State-Level Forecasts for the 2016 US Presidential Elections: Political Economy Model Predicts Hillary Clinton Victory

Bruno Jerôme, *University of Paris, II Panthéon-Assas*
Véronique Jerôme-Speziari, *University of Paris-Sud Saclay*

Is Hillary Clinton able to win a third term for the Democrats in November 2016? Since 1952, Democrats and Republicans have each been successful three times in winning two consecutive terms. However, Ronald Reagan and George H.W. Bush were the only presidents to win three consecutive presidential terms for the Republicans (1981–1993).

Each electoral victory is based on a combination of several factors for which positive determinants outweigh the negative ones. In terms of positives, the current economic conjuncture is favorable to Hillary Clinton. With the country's unemployment rate at 8.2% when Obama started his second term in January 2013, the nation's unemployment rate has decreased to 4.7% in May 2016. In addition, she is helped by Obama's approval rating, which is at 52% as of April 2016 (Gallup). This is still six percentage points higher than what Barack Obama had in 2012. Finally, Hillary Clinton is facing a divided opposition. Her opponent, Donald Trump, is the least-supported Republican nominee since 1980, with an average score of 44.5% among Republicans. In comparison, Mitt Romney won 50.1% of the GOP popular vote (2012), George W. Bush 65.2% (2000), Bob Dole 52.2% (1996) and Ronald Reagan 62.1% (1980).

While the planets seem aligned for Hillary Clinton, the bad news is the Democrats are running for a third consecutive term. Therefore, there is a risk of fatigue on the part of the electorate. Norpoth (2014) explains this phenomenon in terms of cycles; the risk of defeat is at its highest after two (possibly three) terms. Consequently, the competition could be more balanced than expected between the GOP and Democrats, at least with regard to the popular vote (PV).

The electoral vote (EV) has its particular logic since each state follows its own ideological pattern and has a high propensity to be either a stronghold or a swing state. In order to forecast who could be the next President of the United States, we developed a state-by-state political economy model based on local and national data, whose values were known one quarter before the election. After simulating popular votes at the state level, electoral votes (EVs) are deduced for the incumbent party and then aggregated at a national level. As a result, our methodology imitates the indirect process of the Electoral College used in the US presidential election. This state-by-state political economy model predicts that Hillary Clinton could be elected as president of the United States, despite a close race in the PV.

Our approach goes as follow. First, for each state, we calculated the probability that the Democrats will win an absolute majority of the votes. This is a departure from our previous forecast in 2012 (Jérôme and Jérôme-Speziari 2012). Second, to test the robustness of the aggregated forecast, we have cumulated the predicted votes whose probability was exceeding 73% to reach an absolute majority. This leads to classify the states according to whether they are weak, strong, or toss-ups for each camp.

The data used here are those available in the second quarter of 2016.[1] An optimal forecast would normally be made one quarter before the presidential election. However, the state-by-state probability of victory allows us to assume that the Democrats are leading significantly enough in terms of EVs to ensure Clinton's success.

FORECASTING US PRESIDENTIAL ELECTIONS

According to Lewis-Beck and Stegmaier (2013), the 2012 US presidential elections were quite challenging for various forecasting approaches. With the increased demand in forecasts from not only campaigns but also citizens, the supply of forecasts available can be roughly classified into four specific groups: structuralists, aggregators, synthesizers[2] and judges.

Using economic and socio-political data, the seminal works in forecasting US elections (Sigelman 1979; Lewis-Beck and Rice 1982; Rosenstone 1983 and Lewis-Beck and Rice 1984) were mainly structural and consequently based on a real theory of voting (Lewis-Beck 2005). Among them, Rosenstone (1983), Holbrook (1991), and Campbell (1992) build pooling models, whose unit of study is the state, but their approach remained an exception in the literature. State-based models are still neglected. For instance, in the last 2016 *Pollyvote* survey, econometric aggregated models were clearly dominant.

Fundamentally structural, our political economy model (PEM) is also part of this state-based "family"[3] of forecasts,

doi:10.1017/S1049096516001311

but takes into account partisan patterns at the state level, following our first political economy pooled time series models built for the 2002 French presidential elections (Jérôme and Jérôme-Speziari 2001; Jérôme, Jérôme-Speziari, and Lewis-Beck 2003).

