suggest that there is not much heterogeneity in state level preferences (relatively small (σ_η^2)), while the weight on the private signal is very large. Thus, the degree of noise in the voting signal over quality is quite small. This would suggest that the weight voters put on the public voting signal, shown in equation (13), is large; however, this weight increases also with the variance in the initial prior, and as our estimated variance (σ_1^2) is relatively small, it can be shown that the weight on the public voting signal (β_1) is also small in the beginning of the primary season. Consequently, it can be concluded that voters do not put much weight on updates from voting returns, meaning that social learning is not necessary because expected candidate quality is updated privately.

The aforementioned interpretation is opposed to the intuition deriving from the interpretation proposed by Knight and Schiff (2010) that we described above (page 20). Specifically, this alternative interpretation suggests that although the variance in the initial prior is small, voters do not put weight on the initial prior because they put their weight on the private signal that has almost no noise. Thus, social learning does not need to evolve from voting returns, as voters already know everything about candidate quality privately. However, we decided to include both interpretations, mainly because the estimated magnitude of the variance (σ_ε^2) is insignificant.

TABLE 1

MULTINOMIAL LOGIT MODEL - FIRST STAGE OF ESTIMATION

	BASELINE SPECIFICATION			
	5 Periods		4 Periods	
	Obama	Edwards	Obama	Edwards
	(1)	(2)	(3)	(4)
Constant	-0.568**	-0.898**	-0.559**	-0.896**
	[-0.622, -0.499]	[-0.963, -0.812]	[-0.632, -0.454]	[-0.985, -0.802]
AL	0.211	-0.437	0.214	-0.344
	[-0.291, 0.680]	[-1.021, 0.088]	[-0.104, 0.571]	[-0.890, 0.078]
AR	-1.129**	-0.224	-1.131**	-0.255

	[-2.126, -0.488]	[-0.878, 0.202]	[-1.934, -0.635]	[-0.889, 0.218]
AZ	-0.796**	-0.516**	-0.802**	-0.457**
	[-1.313, -0.387]	[-1.181, -0.092]	[-1.242, -0.351]	[-1.025, -0.096]
CA	-0.055	-0.275**	-0.064	-0.298**
	[-0.172, 0.121]	[-0.460, -0.071]	[-0.222, 0.111]	[-0.495, -0.092]
CO	0.018	-0.024	0.015	-0.022
	[-0.343, 0.297]	[-0.490, 0.302]	[-0.365, 0.400]	[-0.539, 0.315]
CT	0.369*	-0.387	0.378*	-0.342
	[-0.017, 0.822]	[-0.970, 0.091]	[-0.035, 0.630]	[-1.017, 0.234]
DC	0.938**	-0.107	0.990**	0.011
	[0.387, 1.458]	[-1.393, 0.853]	[0.281, 1.552]	[-1.059, 1.005]
FL	-0.482**	-0.139	-0.501**	-0.137
	[-0.813, -0.264]	[-0.422, 0.112]	[-0.730, -0.280]	[-0.330, 0.059]
GA	0.423**	-0.079	0.422**	-0.037
	[0.082, 0.784]	[-0.583, 0.277]	[0.135, 0.788]	[-0.437, 0.280]
IA	0.140	0.702**	0.180	0.722**
	[-0.236, 0.527]	[0.364, 1.127]	[-0.420, 0.659]	[0.267, 1.135]
ID	0.562*	0.869**	0.614*	0.808**
	[-0.046, 1.303]	[0.291, 1.490]	[-0.142, 1.227]	[0.276, 1.574]
IL	0.868**	-0.216*	0.889**	-0.166
	[0.692, 1.070]	[-0.425, 0.061]	[0.694, 1.056]	[-0.452, 0.121]
IN	0.198	0.195	0.182	0.209
	[-0.104, 0.404]	[-0.102, 0.533]	[-0.099, 0.463]	[-0.251, 0.504]
KS	0.026	0.297	-0.023	0.331
	[-0.560, 0.624]	[-0.542, 0.881]	[-0.621, 0.703]	[-0.294, 0.801]
KY	0.104	0.438*	0.136	0.528*
	[-0.369, 0.490]	[-0.044, 0.883]	[-0.316, 0.521]	[-0.117, 0.936]
LA	-0.070	0.046	-0.159	0.077
	[-0.663, 0.402]	[-0.550, 0.521]	[-0.601, 0.356]	[-0.550, 0.470]
MA	-0.409**	-0.440**	-0.353**	-0.453**
	[-0.744, -0.077]	[-0.924, -0.054]	[-0.741, -0.041]	[-0.892, -0.088]
MD	0.027	-0.674**	0.021	-0.699**
	[-0.273, 0.330]	[-1.356, -0.238]	[-0.260, 0.289]	[-1.140, -0.285]
ME	-0.050	0.042	-0.032	-0.050
	[-0.474, 0.354]	[-0.652, 0.721]	[-0.705, 0.607]	[-0.673, 0.531]
MI	-0.162	-0.204*	-0.146	-0.147
	[-0.366, 0.127]	[-0.421, 0.012]	[-0.374, 0.089]	[-0.473, 0.163]
MN	-0.006	0.087	-0.016	0.099
	[-0.224, 0.284]	[-0.187, 0.412]	[-0.354, 0.271]	[-0.412, 0.310]
MO	0.165	0.333**	0.113	0.285*
	[-0.132, 0.489]	[0.011, 0.654]	[-0.172, 0.489]	[-0.038, 0.678]
MS	0.276	-0.306	0.237	-0.385
	[-0.430, 0.871]	[-1.318, 0.467]	[-0.430, 0.911]	[-1.281, 0.417]
NC	-0.039	0.476**	-0.043	0.453**
	[-0.380, 0.235]	[0.077, 0.798]	[-0.414, 0.239]	[0.216, 0.724]
NE	-0.203	0.246	-0.142	0.204

