Forecasting the 2012 US Presidential Election: Lessons from a State-by-State Political Economy Model

Article *in* Political Science and Politics · October 2012

DOI: 10.1017/S1049096512000972

CITATIONS
11

READS
51

2 authors:

Bruno Jérôme
Université Panthéon-Assas Paris 2
44 PUBLICATIONS **308** CITATIONS
SEE PROFILE

Veronique Jerome
Université Paris-Sud 11
27 PUBLICATIONS **215** CITATIONS
SEE PROFILE

SYMPOSIUM

Forecasting the 2012 US Presidential Election: Lessons from a State-by-State Political Economy Model

Bruno Jerôme, *University of Paris 2*
Véronique Jerôme-Speziari, *University of Paris Sud 11*

Since 2008, the economic fallout from the subprime mortgage crisis has led to the defeat of a number of incumbents in the world's major democracies. For instance, in the former EU-15, eight countries (including France) have ousted their incumbents in favor of new leaders.[1] The United States is no exception, and the 2012 US presidential election will see Barack Obama running for a second term during difficult economic times. After hitting a high of 10% in October 2009, the nation's unemployment rate decreased to 8.2% in May 2012. Nonetheless, this is still 0.7 percentage point higher than what Ronald Reagan faced in 1984 or what confronted George H.W. Bush in 1992 as they ran for their second terms. Looking at measures of presidential popularity for the month of May since 1980, Barack Obama's approval rating is at 46% in the Gallup polls, which is the third-worst rating after George W. Bush (30% in 2008) and George H.W. Bush (39.4% in 1992). Given Barack Obama's approval rating and the current national unemployment level, must we conclude that Barack Obama is irremediably on the ropes against Mitt Romney in 2012?

To answer this question, we developed a forecasting model based on local and national data, whose values were known well before the election. By obtaining electoral votes (EVs) from simulated popular votes at the state level, this methodology imitates the indirect process of the Electoral College used in the US presidential election. This state-by-state political economy model predicts that Barack Obama will be reelected for a second term, despite an unfavorable economic context. However, our main scenario is supplemented by two additional forecasts based on particular determinants whose values are not, or cannot be, fully captured by our explanatory variables.

First, the data used in this article (written at the end of June 2012) were available in the second quarter of 2012. This poses a small problem because an optimal forecast would normally be made one quarter before the presidential election. Therefore, we ask what the politicoeconomic situation could be like in September 2012 by simulating what would happen if both unemployment and Barack Obama's popularity were to worsen in the third quarter of 2012.

Second, currently we do not know to what extent the next presidential election will be used as an "anti-Obama" referendum by voters, similar to what happened in France to Nicolas Sarkozy. To make this scenario realistic, we have imputed the most unfavorable error margin to the incumbent's forecasted votes obtained from our main scenario.

As a result, in addition to forecasting who could be the next president of the United States, the structure of the state-by-state model enables us to test variations by highlighting the propensity of certain states to be strongholds or swing states.

FORECASTING PRESIDENTIAL ELECTIONS IN THE UNITED STATES

In "The State of Presidential Election Forecasting," R.J. Jones (2008) provides an exhaustive review of 13 methods that have been previously used to predict US presidential elections. According to Michael Lewis-Beck (2005), unlike others, only statistical estimation (forecasting) methods using economic and sociopolitical data are based on a real theory of voting.

Sigelman (1979), Lewis-Beck and Rice (1982), Rosenstone (1983), and Lewis-Beck and Rice (1984) built the first statistical political economy models aimed at forecasting US elections. Among these, Rosenstone's pooling model, whose unit of study is the state, is exceptional. Following Rosenstone, two more pooling models were built: Holbrook's in 1991 and Campbell's in 1992. However, state-based models are still neglected; in the last 2008 *Pollyvote* survey, aggregated models were still dominant.

Our political economy model (PEM) is part of this state-based "family" but adds partisan dynamics at the state level, similar to the first political economy pooled time series models built for the 2002 French presidential election (Jérôme and Jérôme-Speziari 2001; Jérôme, Jérôme-Speziari, and Lewis-Beck 2003).

Our PEM uses data for subnational units and then aggregates these subnational predictions to create an overall prediction. Using subnational data, our forecasting approach differs from all other models in the US forecasting literature (exceptions are previously mentioned). Our approach has the potential to substantially improve predictive accuracy insofar as local context matters for shaping beliefs about the economy and past performance of political actors.