Our PEM uses data disaggregated below the national level and then generates sub-national predictions before aggregating them to create an overall prediction. Accordingly, our forecasting approach differs from almost all other models currently forecasting US elections. Our approach has shown its accuracy by correctly predicting the two-party result for both PV and EV in 2004, 2008, and 2012.[4]

Our PEM uses data disaggregated below the national level and then generates sub-national predictions before aggregating them to create an overall prediction. Accordingly, our forecasting approach differs from almost all other models currently forecasting US elections.

WHY FORECAST US ELECTIONS WITH POOLED TIME-SERIES MODELS?

We can expect the model to perform better insofar as pooling data clearly helps increase both the number of observations in the dataset and the degrees of freedom. A larger number of observations helps minimize biases due to aggregation effects. In addition, we expect benefits from the "properties of aggregation" in that errors can be balanced out. According to Baltagli (1995), a pooling process also allows for research into the "dynamics of change." Indeed, purely cross-sectional or purely longitudinal methods are unable to measure or to exhibit structural effects, which are based on both time and the individual.

Furthermore, as American politician Tip O'Neill used to say, "All politics is local." This is merely an explicit way of saying that voting behavior has a territorial component in the United States. Since the US presidential election is ultimately decided by the Electoral College vote, there is an obvious interest in predicting what could happen in "battleground" or "swing" states. If local economic and social conditions affect voters' electoral behavior in a given state, then we can improve our predictive accuracy at the national level by using pooled time series voting models that make use of state-level data.

Sometimes in particular geographic areas, voting behavior is driven by long-term forces, but these are rarely immutable. First, some local effects depend on parameters such as changes in general economic conditions or/and changes in local economic conditions, which could lead to a significant residential mobility (see Tiebout 1956 and "*voting with one's feet*"). Second, some states are very stable ideologically, whereas others are highly unstable. In light of these idiosyncratic considerations, our state-by-state approach must reflect the evolution of "partisan domination" through time and space. For those purposes, we have built two partisan representation indexes reflecting states' partisan particularities. To achieve this, we distinguish between the entire period under study (1952–2012) and the period from 1980 to 2012, because this is when significant partisan changes occurred in US politics.

From 1952 to 2012, Republicans were dominant in 12 out of 50 states (plus DC) (AK, AZ, ID, IN, KS, MT, NE, ND, OK, SD, VA, WY); this means that the Republicans have a long-term electoral base of 72 electoral votes (EVs), or 13.4% of the total EV. The exact same calculation for the Democrats[5] leads to a long-term electoral base of 42 EVs, or 8% of the total EV (HI, MD, MA, MN, RI, and DC). These "locked-in" 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 observed in CA, CT, DE, IA, IL, ME, MI, NJ, NY, OR, PA, VT, WA, and WI, which translates into 206 EVs, or 38% of the total EVs. For Republicans, their potential new domination may appear in AL, AR, GA, KY, LA, MS, MO, NC, SC, TN, TX, and UT, which amounts to 142 EVs, or 26.4% of the total EVs. We must underline that the Republicans' gains are mostly in the southern states (except MO and UT), where they have captured southern Democrats. Since 2012, it should be noted that GOP is becoming stronger in NC. Finally, seven states[6] remain without strong partisan domination, representing 76 EVs, or 14% of the total EVs.

STATE-BY-STATE VOTE FUNCTION AND VARIABLES

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 in order to gain stability (Lewis-Beck and Stegmaier 2013; Nannestad and Paldam 1994). Thus, Rosenstone (1983) argues that only such an approach can explain why citizens change their vote from one election to the other.

The structural model from which we will compute the ordinary least squares (OLS) regression estimates, along with the operationalization of the variables from 1980 to 2012 in the 50 American states plus DC (N=459), goes as follow:

$$INCV_{i,t} = C + \Delta U_{i,t-n} + PJA_{t-n} + PPI_{i,t-n} + \text{Politics and Institutions}_{i,t-n} + \text{President's local Strongholds}_{i,t-n} + \text{Local Peculiarities}_{i,t-n} + \text{Error}$$

The dependent variable INCVi,t measures the vote share in the ith state (i = 1, ..., 51) and for the tth time period (t = 1980, 1984, ..., 2012) secured by the incumbent party. The outgoing President was a Democrat in 1980, 1996, 2000 and 2012 and Republican in 1984, 1988, 1992, 2004, and 2008.

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.