	[-0.796, 0.415]	[-0.518, 0.975]	[-0.763, 0.372]	[-0.411, 0.831]
NH	0.654*	0.743**	0.665**	0.828**
	[-0.125, 1.220]	[0.094, 1.495]	[0.035, 1.533]	[0.098, 1.672]
NJ	-0.297*	-0.439**	-0.324*	-0.396**
	[-0.620, 0.055]	[-0.712, -0.005]	[-0.660, 0.090]	[-0.806, -0.030]
NM	0.429	0.356	0.377	0.289
	[-0.341, 1.044]	[-0.430, 1.001]	[-0.540, 1.047]	[-0.536, 0.966]
NV	-0.379	-0.486	-0.398	-0.557
	[-1.381, 0.131]	[-1.821, 0.180]	[-0.864, 0.165]	[-1.415, 0.021]
NY	-0.468**	-0.434**	-0.458**	-0.406**
	[-0.755, -0.258]	[-0.676, -0.198]	[-0.675, -0.285]	[-0.603, -0.186]
OH	-0.134	0.343**	-0.128	0.331**
	[-0.339, 0.122]	[0.126, 0.585]	[-0.320, 0.162]	[0.143, 0.545]
OK	-0.612**	-0.020	-0.667**	-0.009
	[-1.126, -0.221]	[-0.437, 0.462]	[-1.362, -0.114]	[-0.554, 0.462]
OR	0.252	0.479**	0.228	0.513**
	[-0.083, 0.681]	[0.145, 0.831]	[-0.183, 0.560]	[0.210, 0.859]
PA	-0.377**	-0.218*	-0.344**	-0.242*
	[-0.706, -0.153]	[-0.465, 0.040]	[-0.586, -0.191]	[-0.456, 0.038]
SC	0.246	0.197	0.250	0.226
	[-0.215, 0.679]	[-0.311, 0.682]	[-0.248, 0.743]	[-0.412, 0.781]
TN	-0.044	0.272*	-0.036	0.323*
	[-0.513, 0.268]	[-0.070, 0.745]	[-0.591, 0.354]	[-0.143, 0.740]
TX	-0.095	-0.347**	-0.098	-0.327**
	[-0.303, 0.087]	[-0.593, -0.131]	[-0.279, 0.108]	[-0.566, -0.059]
UT	-0.080	-0.363	-0.006	-0.386
	[-0.742, 0.498]	[-1.152, 0.377]	[-0.792, 0.584]	[-2.013, 0.497]
VA	0.424**	0.172	0.406**	0.151
	[0.065, 0.841]	[-0.419, 0.616]	[0.005, 0.775]	[-0.312, 0.583]
WA	0.037	0.031	0.073	0.015
	[-0.251, 0.302]	[-0.388, 0.302]	[-0.239, 0.318]	[-0.299, 0.387]
WI	0.187	0.098	0.160	0.068
	[-0.105, 0.479]	[-0.150, 0.413]	[-0.097, 0.515]	[-0.243, 0.399]
WV	-0.416	0.120	-0.491	0.094
	[-1.139, 0.130]	[-0.540, 0.760]	[-1.212, 0.211]	[-0.509, 0.817]

95% bootstrap confidence intervals are indicated in brackets below every coefficient.
*Significant at 10%.
**Significant at 5%.

TABLE 2

ESTIMATED PARAMETERS- SECOND STAGE OF ESTIMATION

	BASELINE SPECIFICATION		DISTANCE SPECIFICATION	
	5 Periods (1)	4 Periods (2)	5 Periods (3)	4 Periods (4)
σ(η)^2	0.220**	0.219**	0.220**	0.221**
	[0.149, 0.283]	[0.167, 0.300]	[0.149, 0.284]	[0.166, 0.300]
σ(1)^2	0.147**	0.145**	0.153**	0.150**
	[0.071, 0.396]	[0.063, 0.483]	[0.088, 0.514]	[0.074, 0.497]
σ(ε)^3	0.000025	0.00001	0.00001	0.00001
	[-0.285, 1.247]	[-0.377, 0.298]	[-0.310, 0.794]	[-0.404, 0.317]

95% bootstrap confidence intervals are indicated in brackets below every coefficient.
**Significant at 5%.

To sum up, the pattern of the results during our primary period along with the small estimate of the variance in the initial prior suggest that voters did not put much weight on the voting returns of the early primaries. This conclusion is also supported by the intuition deriving from the small magnitude of the degree of noise in the signal. Thus our results indicate that the momentum effects that were present in the 2004 primary do not continue in the 2008 primary.

4.2.3 Distance Specification

In this specification we will relax the assumption that state level preferences are observed only by voters of the same state and are unobserved by voters residing in other states. Again we will follow the additional specification Knight and Schiff (2010) developed for their model and adapt it to our analysis. In this case, we will alternatively assume that state level preferences are a combination of both an unobserved part (η_{cs}) and an observed part (X_{cs}) that captures geographic advantages when politicians campaign in their home states. This alternative assumption will affect the vote shares and equation (9) will be transformed as follows:

$$\ln(u_{cst} / u_{0st}) = \eta_{cs} + \gamma X_{cs} + \alpha_t \theta_{cs} + (1-\alpha_t)\mu_{ct} \tag{18}$$

Now voting returns also depend on the (γX_{cs}) component that represents observed preferences, where γ is the weight on observed preferences that we will estimate in the model. Equation (18) will consequently result in a transformation of equation (14), which is one of the main social learning parameters, as follows:

$$\mu_{ct+1} - \mu_{ct} = \frac{\beta_t / N_t}{\alpha_t} \sum_{s \in \Omega_t} [\ln(u_{cst} / u_{0st}) - \gamma X_{cs} - \mu_{ct}] \tag{19}$$

Equation (19) captures the possibility that voters in late states take into account observed state level characteristics. Thus, voting returns showing that a candidate had an advantage in his home state do not necessarily cause momentum.

The baseline empirical model is easily adapted to this specification. The measure we use to represent (X_{cs}), proposed by Knight and Schiff (2010) is the distance between state *s* and the home state of candidate *c* relative to the distance between state *s* and New York, the home state of Clinton. We proceed with the first step of our estimation as in the baseline model, which derives the estimated fixed effects of the preferences. Then, we regress these estimated fixed effects on the distance measure as follows:

$$\eta_{cs} = \alpha_c + \beta rd_{cs} + \varepsilon_{cs} \tag{20}$$

where rd_{cs} is the distance measure and ε_{cs} are the residuals of the regression that we report as an estimate of the unobserved preferences (η_{cs}). The second step of the estimation is the same as in the baseline specification with the exception that the mean is updated using equation (19). The results of this regression are presented in Table 3 below, while the results of the key parameters are presented in columns (3) and (4) of Table 2 above.

4.2.4 Distance Results

As we can see in Table 3, distance has no significant effect on voting intentions. This suggests that voters in late states do not consider it as an important observed factor to incorporate into their expectations. This conclusion is also supported by the estimate of the variance of state level preferences (σ_η^2) (shown in columns (3) and (4) of Table 2) that remained almost the same, suggesting that distance does not capture any observed component of state preferences in our case. The mean in the initial prior, represented again by the constant term in Table 3, is the same as in the baseline specification, because we estimated it in the first step and we then regressed the estimated fixed effects on distance. The variance in the initial prior (σ_1^2) remained quite small, supporting the conclusion of no social learning from our baseline specification. Thus, in our case we cannot conclude that distance has an effect on the voting intentions of late voters, or in other words, that distance captures an observed part of state level preferences.