First, we start with a search for the most significant explanatory variables to be included in our model. Next, we estimate

doi:10.1017/S1049096512000972

voting patterns at the state level; then we switch to aggregate state-by-state predictions in terms of Electoral College vote share to predict the winner of the 2012 US presidential election.

WHY TAKE INTO ACCOUNT POLITICAL AND ECONOMIC DATA AT THE STATE LEVEL?

Pooling data obviously helps increase the number of observations in the dataset as well as the degrees of freedom; thus, we can expect the model to perform better. A larger number of observations also guarantees a minimization of biases due to aggregation effects. In addition, we use the so-called saving properties of aggregation so errors could be balanced. Pooling also allows for research into the "dynamics of change" (Baltagli 1995). It also allows us to discern and to measure effects that are difficult to identify when either purely cross-sectional or purely longitudinal methods are used in isolation from each other. Finally, this approach allows us to come closer to studying behavior on the individual level.

In the United States, voting behavior has a territorial component. Because the US presidential election is normally decided by the Electoral College vote, there is a journalistic emphasis on studying the expected voting behavior of the "swing states." American politician Tip O'Neill used to say, "All politics is local." If local economic and social conditions, affect voters' electoral choices in a given geographic area, then we can improve our predictive accuracy at the national level by using pooled time series voting models that make use of state and regional data.

Sometimes long-term consistencies exist in voting behavior in particular geographic regions, but these are rarely immutable. Furthermore, local effects, like national effects, are triggered by factors such as changes in economic conditions, and, in the extreme, changes in local economic conditions can lead to outward or inward migration (Tiebout 1956). Some states are stable in their ideology, whereas others are very unstable. Given these state-by-state idiosyncrasies, our vote function must reflect the evolution of "partisan domination" through time and space. To achieve this, we built two partisan representation indexes reflecting states' partisan particularities. For this purpose, we distinguish between the entire period under study (1952–2008) and the period 1980–2008 when significant partisan changes occurred in US politics.

From 1952 to 2008, Republicans were dominant in 12 out of 50 states (plus Washington, DC) (Alaska, Arizona, Idaho, Indiana, Kansas, Montana, Nebraska, North Dakota, Oklahoma, South Dakota, Virgina, and Wyoming); this means that Republicans have a long-term electoral base of 72 electoral votes (EVs), or 13.4% of the total EVs. If we perform the exact same calculation for Democrats,[2] we find a long-term electoral base of 43 EVs, or 8% of the total EV (Hawaii, Maryland, Massachusetts, Minnesota, Rhode Island, and Washington, DC). These unassailable states are considered to be electoral strongholds for each party.

However, some states are "swing states" or states that realign over time (see Merrill, Grofman, and Brunell 2008), such that both Democrats and Republicans profit from what we call a "new domination" since 1980. For Democrats, this potential new domination can be seen in California, Connecticut, Delaware, Iowa, Illinois, Maine, Michigan, New Jersey, New York, Oregon, Pennsylvania, Vermont, Washington, and Wisconsin, which translates into 213 EVs, or 39% of the total EVs. For Republicans, their potential new domination may appear in Alabama, Arizona, Georgia, Kentucky, Louisiana, Mississippi, Missouri, North Carolina, South Carolina, Tennessee, Texas, and Utah, which amounts to 137 EVs, or 25.5% of the total EVs. We must highlight that the Republicans' gains are mostly in the southern states (except Missouri and Utah), where they have captured "southern Democrats." Finally, six

Table 1

2012 US Presidential Election Political Economy Model Pooled Time Series 50 States and DC (1980–2008)

DEPENDENT VARIABLE		
Incumbent Vote Share at Presidential Election (INCV)		
INDEPENDENT VARIABLES	**COEF.**	**T-STAT.**
Constant	39.71*	31.66
Unemployment Change (BU)	−0.56*	−3.59
President's Job Approval (PJA)		
PJA2	0.28*	11.19
PJA0	0.19*	7.03
Local Partisan Domination (LPD)		
LPD5208	8.52*	17.17
LPD8008[a]	5.23*	11.07
For the above variables see Appendix		
INDV	−0.75*	−21.46
GA80	11.34*	2.51
AR92	−11.83*	−2.63
AR96	7.53**	1.67
HI-ILO8	−12.17*	−3.77
RHSC	6.53*	6.18
DHSC	3.68*	2.45
DLSC	−10.58*	−7.71
DCDS	24.58*	9.29
DCRS	−31.19*	−14.96
UTO8	−8.72*	−1.90
NV80	−17.28*	−3.84
VTO4	−11.75*	−2.61
R-squared	0.85	
Adjusted R-squared	0.84	
S.E.R	4.47	
N	408	