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 (i.e. more unemployment) should cost the incumbent some of their vote share. The second variable (PJA) is the Gallup President's Job Approval, at a national level, six months before the election. For the outgoing 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 or not the incumbent is seeking a second term. Thus, PJA2 is the president's popularity when 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 discussion above, the third variable (PPI) takes into account the characteristics of the partisan cycle in each state. PPI is divided into two local partisan domination variables, one named PPI5212 to code states having significant partisan domination since 1952 and another one called PPI8012, or "new domination" since 1980, which includes recent southern Republican strongholds. More precisely, PPI5208 gives for each state over the 1952–2012 period the rate of success for each party when this rate exceeded 85%[7] for the Republicans and 63% for the Democrats (this variable takes a zero value otherwise). When the ideology of the state is the same as the incumbent, it carries a positive sign; otherwise, it carries a negative sign. Substantively, this means that an incumbent will be rewarded by states that are ideologically similar and punished by states that are ideologically dissimilar. The PPI8012 variable was constructed in the same way, except that the rate of success is calculated for 1980–2012.

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 their home state advantage) and local electoral peculiarities such as scores usually deviating from the standard (Democrats in DC for instance). Unlike our 2012 forecasting model, a variable for the opposition nominee's vote share during their party's primaries (OPPVP) is included, asserting that the higher their score is, the more threatened the incumbent candidate could be. As mentioned above in preliminary discussion, this could be a weakness for Donald Trump in some cases.[8]

Empirical Results

The results for our vote function are presented in table 1. With regard to the main variables, the US presidential vote function provides good statistical results. This model explains 84% of the variance and all the coefficients are statistically significant (t-ratios at 0.05, two-tailed) except UT08 t-ratio at 0.10. The standard error of regression (SER) is 4.64. This means that each estimated score in each state has to be interpreted

Table 1

2016 US Presidential Election State-by-State Model. Pooled Time Series. 50 States and DC (1980–2012)

DEPENDENT VARIABLE		
Incumbent Vote Share at Presidential Election (INCV)		
INDEPENDANT VARIABLES	**COEF.**	**T-STAT.**
Constant (C)	41.34*	*30.73*
Unemployment Change (ΔU)	-0.36*	*-2.26*
President's Job Approval (PJA)		
PJA2	0.28*	*10.61*
PJA0	0.21*	*7.53*
Local Partisan Pattern Index (LPPI)		
LPPI5212	8.4*	*17.26*
LPPI8012 (a)	4.74*	*12.48*
Opposition Nominee Vote share at Primaries (OPPVP)	-0.05*	*-5.65*
For the above variables see Appendix		
INDV	-0.73*	*-21.17*
GA80	13.13*	*2.8*
AR92	-9.82*	*-2.1*
AR96	9.81*	*2.09*
HI08	-14.62*	*-3.07*
HI12	10.65*	*2.27*
RHSC	6.84*	*6.27*
DHSC	4.68*	*3.67*
DLSC	-11.89*	*-9.56*
DCDS	27.42*	*11.43*
DCRS	-30.6*	*-14.23*
VT04	-11.49	*-2.46*
UT08	-7.56**	*-1.57*
R-squared	0.84	
Adjusted R-squared	0.83	
S.E.R	4,64	
N	459	

(a) Including Southern Republican new Strongholds
(*) Significance at a 5% level or above (two-tails)
(*) Significance at a 10% level (two-tails)

Table 2

State-by-State Political Economy Model. 2016 Forecast (July 2016). Popular Votes and Electoral Votes for Democrats

States + Dist Col.	PV (a) DEM	EV (b) DEM	EV GOP
AL	45.6		9
AK	30.9		3
AZ	42.8		11
AR	48.1		6
CA	53.3	55	
CO	53.7	9	
CT	53.5	7	
DE	53.5	3	
FL	51.0	29	
GA	47.6		16
HI	58.1	4	
ID	32.1		4
IL	54.4	20	
IN	43.4		11
IA	54.3	6	
KS	43.8		6
KY	47.8		8
LA	46.7		8
ME	55.2	4	
MD	60.8	10	
MA	61.6	11	
MI	55.1	16	
MN	58.5	10	
MS	46.2		6
MO	47.4		10
MT	45.3		3
NE	41.6		5
NV	51.5	6	
NH	54.1	4	
NJ	57.7	14	
NM	52.2	5	
NY	58.9	29	
NC	47.2		15
ND	44.3		3
OH	51.3	18	
OK	43.2		7
OR	54.2	7	
PA	53.5	20	
RI	61.7	4	
SC	46.6		9
SD	43,0		3

(*continued*)

Table 2 (Continued)

States + Dist Col.	PV (a) DEM	EV (b) DEM	EV GOP
TN	47.8		11
TX	46.8		38
UT	35.4		6
VT	54.2	3	
VA	44.3		13
WA	52.8	12	
WV	46.0		5
WI	55.1	10	
WY	32.3		3
DC	86.4	3	
Maj. = 270	**50.1**	**319**	**219**

(a) Popular Vote (b) Electoral Vote

with a ± 4.64 margin of error. Furthermore, all of the coefficients for the explanatory variables in the presidential vote function (for incumbents) are of the expected sign.