TABLE 3

MULTINOMIAL LOGIT MODEL - FIRST STAGE OF ESTIMATION

	DISTANCE SPECIFICATION			
	5 Periods		4 Periods	
	Obama	Edwards	Obama	Edwards
	(1)	(2)	(3)	(4)
Constant	-0.568**	-0.898**	-0.559**	-0.896**
	[-0.622, -0.499]	[-0.963, -0.812]	[-0.632, -0.454]	[-0.985, -0.802]
AL	0.221	-0.439	0.216	-0.332*
	[-0.177, 0.663]	[-1.048, 0.140]	[-0.102, 0.576]	[-0.881, 0.101]
AR	-1.112**	-0.225	-1.117**	-0.231
	[-1.972, -0.491]	[-0.954, 0.198]	[-1.972, -0.586]	[-0.786, 0.223]
AZ	-0.778**	-0.507**	-0.756**	-0.464**
	[-1.271, -0.388]	[-1.072, -0.096]	[-1.193, -0.334]	[-1.002, -0.175]
CA	-0.027	-0.277**	-0.049	-0.298**

	[-0.170, 0.146]	[-0.430, -0.069]	[-0.206, 0.135]	[-0.494, -0.094]
CO	0.040	-0.033	0.061	-0.008
	[-0.393, 0.312]	[-0.508, 0.306]	[-0.339, 0.476]	[-0.612, 0.330]
CT	0.316*	-0.419	0.350*	-0.397
	[-0.034, 0.769]	[-1.007, 0.054]	[-0.063, 0.641]	[-1.048, 0.127]
DC	0.922**	-0.096	0.938**	-0.004
	[0.352, 1.474]	[-1.409, 0.854]	[0.293, 1.494]	[-1.060, 1.002]
FL	-0.484**	-0.123	-0.504**	-0.136
	[-0.817, -0.232]	[-0.375, 0.114]	[-0.739, -0.290]	[-0.313, 0.074]
GA	0.416**	-0.046	0.426**	-0.010
	[0.057, 0.783]	[-0.558, 0.291]	[0.179, 0.786]	[-0.410, 0.309]
IA	0.149	0.718**	0.176	0.724**
	[-0.235, 0.578]	[0.363, 1.093]	[-0.431, 0.685]	[0.236, 1.133]
ID	0.573*	0.857**	0.635*	0.796**
	[-0.047, 1.191]	[0.334, 1.580]	[-0.133, 1.249]	[0.321, 1.566]
IL	0.889**	-0.214*	0.907**	-0.162
	[0.716, 1.096]	[-0.472, 0.065]	[0.692, 1.148]	[-0.439, 0.061]
IN	0.204	0.158	0.211	0.218
	[-0.060, 0.412]	[-0.100, 0.534]	[-0.077, 0.468]	[-0.249, 0.522]
KS	0.037	0.332	-0.001	0.308
	[-0.553, 0.639]	[-0.537, 0.883]	[-0.583, 0.658]	[-0.294, 0.807]
KY	0.107	0.472*	0.127	0.528**
	[-0.359, 0.548]	[-0.034, 0.894]	[-0.312, 0.562]	[-0.140, 0.975]
LA	-0.065	0.075	-0.124	0.092
	[-0.660, 0.418]	[-0.554, 0.536]	[-0.543, 0.378]	[-0.450, 0.530]
MA	-0.421**	-0.481**	-0.415**	-0.466**
	[-0.762, -0.146]	[-0.967, -0.091]	[-0.645, -0.072]	[-0.914, -0.097]
MD	-0.006	-0.707**	-0.005	-0.706**
	[-0.314, 0.337]	[-1.362, -0.377]	[-0.300, 0.291]	[-1.204, -0.289]
ME	-0.073	0.002	-0.083	-0.090
	[-0.628, 0.376]	[-0.660, 0.682]	[-0.753, 0.540]	[-0.736, 0.502]
MI	-0.166	-0.212**	-0.143	-0.165
	[-0.374, 0.119]	[-0.450, -0.003]	[-0.375, 0.101]	[-0.498, 0.149]
MN	0.003	0.079	-0.003	0.105
	[-0.228, 0.298]	[-0.182, 0.400]	[-0.348, 0.270]	[-0.428, 0.275]
MO	0.169	0.335**	0.117	0.295*
	[-0.133, 0.512]	[0.015, 0.664]	[-0.155, 0.411]	[-0.031, 0.683]
MS	0.277	-0.308	0.194	-0.422
	[-0.481, 0.881]	[-1.438, 0.595]	[-0.393, 0.935]	[-1.258, 0.412]
NC	-0.067	0.497**	-0.045	0.466**
	[-0.386, 0.215]	[0.095, 0.812]	[-0.411, 0.200]	[0.203, 0.740]
NE	-0.174	0.245	-0.145	0.203
	[-0.795, 0.435]	[-0.387, 0.872]	[-0.752, 0.372]	[-0.379, 0.833]
NH	0.654*	0.707**	0.610**	0.776**
	[-0.125, 1.220]	[0.099, 1.444]	[0.063, 1.480]	[0.040, 1.607]
NJ	-0.297**	-0.456**	-0.369*	-0.435**

	[-0.635, -0.032]	[-0.725, -0.052]	[-0.693, 0.082]	[-0.830, -0.034]
NM	0.428	0.354	0.428	0.291
	[-0.306, 1.098]	[-0.460, 1.101]	[-0.516, 1.075]	[-0.532, 0.982]
NV	-0.333	-0.485	-0.375	-0.527
	[-1.355, 0.154]	[-1.852, 0.181]	[-0.890, 0.153]	[-1.399, 0.026]
NY	-0.519**	-0.436**	-0.519**	-0.443**
	[-0.793, -0.269]	[-0.741, -0.204]	[-0.731, -0.279]	[-0.628, -0.193]
OH	-0.141	0.340**	-0.145	0.328**
	[-0.349, 0.103]	[0.118, 0.579]	[-0.351, 0.164]	[0.139, 0.543]
OK	-0.592**	-0.014	-0.642**	-0.005
	[-1.110, -0.238]	[-0.463, 0.470]	[-1.287, -0.047]	[-0.489, 0.472]
OR	0.290	0.497**	0.249	0.512**
	[-0.073, 0.720]	[0.140, 0.819]	[-0.138, 0.624]	[0.207, 0.868]
PA	-0.400**	-0.254*	-0.374**	-0.255*
	[-0.616, -0.182]	[-0.479, 0.006]	[-0.600, -0.158]	[-0.482, 0.014]
SC	0.234	0.183	0.254	0.231
	[-0.231, 0.646]	[-0.301, 0.701]	[-0.267, 0.743]	[-0.391, 0.775]
TN	-0.021	0.293*	-0.028	0.338*
	[-0.498, 0.288]	[-0.059, 0.687]	[-0.575, 0.361]	[-0.140, 0.745]
TX	-0.071	-0.325**	-0.072	-0.306**
	[-0.246, 0.143]	[-0.568, -0.106]	[-0.269, 0.173]	[-0.557, -0.032]
UT	-0.028	-0.334	-0.001	-0.395
	[-0.704, 0.528]	[-1.151, 0.378]	[-0.758, 0.676]	[-2.012, 0.501]
VA	0.413**	0.178	0.396*	0.140
	[0.104, 0.825]	[-0.238, 0.630]	[-0.062, 0.738]	[-0.309, 0.593]
WA	0.068	0.038	0.091	-0.0003
	[-0.250, 0.318]	[-0.399, 0.304]	[-0.252, 0.317]	[-0.291, 0.377]
WI	0.185	0.102	0.184	0.065
	[-0.100, 0.415]	[-0.154, 0.380]	[-0.092, 0.512]	[-0.257, 0.381]
WV	-0.441	0.092	-0.475	0.124
	[-1.154, 0.107]	[-0.540, 0.760]	[-1.238, 0.176]	[-0.474, 0.797]
distance	0.002	0.002	0.003	0.003
	[-0.004, 0.011]	[-0.004, 0.011]	[-0.005, 0.011]	[-0.005, 0.011]

95% bootstrap confidence intervals are indicated in brackets below every coefficient.
*Significant at 10%.
**Significant at 5%.