[a]Including Southern Republican new Strongholds

*Significance at a 1% level (two-tails)

**Significance at a 5% level (two-tails)

Table 2

Ex-Post Forecasts Accuracy

YEAR	INCUMBENT	CORRECT PREDICTION	POPULAR VOTE (%)			ELECTORAL VOTE		
			Forecast	Actual	Error (F – A)	Forecast	Actual	Error (F – A)
1980	D	Yes	45.82	41	4.82	32	49	–17
1984	R	Yes	59.76	59	0.75	535	525	10
1988	R	Yes	51.36	53.4	–2.04	294	426	–132
1992	R	Yes	37.91	37.5	0.41	231	168	63
1996	D	Yes	49.66	49.24	0.42	319	379	–60
2000	D	No	49.83	48.41	1.42	304	266	38
2004	R	Yes	51.34	50.7	0.64	299	286	13
2008	R	Yes	46.7	45.66	1.04	163	173	–10
2012	D	?	51.6	?	?	324	?	?

states remain without strong partisan domination, and they represent 48 EVs, or 9% of the total EVs.

A STATE-BY-STATE POLITICAL ECONOMY FORECASTING MODEL FOR US PRESIDENTIAL ELECTIONS

Models built on vote functions are based on a theory of governmental responsibility (Key 1966). Vote functions offer an explicative model of voters' choices and preferences and provide a model of votes for an incumbent based on political and economic performances. Vote functions need to include "good" political and economic variables side-by-side to gain stability (Nannestad and Paldam 1994). In addition, pure economic vote functions (see Fair 1978) are regularly underspecified and lead to significant errors when these are used in electoral forecasting. Thus, Rosenstone (1983, 45) argues that an aggregated voting model that keeps only one of the above-mentioned possibilities could not explain why citizens change their vote from one election to the other.

Next, we include the structural model from which we compute the ordinary least squares (OLS) regression estimates, along with the operationalization of the variables from 1980–2008 in the 50 American states plus Washington, DC ($N = 408$):

$$\begin{aligned} \mathrm{INCV}_{i,t} = {} & \mathrm{C} + \Delta\mathrm{U}_{i,t-n} + \mathrm{PJA}_{t-n} + \mathrm{LPD}_{i,t-n} \\ & + \text{Politics and Institutions}_{i,t-n} \\ & + \text{President's Local Strongholds}_{i,t-n} \\ & + \text{Local Electoral Peculiarities}_{i,t-n} + \text{Error} \end{aligned}$$

The dependent variable $\mathrm{INCV}_{i,t}$ measures the vote share in the ith state ($i = 1,\ldots,51$) and for the tth time period ($t =$ 1980, 1984, …, 2008) going to the ruling political party. The incumbent president was a Democrat in 1980, 1996, and 2000 and was a Republican in 1984, 1988, 1992, 2004, and 2008.

For the explanatory variables, the first variable (ΔU) is the change in the local (i.e., state-level) unemployment rate from the month after the president was elected to the month prior to the next presidential election. A positive change in state unemployment should cost the incumbent a portion of his vote share. The second variable is the Gallup President's Job Approval (PJA), at a national level, six months before the election. For the outgoing majority candidate, the higher the president's popularity is, the higher the electoral premium. However, the impact of popularity should not be the same and depends on whether the incumbent is seeking a second term or not. Thus, PJA2 is the president's popularity when he is seeking a second term (and 0 otherwise). PJA0 is the President's Job Approval when the incumbent is not running for a second term. Following the above discussion, the third variable, Local Partisan Domination (LPD), takes into account the characteristics of the partisan cycle in each state. LPD is divided into two local partisan domination variables, one named LPD5208 to code states having significant partisan domination since 1952 and another one called LPD8008, or "new domination" since 1980, which includes recent southern Republican strongholds. More precisely, for each state during the 1952–2008 period, LPD5208 gives the rate of success for each party when this rate exceeded 85%[3] for the Republicans and 61% for the Democrats (this variable takes a 0 value otherwise). When the ideology of the state is the same as the incumbent, it carries a "+" sign; otherwise, it carries a "–" sign. This means that an incumbent will be rewarded by states that are ideologically similar and punished by states that are ideologically dissimilar. The LPD8008 variable was constructed in the same way, except that the rate of success is calculated for 1980–2008.