Among the variables used in the forecast, the constant provides an estimate for the incumbent's electoral base, which is 41.34% of the long-run vote share. On average, a one-point rise in the unemployment rate costs the incumbent 0.36% of the votes in a given state. In the case of new candidate (from the incumbent president's same party), a popularity score of 50% for the current president yields 10.5% more of the popular votes. However, in the case of a incumbent president seeking a second term, a 50% approval score yields 14% more of the popular vote.

On average, the outgoing majority gains 8.4%, depending on the party's historical strength in a particular state (PPI5212). However, in strongholds such as these, the opposition candidate suffers a 8.4% electoral disadvantage. In "new partisan domination" states (PPI8012), the outgoing majority gains an average of 4.7% in their strongholds, while the opposition candidate suffers a 4.7% electoral loss. Finally, a vote share of 50% for the opposition party nominee during that party's primaries (OPPVP) costs the incumbent party's candidate 2.5% of the popular vote. For instance, a vote share of 36% for Donald Trump during the GOP primaries in Ohio only costs Hillary Clinton 1.8% of the popular vote, all else being equal.

Forecasting and Discussion

The 2016 pre-election values have been plugged into the model in order to predict the popular votes in each state. Table 2 shows the incumbent state-by-state vote share forecasts. The forecasting model (first created to predict 2004, 2008 and 2012 elections) gives a very narrow absolute majority for the Democrats with 50.1% of the popular vote (-1% compared with 2012) and 319 EVs (- 13 compared with 2012). This simulation shows a slight decrease in the popular vote (-0.2) from our previous forecast calculated in March 2016.

Thus, if the elections were held in September 2016, the Democrats would obtain a majority of the popular vote in 26 states (-1 compared with 2012 and -3 compared with 2008).

According to the model, the Republicans would only regain Virginia (13 EVs).

How do we explain such a result when compared to 2012? With regard to aggregated PV, the economic situation is not the same. In addition, current president's credibility is rather high (52% were satisfied with his performance in April 2016, Gallup USA Today) when compared with his own approval rating in March 2012 (when 46% were satisfied). However, Hillary Clinton is not the outgoing president and will not profit from the maximum bonus that she could expect.

The translation into votes with PJAo highlights an average electoral loss of 3.6 points per state compared to the exact same calculation with PJA2[9] (credibility when the current president is not outgoing). Such a gap erodes dramatically what is gained thanks to good economic situation (measured by a favorable change in unemployment).

Otherwise, the estimated coefficients for the PPI indexes (PPI5212 and PPI8012) do not differ significantly from those of our 2012 forecast, showing a near status quo in partisan patterns since 2008. This demonstrates that Hillary Clinton's insufficient popularity bonus prevents her to be clearly above the 50% bar.

When looking at our EV forecast, Hillary Clinton leads by 100 EVs over Donald Trump. As a consequence, how is it that the forecast is robust enough at a national level but not overall state-by-state? More specifically, are certain swing states still able to change the expected result?

Notwithstanding unforeseen shocks, Hillary Clinton's chances of victory seem to be solid in terms of EVs. But one must expect a close competition in terms of PV.

As a check, we have calculated the probability for the Democrats of attaining an absolute majority of the votes given that our model produced a -4.64 point margin of error in each state. The results are shown in figure 1.

According to the probabilities, the Democrats have absolutely no chance of winning an absolute majority in 25 states, which would have netted them 219 EVs (probability below 35%). At the opposite end of the spectrum, they have a strong chance of winning in 25 states (plus DC), totaling 309 EVs (probability above 58%). This result seems to provide a sufficiently safe margin for the Democrats. Nevertheless, what would occur if we tighten the conditions by pushing probability of reaching an absolute majority up to 73%?