At this point we have completed our social learning analysis for the 2008 Democratic primary, using the theoretical framework and the empirical model Knight and Schiff (2010) developed. Our findings suggest that there is no social learning from the voting returns that leads to momentum effects. In the next part we develop an additional specification that examines some sources of voting behavior related to demographics.

5. ADDITIONAL SPECIFICATION: SOURCES OF VOTING BEHAVIOR

From the previous social learning analysis it was made clear that one of the most interesting questions concerning elections is voting behavior. Related research is focused on examining why voters voted the way they did, what are the factors that influenced their decision, or what are the implications of these results. According to Prysby and Scavo (2005), campaign events and incidents are not sufficient to answer these questions, but a deeper analysis of the voting behavior together with an understanding of the unique aspects of elections is needed to create a complete explanation. The same authors support that there are two major concerns arising while studying voting behavior. One concern lies in the explanation of the election result by identifying the sources of individual voting behavior, while the other concern focuses on changes in voting patterns over time. These two concerns are complementary; nevertheless they form different research questions.

While the main focus of our study was on examining social learning and potential momentum effects in the 2008 US presidential primaries, we decided to include an additional specification that examines voting behavior by identifying some of its sources. According to The BBC's study on electoral systems, voting and political attitudes in the UK, voting behavior is influenced by social class, geography, age and background, issue voting and media. Motivated by this study, we decided to test how voting behavior is affected by demographics. Specifically, we want to examine how demographic characteristics of voters influence the individuals' preference in voting for a specific candidate.

5.1 Demographic characteristics

This empirical application about voters' demographic characteristics also focuses on the 2008 Democratic primary. Our interest lies in testing the potential effect of demographic characteristics on individual candidate preference before the first primary (Iowa caucus) takes place.

5.1.1 Data

We chose to focus on the pre-primary season only, in order to see the effects of the demographics on candidate preferences without any influence from potential momentum effects and social learning that could arise during the primary season. For that reason, we use individual preferences from daily opinion polls on a rolling cross section covering the pre-primary season (2/10/2007- 1/1/2008), with the data taken from the Wave 1 of the 2008 Online Dataset of the National Annenberg Election Survey. We decided that for this part of our analysis the data from the online version are sufficient to answer our empirical question, thus we did not find it instructive to merge them with data from the phone version of the survey. Again we focus on three candidates, Clinton, Edwards and Obama, where Clinton is the baseline candidate.

5.1.2 Empirical Model

As the variable of interest (candidate preference) is of a discrete nature, the most suitable econometric model to use is the multinomial logit model. In this model we include several variables showing individual demographic characteristics in order to test whether these characteristics affect individual preferences on a specific candidate. The base model includes gender, age, education, income and racial identity/ethnicity. Most of these demographic variables are categorical and not continuous, in which case we should use a dummy variable for every category of each demographic characteristic[18]. Specifically, the model includes a male dummy (1 if male, 0 if female), age, eight dummies for each level of education[19], eighteen dummies for each level of income[20] and four dummies for ethnicity[21]. Thus the base model for the demographics specification is as follows:

$$P(preference_i = j) = \frac{\exp(\beta_j^{const} + age_i\beta_j^{age} + male_i\beta_j^{male} + educ_i^{'}\beta_j^{educ} + inc_i^{'}\beta_j^{inc} + ethnic_i^{'}\beta_j^{ethnic})}{\sum_{k=1}^{3}\exp(X_i^{'}\beta_k)}$$

where i stands for the individuals ($i=1,\ldots,n$, $n=8{,}627$) and j for the three candidates ($j=1,2,3$), $preference_i$ is the individual's candidate preference, β_j^{const} is the constant term, age_i is the age variable, which is continuous, $male_i$ is the male dummy, and $educ_i'$, inc_i', and $ethnic_i'$ are vectors of the category dummies for education, income and ethnicity respectively.

Apart from the base model, we also ran an extended model including metropolitan area as a demographic characteristic by adding the dummy $m.area_i$ (1 if the individual lives in metropolitan area, 0 otherwise). When we compare the log likelihoods of both the base and the extended model by simply running a likelihood ratio (lr) test, we can see that the null hypothesis that the two models are not statistically different is rejected, or in other words there is a statistically significant difference between the two models. For that reason, we can conclude that the extended model fits the data significantly better than the base model in the 5% significance level because it

[18] We excluded one category of each variable from the econometric model to keep it as the reference category, in order to avoid the dummy variable trap that would cause exact multicollinearity.

[19] The levels of education are nine: 1=Grade 8 or lower, 2=Some high school, no diploma, 3=High school diploma or equivalent, 4=Some college, no degree, 5=Associate degree, 6=Bachelor's degree, 7=Master's degree, 8=Professional degree, 9=Doctorate degree. The omitted category that acts as a reference group is education level 1.

[20] The levels of income as given from the NAES Questionnaire are: 1=Less than $5,000, 2=$5,000 to $7,499, 3=$7,500 to $9,999, 4=$10,000 to $12,499, 5=$12,500 to $14,999, 6=$15,000 to $19,999, 7=$20,000 to $24,999, 8=$25,000 to $29,999, 9=$30,000 to $34,999, 10=$35,000 to $39,999, 11=$40,000 to $49,999, 12=$50,000 to $59,999, 13=$60,000 to $74,999, 14=$75,000 to $84,999, 15=$85,000 to $99,999, 16=$100,000 to $124,999, 17=$125,000 to $149,999, 18=$150,000 to $174,999, 19=$175,000 or more and the omitted category that acts as the reference group is income level 1.

[21] The categories of racial identity are: 1= Non-Hispanic White, 2=Non-Hispanic Black, 3=Non-Hispanic other race, 4=Hispanic, 5=two or more races, non-Hispanic. The omitted category that acts as reference group is the first category.

explains a bigger part of the alterability of the dependent variable. As the statistically significant difference between the two models is in the 5% significance level and not in the 1%, we decided to illustrate both models, the results of which can be found in columns 1-4 of Table 4 below. The complete results regarding these models (including standard errors and z-values) can be found in the Appendix (Tables 4A-4B).