Some other explanatory variables used to build the econometric model deal with politics and institutions (i.e., the electoral weight of independent candidates), the president's electoral strongholds (i.e., measuring home state advantage), and local electoral peculiarities such as scores usually deviating from the standard (Democrats in Washington, DC, for instance). The specification of these variables whose estimated coefficients are not used to forecast the 2012 election is described in the appendix.

EMPIRICAL RESULTS

Overall, if one focuses on the main variables, the aforementioned US presidential vote model gives good statistical results (see table 1). This model explains 85% of the variance, and all the coefficients are statistically significant (t-ratios at 0.05, two-tailed) except AR96 t-ratio at 0.10.

The standard error of regression (SER) is 4.47. This means that each estimated score in each state has to be interpreted with a ±4.47 error margin. Furthermore, all of the coefficients for the explanatory variables in the presidential vote function (for incumbents) have the expected signs.

Among the variables kept for forecasting, we underline the fact that the constant provides an estimate for the incumbent's electoral base of 39.71% of the long-run vote share. On average, a one-point rise in the unemployment rate costs the incumbent 0.56% of the votes in a given state. In case of a second term, a 50% approval score yields 14% more votes. However, in the case of new candidate (from the incumbent's same party), a popularity score of 50% for the president yields only 9.50% more votes.

The outgoing majority on average gains 8.52%, depending on the party's historical strength in a particular state (LPD5208). Conversely, in such strongholds, the opposition candidate suffers a 8.52% electoral disadvantage. In what we call "new partisan domination" states (LPD8008) measured between 1980 and 2008, the outgoing majority on average gains 5.23% in their strongholds, while the opposition candidate suffers a 5.23% electoral loss.

To see if we can trust this model, we can test it through electoral "backcasting" for previous elections. The observed and forecasted aggregated electoral results for the incumbent are displayed in table 2. Except for 2000, both electoral and popular votes are adequately predicted and the winner is correctly determined.

FORECASTING, SCENARIOS, AND DISCUSSION

The 2012 preelection values have been plugged into the model to predict the popular votes in each state. Table 3 shows the incumbent state-by-state vote share forecasts according to two scenarios. The first one, our main scenario (column 1), is based on economic (change in unemployment) and political (Obama's approval job rate) factors available in June 2012. Our second scenario (column 2) hypothesizes that Barack Obama may be threatened by an "anti-Obama" referendum similar to what befell Nicolas Sarkozy in France. This specific case aims at measuring indirectly what would happen if concerns not accounted for in our model, such as

Table 3

State-by-State Political Economy Model 2012 Forecast (Sept. 2012) Popular Votes and Electoral Votes for Democrats

SCENARIOS	MAIN (1)			ALTERNATIVE (2) ERROR MARGIN = –4.47		
States + Dist. Col.	PV[a] DEM	EV[b] DEM	EV GOP	PV DEM	EV DEM	EV GOP
AL	48.1		9	43.6		9
AK	35.1		3	30.6		3
AZ	44.2		11	39.7		11
AR	48.5		6	44.1		6
CA	55.8	55		51.3	55	
CO	52.3	9		47.9		9
CT	56.0	7		51.6	7	
DE	56.2	3		51.7	3	
FL	52.8	29		48.3		29
GA	48.9		16	44.4		16
HI	60.1	4		55.6	4	
ID	33.8		4	29.3		4
IL	52.5	20		48.0		20
IN	46.2		11	41.7		11
IA	56.0	6		51.6	6	
KS	44.8		6	40.4		6
KY	49.2		8	44.7		8
LA	48.0		8	43.5		8
ME	56.6	4		52.1	4	
MD	62.2	10		57.7	10	
MA	63.5	11		59.0	11	
MI	58.0	16		53.6	16	
MN	60.9	10		56.4	10	
MS	47.6		6	43.2		6
MO	49.4		10	45.0		10
MT	45.3		3	40.8		3
NE	45.2		5	40.7		5
NV	51.5	6		47.1		6
NH	53.0	4		48.5		4
NJ	59.0	14		54.5	14	
NM	51.8	5		47.4		5
NY	59.8	29		55.3	29	
NC	48.4		15	44.0		15
ND	45.8		3	41.3		3
OH	53.8	18		49.3		18
OK	44.8		7	40.4		7
OR	57.7	7		53.2	7	
PA	52.7	20		48.3		20

(continued)

Table 3 (Continued)