Figure 1

Probability to achieve an absolute majority of the votes cast in States for Democrats

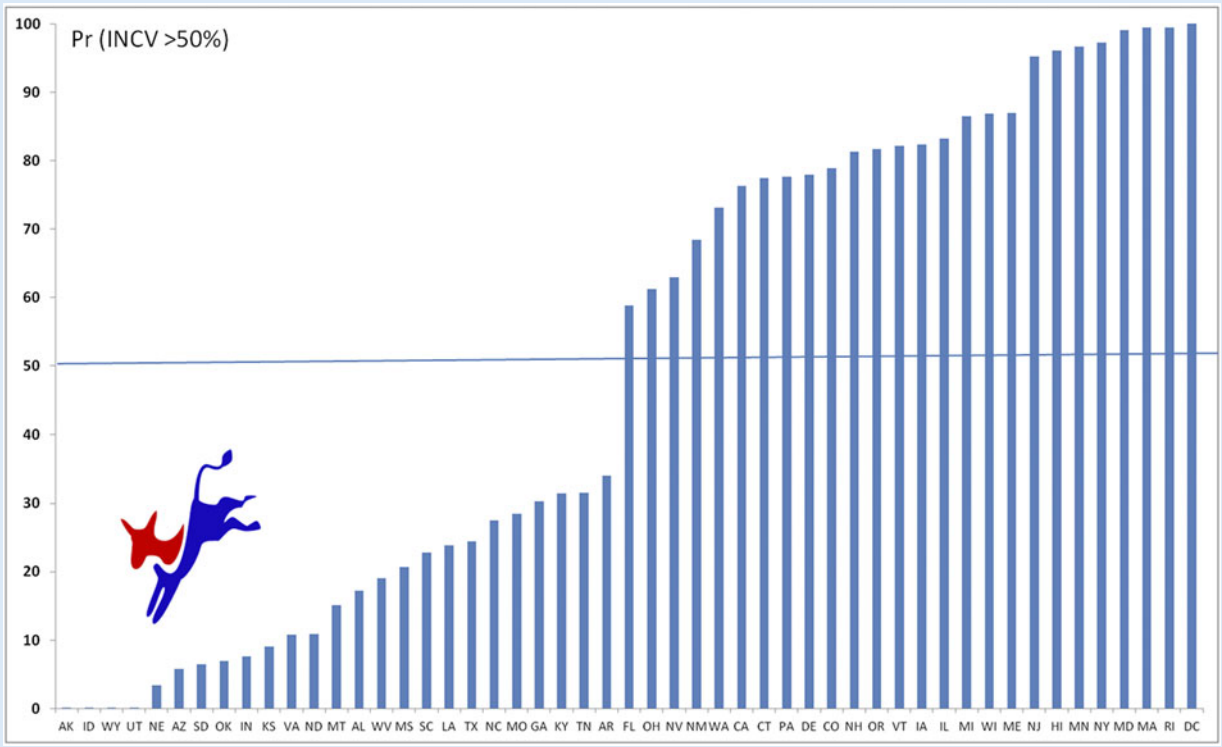

Figure 2

Probability to achieve an absolute majority exceeding 73% of the votes cast (cumulated Electoral Votes)

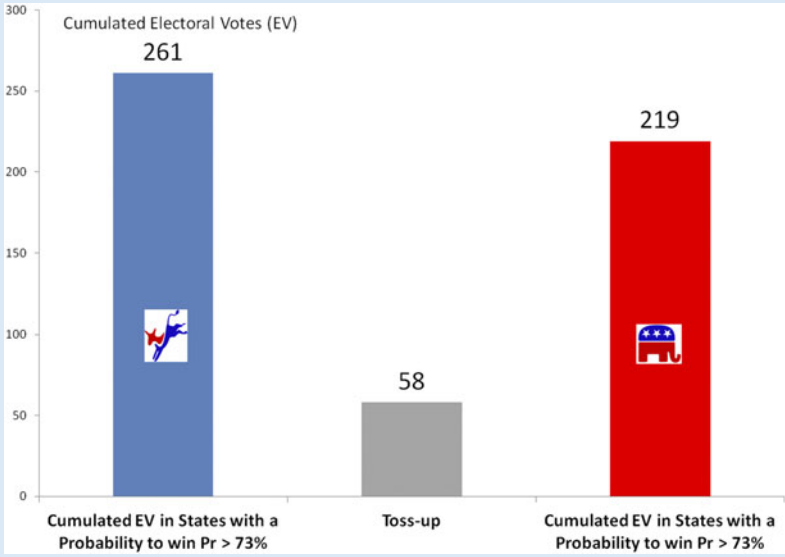

Figure 2 suggests that the Democrats could count on an electoral base of 261 EVs versus 219 for GOP. As a consequence, four states become toss-ups, representing 58 EVs (FL, OH, NM and NV). In such a situation, only FL, OH, NM and NV are the most "fragile" states for Democrats. However, at worst, Hillary Clinton would have to pick nine EVs among the four remaining states to become the 45th President of the United States.