5.1.3 Demographics Results

As the models are multinomial logit models, the coefficients regarding each candidate (Edwards and Obama) will be interpreted compared to the baseline candidate (Clinton). Also, keeping in mind that the candidate-specific estimated coefficients are the log odds of preferring this specific candidate over the baseline candidate will help in the interpretation of the results. The results of the base model are shown in columns 1 and 2 while the results of the extended model are shown in columns 3 and 4 of Table 4. The p-values are reported in brackets under every coefficient.

With regard to the base specification, the coefficient of the candidate-specific constant term shows that Clinton has a substantial lead over Edwards and Obama in the individual preferences. These results are in accordance with Figures 1-3 illustrated in the social learning analysis above, which show that Clinton had a substantial lead over Edwards and Obama in voting intentions during the month preceding the Iowa caucus. Age does not seem to affect the odds of preferring Edwards over Clinton, while it is statistically significant for the preference between Obama and Clinton. As shown in column 2, its coefficient suggests that keeping the rest of the predictor variables at a fixed value, an increase in age by one unit slightly decreases the odds of preferring Obama over Clinton.

The coefficient of the male dummy suggests that Edwards and Obama have substantial advantage over Clinton, as it demonstrates that the odds of preferring Edwards over Clinton for males are almost 50% (49.9%) higher than the odds for females, while the odds of preferring Obama over Clinton for males are almost 18% higher than the odds for females[22], ceteris paribus. Education seems to have no statistically significant effect on the probability of preferring Edwards over Clinton; on the contrary it seems that for high levels of education, the more educated the individual, the higher the odds to prefer Obama over Clinton. To be more specific, it can be seen that obtaining an education level equivalent to college or higher

[22] The odds of preferring Edwards over Clinton for males (male1=1) over the odds of preferring Edwards over Clinton for females (male1=0) is exp(0.4049845)=1.499, which in terms of percent change means that the odds for males are 50% higher than the odds for females to choose Edwards over Clinton. The odds of preferring Obama over Clinton for males over the odds of preferring Obama over Clinton for males is similarly estimated as exp(0.1628935)=1.177, which again in percentage change terms means that the odds for males are 18% higher than for females to choose Obama over Clinton. All the following odds that are shown in this section are calculated in the same way, thus their estimation will not be reported.

(education lever 4 or higher) increases the odds of preferring Obama over Clinton ceteris paribus. Indicatively, the coefficient of education level 6 (predictor variable *educ6*) shows that the odds of preferring Obama over Clinton for an individual who obtains a bachelor's degree are 129% higher than the odds for an individual who does not have this level of education, ceteris paribus.

Regarding the effect of different levels of income on the probability of preferring a specific candidate over the baseline one, the results suggest that only the higher levels of income have a statistically significant effect on candidate preference. To be more specific, as we can see in column 1, only the coefficient of income level 15 (predictor variable *inc15*) is statistically significant, and it suggests that the odds of preferring Edwards over Clinton for an income level between $85,000 and $100,000 are 49% higher than the odds of choosing Edwards over Clinton for another level of income, ceteris paribus. Column 2 shows that more levels of income, although only the higher levels again, have an effect on candidate preference in the case of Obama relative to Clinton. If we take income level 15 again, we can see that keeping all the other predictor variables at fixed levels, the odds of preferring Obama to Clinton for this level of income are 65% higher than the odds for another level of income.

The coefficients for racial identity/ethnicity indicate some very interesting results as they are statistically significant for all ethnicities except for the one referring to people with two or more races (*ethnic5*), which was expected as they could belong to several groups of ethnicities, so they could not affect the probability of a candidate preference. Analytically, in column 1 the coefficient for black people, represented by predictor variable *ethnic2*, is negative and statistically significant at the 1% level and suggests that Clinton has the lead in preferences when the choice comes to Edwards and Clinton, with the odds for black people preferring Edwards to Clinton being 76% lower than the odds for non-black people. Looking at the coefficient of the same dummy variable in column 2 we can see that now it is positive and statistically significant at the 1% level. This is in accordance to expectations, as it makes sense that black people would tend to prefer Obama over Clinton because he is African American. Specifically, the odds of choosing Obama over Clinton for black people are 180% higher than the odds of making the same choice for non black people. Meanwhile, Clinton seems to have the lead in the preferences of individuals with Hispanic or other non-Hispanic racial identity ceteris paribus, as both coefficients for the variables *ethnic3* and *ethnic4* in columns 1 and 2 have a negative sign, meaning that the odds of choosing either Edwards or Obama over Clinton are lower for people of these racial identities.

The results of the extended model that also includes a dummy variable for metropolitan area are shown in columns 3 and 4 of Table 4. By looking at the coefficients it can be concluded that the results are similar to the base model for all the predictor variables, with a slight increase in the coefficients for the extended model, which in turn leads to a slight increase in the odds for the candidate preferences reported above. What is interesting in this model is the coefficient of *m.area1* in column 3, which is negative and statistically significant at the 1% level. This means that the odds of preferring Edwards over Clinton for an individual who lives in a metropolitan area are 20% lower than the odds of making the same choice for an individual who does not live in a metropolitan area. On the contrary, living in a metropolitan area does not seem to have any effect on the probability of choosing between Obama and Clinton as the coefficient of *m.area1* in column 4 is not statistically significant.

TABLE 4
Candidate Preference on Demographic Characteristics
Multinomial Logit Model

	Base Specification		Includes Met. Area	
	Edwards	Obama	Edwards	Obama
	(1)	(2)	(3)	(4)
constant	-0.9162977***	-0.9990069***	-0.7482644**	-0.988078***
	[0.006]	[0.004]	[0.028]	[0.005]
age	0.0004052	-0.0142715***	0.0004886	-0.0142525***
	[0.838]	[0.000]	[0.806]	[0.000]
male1	0.4049845***	0.1628935***	0.4053796***	0.1629428***
	[0.000]	[0.002]	[0.000]	[0.002]
educ2	-0.2442469	0.5018916	-0.2354872	0.5028281
	[0.405]	[0.118]	[0.422]	[0.117]
educ3	-0.1028757	0.2401515	-0.0904235	0.2412234
	[0.700]	[0.432]	[0.735]	[0.430]
educ4	-0.1111432	0.6298659**	-0.0869508	0.6317858**
	[0.678]	[0.039]	[0.746]	[0.038]
educ5	-0.1333525	0.598483*	-0.1120988	0.6001634*
	[0.632]	[0.055]	[0.688]	[0.054]
educ6	-0.1130776	0.8263117***	-0.0802286	0.8286518***
	[0.676]	[0.007]	[0.767]	[0.007]
educ7	-0.1945108	0.7517157**	-0.1665005	0.7535996**
	[0.484]	[0.016]	[0.550]	[0.015]
educ8	-0.0311849	1.065274***	0.0007392	1.067482***
	[0.927]	[0.002]	[0.998]	[0.002]
educ9	-0.0254189	0.669464*	-0.0017392	0.6708895*
	[0.940]	[0.067]	[0.996]	[0.067]
inc2	0.0046561	-0.0466919	0.0066292	-0.0466309
	[0.987]	[0.851]	[0.981]	[0.851]
inc3	-0.1327661	0.3883384	-0.1625512	0.3858378
	[0.655]	[0.113]	[0.585]	[0.115]
inc4	-0.1175191	0.1296273	-0.1334688	0.1287454
	[0.660]	[0.582]	[0.618]	[0.585]
inc5	-0.3102425	0.2848136	-0.3246134	0.2838351
	[0.266]	[0.215]	[0.245]	[0.216]
inc6	0.0810844	0.2847508	0.0642178	0.283833
	[0.730]	[0.167]	[0.785]	[0.169]