SCENARIOS	MAIN (1)			ALTERNATIVE (2) ERROR MARGIN = −4.47		
States + Dist. Col.	PV[a] DEM	EV[b] DEM	EV GOP	PV DEM	EV DEM	EV GOP
RI	62.5	4		58.0	4	
SC	48.7		9	44.2		9
SD	45.1		3	40.7		3
TN	49.5		11	45.0		11
TX	47.5		38	43.0		38
UT	36.5		6	32.1		6
VT	57.4	3		52.9	3	
VA	45.9		13	41.4		13
WA	56.7	12		52.2	12	
WV	52.1	5		47.6		5
WI	57.0	10		52.5	10	
WY	33.6		3	29.1		3
DC	85.8	3		81.4	3	
Maj. = 270	51.6	324	214	47.2	208	330

[a]PV: Popular Vote; [b]EV: Electoral Vote

the unpopularity of health-care reform or a "racial cost" (see Lewis-Beck, Tien, and Nadeau 2010), were to have a deep and negative impact on Obama's vote share. To simulate this, we have imputed a −4.47 percentage points error margin to the incumbent's forecasted votes issued from our main scenario.

With regard to the main scenario, the forecasting model (launched to predict 2004 and 2008 elections) gives an absolute majority for the Democrats with 51.6% of the popular votes (−1.2% compared with 2008) and 324 EVs (−39 compared with 2008). This simulation shows a slight increase in popular vote (+0.4) from our previous forecast calculated in August 2011. However, one can observe a near status quo regarding the electoral vote.

Thus, if the elections were held in September 2012, the Democrats would obtain a majority of the popular votes in 27 states (−2 compared with 2008). According to the forecast, the Republicans would regain two of its southern strongholds that it lost in 2008: North Carolina (15 EVs) and Virginia (13 EVs). To this, we can also add a victory in Indiana (11 EVs). However, Republicans would lose West Virginia (5 EVs).

How do we explain this result? Basically, after peaking at 10% in October 2009, the unemployment rate has been decreasing and has stabilized at around 8.2% in May and June (compared to 7.8% when Obama took office). This evolution enables the Democrats to resist in most of their historical and "new" strongholds. In addition, the incumbent president's credibility is still rather high (46.4% were satisfied with his performance in March 2012, Gallup *USA Today*), compared with George W. Bush's approval rating in March 2007 (when 33% were satisfied). Translating this into votes, this means on average an electoral gain of 3.9[4] points per state compared with George W. Bush's situation at the same time.

However, if we uniformly apply a −4.47 point error margin to Democrats in each state to measure factors not accounted for in our explanatory variables, we find that Mitt Romney could pull off a backdoor victory in the 2012 election with 330 EVs and 52.8% of the vote share. First, this result shows that Obama's lead is rather fragile if an "anything but Obama" strategy dominates. Second, this simulation shows that nine states, Colorado, Florida, Illinois, Nevada, New Hampshire, New Mexico, Ohio, Pennsylvania, and West Virginia, could be decisive. Among them, four states (Florida, Illinois, Ohio, and Pennsylvania) account for 87 EVs.

If the likelihood of an "anti-Obama" referendum turns out to be smaller than expected, is there a baseline in terms of popularity and unemployment to guarantee a second term for the incumbent in November? Ceteris paribus, according to our calculations, if the unemployment rate exceeds 9.4% in September on the national level (e.g., a + 1.2 point change in unemployment on average) and if Barack Obama's job approval falls to 39%, then Mitt Romney would mathematically win 274 EVs and 51.1% of the popular vote. Moreover, the Republicans would win Colorado, Illinois, Nevada, New Mexico, Pennsylvania, and West Virginia.

Now a question arises: with five months remaining until the next election, will the European debt crisis and weakening growth worldwide significantly change Barack Obama's chances of victory? Notwithstanding extenuating circumstances or external and unforeseen shocks, one can doubt it, and the Democrats should have a second term in the White House. However, as our state-by-state model shows, the Democrats will have to keep an eye on Florida, Illinois, Ohio, and Pennsylvania to maximize their chances of reelection. ■

NOTES

1. The others are Denmark, Greece, Italy, Ireland, Portugal, Spain, and the United Kingdom.
2. Here the threshold giving the rate of electoral success has to be above 61% (average = 38%; standard error = 23%).
3. The threshold victory threshold is based on the last 15 presidential elections and is calculated in the following way: 85% = 62% (average rate of Republican victories over 1958–2008) + 23% (standard error). For the Democrats: 61% = 38% (average rate of Democrat victories over 1958–2008) + 23% (standard error).
4. $(46.4 - 33) \times 0.29 = -3.9$

REFERENCES

Baltagli, Badi H. 1995. *Econometric Analysis of Panel Data.* New York: Wiley.