inc7	0.1235605	0.2581087	0.1061169	0.256878
	[0.587]	[0.199]	[0.641]	[0.202]
inc8	0.0361573	0.0710997	0.0255471	0.0705763
	[0.873]	[0.723]	[0.910]	[0.725]
inc9	-0.1230039	0.192723	-0.1356664	0.1922487
	[0.587]	[0.325]	[0.549]	[0.327]
inc10	0.0304495	0.2415132	0.0218953	0.2412783
	[0.890]	[0.209]	[0.921]	[0.210]
inc11	0.1507545	0.3481389*	0.1474419	0.3479381*
	[0.476]	[0.060]	[0.486]	[0.060]
inc12	0.0507012	0.2286244	0.046236	0.2283401
	[0.815]	[0.229]	[0.831]	[0.230]
inc13	0.0948956	0.2487476	0.0924227	0.2484639
	[0.656]	[0.182]	[0.664]	[0.182]
inc14	0.1035862	0.2753395	0.1020236	0.2752423
	[0.649]	[0.166]	[0.654]	[0.166]
inc15	0.4016568*	0.5000968**	0.4049695*	0.5004819**
	[0.078]	[0.013]	[0.076]	[0.013]
inc16	0.0883189	0.41867**	0.0984163	0.4194905**
	[0.703]	[0.037]	[0.671]	[0.037]
inc17	-0.1208997	0.1554056	-0.1071008	0.1563257
	[0.650]	[0.494]	[0.688]	[0.492]
inc18	0.0017687	0.2905067	0.0157919	0.2915958
	[0.995]	[0.273]	[0.960]	[0.271]
inc19	0.3780689	0.6074255**	0.3918034	0.6086095**
	[0.178]	[0.014]	[0.163]	[0.013]
ethnic2	-1.434537***	1.029921***	-1.41429***	1.031285***
	[0.000]	[0.000]	[0.000]	[0.000]
ethnic3	-1.20583***	-0.3803483**	-1.199309***	-0.3798852**
	[0.000]	[0.019]	[0.000]	[0.019]
ethnic4	-0.9047519***	-0.2586076**	-0.8856702***	-0.2573201**
	[0.000]	[0.012]	[0.000]	[0.013]
ethnic5	-0.2356955	0.122028	-0.2256224	0.1227191
	[0.235]	[0.487]	[0.256]	[0.485]
m.area1			-0.2234306***	-0.0154965
			[0.007]	[0.854]
Number of obs	8627		8627	
Prob>chi2	0.0000		0.0000	
Pseudo R2	0.0534		0.0538	

P-values are indicated in brackets below every coefficient.
*Significant at 10%
**Significant at 5%
***Significant at 1%.

The analysis of the demographics specification indicates that demographic characteristics indeed have an effect on the probability of preferring a specific candidate over the other. Nevertheless, these explanatory variables are not the only ones that affect candidate preferences, as voting behavior can be influenced by several sources other than demographics, such as campaign events, advertisements, media coverage of candidate debates, candidate political agendas and so on.

6. CONCLUSION

In this study we have tried to examine whether the momentum effects found by Knight and Schiff (2010) in the 2004 Democratic primary appear also in the 2008 Democratic primary. We have adopted the theoretical framework for social learning the aforementioned authors developed in their paper, and we adjusted their econometric model to fit our sample in order to estimate four key social learning parameters with a two-step estimation approach.

The results of the baseline specification suggest that there was no social learning in the 2008 Democratic primary. Specifically, the mean of the initial prior estimated in the first step indicates that Clinton was favored over Obama and Edwards in the pre-primary season. Then, although she underperformed in Iowa while Obama and Edwards outperformed, late voters did not update their preferences appropriately, and Obama did not enjoy momentum effects in the primary season we focused on. The small estimate of the variance in the initial prior also indicates that voters were unresponsive to the release of voting returns.

The results regarding social learning were not altered when we relaxed the assumption of unobserved preferences in the distance specification. We estimated the baseline model including a distance measure that captures a potential observed component of preferences. The results showed that in our case distance did not have a significant effect on estimated preferences. Thus, we cannot conclude that distance captures an observed part of preferences.

Despite the social learning analysis that was the main goal of our study, we also developed an additional specification regarding some sources of voting behavior related to demographics. The results indicate that demographic characteristics such as age, gender, level of education, level of income and region of residence have significant effects on the probability to choose among candidates.

As a final remark, we would like to highlight that there is ample room for further investigation of our analysis in many directions. To be more specific, one extension would be to develop a model of social learning that allows entry and exit of candidates[23]. This way, we would be able to include more periods to our sample allowing some candidates that were viable in the beginning of the primary season to withdraw later. For instance, this could be applied in our baseline specification, where

[23] This extension was proposed by Knight and Schiff (2010). However, we found it instructive to mention it here as it would apply in our case.

we would be able to include Super Tuesday in our sample as well, even though Edwards withdrew from the race just before Super Tuesday. Another extension of our study would be to include a time trend in the baseline model, like Knight and Schiff (2010) did in their paper. This extension would relax the assumption that voter preferences are time independent and would allow late voters to incorporate this trend into their expectations. Finally, an extension regarding our additional specification would be to include media in the model and see whether and how media coverage and advertising affect candidate preferences.

REFERENCES

- Adkins, R.E., and Dowdle, A.J., (2001). "How Important Are Iowa and New Hampshire to Winning Post-Reform Presidential Nominations?". *Political Research Quarterly*, Vol. 54, No. 2, 431-444.
- Ali, S.N., and Kartik, N., (2006). "A Theory of Momentum in Sequential Voting". *UC San Diego, Stanford University*.
- Ali, S.N., and Kartik, N., (2010). "Observational Learning with Collective Preferences". *Working paper, UC San Diego, Columbia University.*
- Banerjee, A.V., (1992). "A Simple Model of Herd Behavior". *The Quarterly Journal of Economics,* Vol. 107, No. 3, 797-817.
- Bartels, L.M., (1987). "Candidate Choice and the Dynamics of the Presidential Nominating Process". *American Journal of Political Science*, Vol. 31, No.1, 1-30.
- Bartels, L.M., (1988). "Presidential Primaries and the Dynamics of Public Choice". *Princeton University Press.*
- BBC, Modern Studies, Political Issues in the UK, "Electoral Systems, Voting and Political Attitudes". http://www.bbc.co.uk/bitesize/higher/modern/uk_gov_politics/elect_vote/revision/2/
- Belleflamme, P., and Peitz, M., (2010). "Industrial Organization Markets and Strategies". *Cambridge University Press*, New York.
- Bikhchandani, S., Hirshleifer, D., and Welch, I., (1992). "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades". *The Journal of Political Economy,* Vol. 100, Issue 5, 992-1026.
- Cai, H., Chen, Y., and Fang, H., (2008). "Observational Learning: Evidence from a Randomized Natural Field Experiment". *Yale University.*
- CNN Election Center 2004, 2008, www.cnn.com.
- Dekel, E., and Piccione, M., (2000). "Sequential Voting Procedures in Symmetric Binary Elections'. *Journal of Political Economy*, Vol. 108, No. 1, 34-55.
- Deltas, G., Herrera, H., and Polborn, M.K., (2010). "A Theory of Learning and Coordination in the Presidential Primary System". *Working paper*.
- Glaeser, E.L., and Sacerdote, B., (2007). "Aggregation Reversals anf the Social Formation of Beliefs". *NBER working paper*.
- Goeree, J.K., Agranov, M., Romero, J., and Yariv, L., (2012). "What Makes Voters Turn Out: The Effects of Polls and Beliefs". *Working paper No. 67, University of Zurich*.
- Hummel, P., and Knight B., (2012). "Sequential or Simultaneous Elections? A Welfare Analysis". *NBER Working Paper*, No. 18076.
- Klumpp, T., and Polborn, M.K., (2005). "Primaries and the New Hampshire Effect". *Working paper*, *Indiana University*, *University of Illinois*.
- Knight, B., and Schiff, N., (2007). "Momentum and Social Learning in Presidential Primaries". *NBER Working paper*, No. 13637.

- Knight, B., and Schiff, N., (2010). "Momentum and Social Learning in Presidential Primaries". *Journal of Political Economy*, Vol. 118, No. 6, 1110-1150.
- Moore, D., McCabe, G., Duckworth, W., and Alwan, L. 2009. "The Practice of Business Statistics", 2nd Edition, *W. H. Freeman and Company, New York.*
- Mordkoff (2000, 2011) http://www2.psychology.uiowa.edu/faculty/mordkoff/GradStats/part%20I/I.07%20normal.pdf
- Prysby, Charles, Carmine Scavo, American Political Science Association, and Inter-university Consortium for Political and Social Research. SETUPS: Voting Behavior: The 2004 Election http://www.icpsr.umich.edu/icpsrweb/instructors/setups/voting.jsp
- Selman, D., (2010). "Optimal Sequencing of Presidential Primaries". *Boğaziçi University, University of Pennsylvania (2008).*
- Smith, L., and Sorensen, P., (2000). "Pathological Outcomes of Observational Learning". *Econometrica*, Vol. 68, No. 2, 371-398.
- Steger, W.P., Dowdle, A.J., and Adkins, R.E., (2004). "The New Hampshire Effect in Presidential Nominations". *Political Research Quarterly*, Vol. 57, No. 3, 375-390.
- Strumpf, K.S., (2002). "Strategic Competition in Sequential Election Contests". *Public Choice*, Vol. 111, Issue 3-4, 377-397.
- Verbeek, M. 2012. "A Guide to Modern Econometrics", 4th edition, *John Wiley &Sons Ltd. Publication*.
- Welch, I., (1992). "Sequential Sales, Learning, and Cascades". *The Journal of Finance*, Vol. 47, No. 2, 695-732.

APPENDIX

TABLE 4A
Candidate Preference on Demographic Characteristics
Multinomial Logit Model - Base Specification

	Edwards			Obama		
	Coefficient/ [p-values]	standard error	z value	Coefficient/[p-values]	standard error	z value
	(1)	(2)	(3)	(4)	(5)	(6)
constant	-0.9162977***	0.3335112	-2.75	-0.9990069***	-0.9990069	-2.86
	[0.006]			[0.004]		
age	0.0004052	0.0019878	0.20	-0.0142715***	0.0018195	-7.84
	[0.838]			[0.000]		
male1	0.4049845***	0.0588963	6.88	0.1628935***	0.0537844	3.03
	[0.000]			[0.002]		
educ2	-0.2442469	0.2931524	-0.83	0.5018916	0.3207005	1.56
	[0.405]			[0.118]		
educ3	-0.1028757	0.2666214	-0.39	0.2401515	0.3053167	0.79
	[0.700]			[0.432]		
educ4	-0.1111432	0.2681155	-0.41	0.6298659**	0.3045723	2.07
	[0.678]			[0.039]		
educ5	-0.1333525	0.2784416	-0.48	0.598483*	0.3113488	1.92
	[0.632]			[0.055]		
educ6	-0.1130776	0.2705933	-0.42	0.8263117***	0.3058466	2.70
	[0.676]			[0.007]		
educ7	-0.1945108	0.2780086	-0.70	0.7517157**	0.3108206	2.42
	[0.484]			[0.016]		
educ8	-0.0311849	0.3389685	-0.09	1.065274***	0.3509158	3.04
	[0.927]			[0.002]		
educ9	-0.0254189	0.3382269	-0.08	0.669464*	0.3656994	1.83
	[0.940]			[0.067]		
inc2	0.0046561	0.2827279	0.02	-0.0466919	0.2484564	-0.19
	[0.987]			[0.851]		
inc3	-0.1327661	0.2975582	-0.45	0.3883384	0.2448025	1.59
	[0.655]			[0.113]		
inc4	-0.1175191	0.2674377	-0.44	0.1296273	0.2355208	0.55
	[0.660]			[0.582]		
inc5	-0.3102425	0.2788567	-1.11	0.2848136	0.2294884	1.24
	[0.266]			[0.215]		
inc6	0.0810844	0.2349034	0.35	0.2847508	0.2062536	1.38
	[0.730]			[0.167]		
inc7	0.1235605	0.2275157	0.54	0.2581087	0.2011134	1.28
	[0.587]			[0.199]		
inc8	0.0361573	0.2264852	0.16	0.0710997	0.2004953	0.35