Campbell, James E. 1992. "Forecasting the Presidential Vote in the States." *American Journal of Political Science* 36: 386–407.

Fair, Ray C. 1978. "The Effect of Economic Events on Votes for President." *Review of Economics and Statistics* 64: 159–72.

Holbrook, Thomas M. 1991. "Presidential Elections in Space and Time." *American Journal of Political Science* 35: 91–109.

Jérôme, Bruno, and Véronique Jérôme-Speziari. 2001. "La France réélirait Chirac." In *Ils ont trouvé la formule qui donne le résultat des élections*, *L'Expansion* 656, 8–21 November: 74–75.

Jérôme, Bruno, Véronique Jérôme-Speziari, and Michael S. Lewis-Beck. 2003. "Reordering the French Election Calendar: Forecasting the Consequences for 2002." *European Journal of Political Research* 423: 425–40.

Jones, Randall J. 2008. "The State of Presidential Election Forecasting: The 2004 Experience." *The International Journal of Forecasting* 24 (2): 310–21.

Key, Valdimer. O. 1966. *The Responsible Electorate*. New York: Vintage.

Lewis-Beck, Michael S. 2005. "Election Forecasting: Principles and Practice." *British Journal of Politics and International Relations* 7: 145–64.

Lewis-Beck, Michael S., and Martin Paldam. 2000. "Economic Voting: An Introduction." *Electoral Studies* 19 (2/3): 113–22.

Lewis-Beck, Michael S., and Tom W. Rice. 1982. "Presidential Popularity and Presidential Vote." *Opinion Quarterly* 46: 534–37.

———. 1984. "Forecasting Presidential Elections: A Comparison of Naive Models." *Political Behavior* 61: 9–21.

Lewis-Beck, Michael S., Charles Tien, and Richard Nadeau. 2010. "Obama's Missed Landslide: A Racial Cost?" *PS: Political Science & Politics* 43 (1): 69–76.

Merrill, Samuel, Bernard Grofman, and Thomas L. Brunell. 2008. "Cycles in American National Electoral Politics, 1854–2006: Statistical Evidence and an Explanatory Model." *American Journal of Political Science* 102 (1): 1–17.

Nannestad, Peter, and Martin Paldam. 1994. "The VP-function: A Survey of the Literature on Vote and Popularity Functions after 25 years." *Public Choice* 79 (3/4): 213–45.

Rosenstone, Steven J. 1983. *Forecasting Presidential Elections*. New Haven, CT: Yale University Press.

Sigelman, Lee. 1979. "Presidential Popularity and Presidential Elections." *Public Opinion Quarterly* 43: 532–34.

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64: 416–24.

APPENDIX: Specification of Additional Explanatory Variables

INDV: vote share for the Independent or Green candidates who had a real "nuisance" power for the incumbents (0 otherwise). [John Anderson Independent candidate in 1980, Ross Perot in 1992 and 1996, and Ralph Nader in 2000]. This variable shows that one point gain in votes for Independent/Green candidates costs 0.75% points to the incumbent (whichever party he may be from).

Dummy variables scored 1 (0 otherwise) measuring significant home state advantage: **GA80** Georgia in 1980 for Jimmy Carter, **AR92** Arkansas in 1992 for Bill Clinton as an opponent, **AR96** Arkansas in 1996 for Bill Clinton as an incumbent, **HI-IL08** Hawaii and Illinois for Barack Obama in 2008.

RHSC and **DHSC**: dummy variables scored 1 (0 otherwise) to take into account states where Republicans (R) or Democrats (D) have systematically high scores.

DLSC: dummy variable scored 1 to code states where Democrats have systematically low scores (0 otherwise). *This variable is not significant for Republicans*.

DCDS and **DCRS**: dummy variable scored 1 (0 otherwise) in District of Columbia when Democrats (D) or Republicans (R) are incumbents.

UT08: in 2008 in the Republican stronghold of Utah, McCain competed against five candidates who gathered almost 24% of the total vote share. This variable is scored 1 in 2008 (0 otherwise).

NV80 and **VT04** are outliers: dummy variables scored 1 (0 otherwise) for Nevada in 1980 and Vermont in 2004.