	[0.873]			[0.723]		
inc9	-0.1230039	0.2264069	-0.54	0.192723	0.1959719	0.98
	[0.587]			[0.325]		
inc10	0.0304495	0.2200042	0.14	0.2415132	0.1924208	1.26
	[0.890]			[0.209]		
inc11	0.1507545	0.2116629	0.71	0.3481389*	0.1848167	1.88
	[0.476]			[0.060]		
inc12	0.0507012	0.2165851	0.23	0.2286244	0.1901745	1.20
	[0.815]			[0.229]		
inc13	0.0948956	0.2127488	0.45	0.2487476	0.1862903	1.34
	[0.656]			[0.182]		
inc14	0.1035862	0.2277966	0.45	0.2753395	0.1986035	1.39
	[0.649]			[0.166]		
inc15	0.4016568*	0.2280749	1.76	0.5000968**	0.2010978	2.49
	[0.078]			[0.013]		
inc16	0.0883189	0.2314464	0.38	0.41867**	0.2006981	2.09
	[0.703]			[0.037]		
inc17	-0.1208997	0.2662603	-0.45	0.1554056	0.2273659	0.68
	[0.650]			[0.494]		
inc18	0.0017687	0.3119672	0.01	0.2905067	0.2647747	1.10
	[0.995]			[0.273]		
inc19	0.3780689	0.2810042	1.35	0.6074255**	0.2461913	2.47
	[0.178]			[0.014]		
ethnic2	-1.434537***	0.1400094	-10.25	1.029921***	0.066223	15.55
	[0.000]			[0.000]		
ethnic3	-1.20583***	0.2375407	-5.08	-0.3803483**	0.1616395	-2.35
	[0.000]			[0.019]		
ethnic4	-0.9047519***	0.1340831	-6.75	-0.2586076**	0.1029607	-2.51
	[0.000]			[0.012]		
ethnic5	-0.2356955	0.198341	-1.19	0.122028	0.1755003	0.70
	[0.235]			[0.487]		
Number of obs	8627					
LR chi2 (64)	937.99					
Prob>chi2	0.0000					
Pseudo R2	0.0534					
Log likelihood	-8316.5453					

TABLE 4B

Candidate Preference on Demographic Characteristics

Multinomial Logit Model - Extended Specification (incl. Met. Area)

	Edwards			Obama		
	Coefficient/ [p-values]	standard error	z value	Coefficient/[p-values]	standard error	z value
	(1)	(2)	(3)	(4)	(5)	(6)
constant	-0.7482644** [0.028]	0.3396671	-2.20	-0.988078*** [0.005]	0.355738	-2.78
age	0.0004886 [0.806]	0.0019907	0.25	-0.0142525*** [0.000]	0.0018189	-7.84
male1	0.4053796*** [0.000]	0.0589356	6.88	0.1629428*** [0.002]	0.0537829	3.03
educ2	-0.2354872 [0.422]	0.2934135	-0.80	0.5028281 [0.117]	0.3207064	1.57
educ3	-0.0904235 [0.735]	0.2669177	-0.34	0.2412234 [0.430]	0.3053448	0.79
educ4	-0.0869508 [0.746]	0.2685214	-0.32	0.6317858** [0.038]	0.3046665	2.07
educ5	-0.1120988 [0.688]	0.2788045	-0.40	0.6001634* [0.054]	0.3114147	1.93
educ6	-0.0802286 [0.767]	0.271127	-0.30	0.8286518*** [0.007]	0.3060276	2.71
educ7	-0.1665005 [0.550]	0.2784409	-0.60	0.7535996** [0.015]	0.3109448	2.42
educ8	0.0007392 [0.998]	0.3394182	0.00	1.067482*** [0.002]	0.3510716	3.04
educ9	-0.0017392 [0.996]	0.3386168	-0.01	0.6708895* [0.067]	0.3657997	1.83
inc2	0.0066292 [0.981]	0.2828053	0.02	-0.0466309 [0.851]	0.2484533	-0.19
inc3	-0.1625512 [0.585]	0.2978764	-0.55	0.3858378 [0.115]	0.2449333	1.58
inc4	-0.1334688 [0.618]	0.2676826	-0.50	0.1287454 [0.585]	0.2355746	0.55
inc5	-0.3246134 [0.245]	0.2790168	-1.16	0.2838351 [0.216]	0.229521	1.24
inc6	0.0642178 [0.785]	0.2350575	0.27	0.283833 [0.169]	0.2063387	1.38
inc7	0.1061169 [0.641]	0.2277092	0.47	0.256878 [0.202]	0.201199	1.28
inc8	0.0255471 [0.910]	0.226634	0.11	0.0705763 [0.725]	0.2005315	0.35
inc9	-0.1356664 [0.549]	0.2265605	-0.60	0.1922487 [0.327]	0.1960149	0.98
inc10	0.0218953	0.2201222	0.10	0.2412783	0.1924402	1.25

	[0.921]			[0.210]		
inc11	0.1474419	0.2117434	0.70	0.3479381*	0.1848228	1.88
	[0.486]			[0.060]		
inc12	0.046236	0.2166819	0.21	0.2283401	0.1901792	1.20
	[0.831]			[0.230]		
inc13	0.0924227	0.2128304	0.43	0.2484639	0.186296	1.33
	[0.664]			[0.182]		
inc14	0.1020236	0.2278751	0.45	0.2752423	0.1986053	1.39
	[0.654]			[0.166]		
inc15	0.4049695*	0.2281761	1.77	0.5004819**	0.2011036	2.49
	[0.076]			[0.013]		
inc16	0.0984163	0.231552	0.43	0.4194905**	0.2007232	2.09
	[0.671]			[0.037]		
inc17	-0.1071008	0.2663607	-0.40	0.1563257	0.2274072	0.69
	[0.688]			[0.492]		
inc18	0.0157919	0.3120963	0.05	0.2915958	0.2648215	1.10
	[0.960]			[0.271]		
inc19	0.3918034	0.2811269	1.39	0.6086095**	0.2462397	2.47
	[0.163]			[0.013]		
ethnic2	-1.41429***	0.1402342	-10.1	1.031285***	0.0665808	15.49
	[0.000]			[0.000]		
ethnic3	-1.199309***	0.237603	-5.05	-0.3798852**	0.16167	-2.35
	[0.000]			[0.019]		
ethnic4	-0.8856702***	0.1342813	-6.60	-0.2573201**	0.1031492	-2.49
	[0.000]			[0.013]		
ethnic5	-0.2256224	0.1984663	-1.14	0.1227191	0.1755531	0.70
	[0.256]			[0.485]		
m.area1	-0.2234306***	0.0833648	-2.68	-0.0154965	0.0841303	-0.18
	[0.007]			[0.854]		
Number of obs	8627					
LR chi2 (66)	945.44					
Prob>chi2	0.0000					
Pseudo R2	0.0538					
Log likelihood	-83.128.215					

TABLE 5

ELECTION DATES AND TIME PERIODS OF PRIMARY

5 Periods			4 Periods		
State	Period	Date	State	Period	Date
IA	**1**	**3/1/2008**	**IA**	**1**	**3/1/2008**
NH	**2**	**8/1/2008**	**NH**	**2**	**8/1/2008**
MI		15/1/2008	MI		15/1/2008
NV	**3**	**19/1/2008**	**NV**	**3**	**19/1/2008**
SC	**4**	**26/1/2008**	**SC**	**4**	**26/1/2008**
FL	**5**	**29/1/2008**	FL		29/1/2008