The above results and our forecast clearly suggest that improvement in the economy had real local effects in numerous key states during Obama's second term. A second parameter seems to play a noticeable role in the forecast, which is the division among the Republicans.

As we have pointed out, Donald Trump has been the least supported nominee at the GOP primaries in four decades. In the Cleveland GOP convention, Donald Trump won, but with the most delegates opposed to his nomination since 1976.

Undoubtedly, Hillary Clinton should take this opportunity to win. Notwithstanding unforeseen shocks, Hillary Clinton's chances of victory seem to be solid in terms of EVs. But one must expect a close competition in terms of PV.

In this respect, a last unknown factor exists: are independent candidates Gary Johnson (Libertarian) and Jill Stein (Green) going to run? Our state-by-state model shows that for each one percentage point gained by an independent candidate, the incumbent loses 0.73. In July 2016, pollsters predict on average 8.5% for Johnson and 4% for Stein, meaning that the main candidates will be penalized. In the end, independent candidacies may add a little more uncertainty in determining who is going to win the popular vote but not about the identity of who will win the next presidential election in the United States. ■

NOTES

1. This article was written at the end of July 2016.
2. On this approach, see Lewis-Beck, Nadeau and Bélanger (2016)
3. For other sub-national models see Klarner (2012) and Berry and Bickers (2012).
4. See Jérôme and Jérôme-Speziari (2004 ; 2012 ; 2013) and Pottier (2008).
5. Here the threshold giving the rate of electoral success has to be above 63% (average = 39%; standard error = 24%).
6. CO, FL, NV, NH, NM, OH, WV
7. The victory threshold is based on the last 15 presidential elections and is calculated in the following way: 85% = 61% (average rate of Republican victories over 1952–2012) + 24% (standard error). For the Democrats: 63% = 39% (average rate of Democrat victories over 1952–2012) + 24% (standard error).
8. The specification of the variables whose estimated coefficients are not used to forecast the 2012 election is described in the Appendix.
9. With PJA2 an outgoing president would gain 52 x 0.28 = 14.56 on average, while with PJA0, Hillary Clinton should get 52 x 0.21 = 10.92. The mean loss is 10.92-14.56 = -3.64 in each state.

REFERENCES

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

Berry, Michael J. and Kenneth N. Bickers. 2012. "Forecasting the 2012 Presidential Election with State-Level Economic Indicators." *PS: Political Science & Politics* 45 (4): 669–74.

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

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: 425–440.

Jérôme, Bruno and Veronique Jérôme-Speziari. 2004. "Bush l'emporterait sur Kerry". *La lettre de L'Expansion.* 25 October. 1730: 2.

Jerôme, Bruno and Véronique Jerôme-Speziari. 2012. "Forecasting the 2012 US Presidential Election: Lessons from a State-by-State Political Economy Model." *PS: Political Science & Politics* 45 (4): 663–668.

———. 2013. "Why the state-by-state political economy model did it right." *PS: Political Science & Politics* 46 (1): 45–6.

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

Klarner, Carl. 2012. "State-Level Forecasts of the 2012 Federal and Gubernatorial Elections." *PS: Political Science & Politics* 45 (4): 655–62.

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 Tom W. Rice. 1982. "Presidential Popularity and Presidential Vote." *Opinion Quarterly* 46: 534–37.

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

Lewis-Beck, Michael S. and Mary Stegmaier. 2013. "The VP-Function Revisited: A Survey of the Literature on Vote and Popularity Functions after over 40 years." *Public Choice* 157 (3): 367–85.

Lewis-Beck, Michael S., Richard Nadeau and Éric Bélanger. 2016. "The British General Election: Synthetic Forecasts." *Electoral Studies* 41: 264–268

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.

Norpoth, Helmut. 2014. "The Electoral Cycle." *PS: Political Science & Politics* 47 (2): 332–5.

Pottier, Jean.-Marie. 2008. "Barack Obama, Chouchou des prévisionnistes." 31 octobre. www.challenges.fr. http://www.challenges.fr/monde/20081031.CHA7788/barack-obama-chouchou-des-previsionnistes.html

Rosenstone, Steven J. 1983. *Forecasting Presidential Elections.* New Haven: 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 0.73% points to the incumbent (whichever party he may be from).

OPPVP : opposition nominee vote share at primaries. This variable shows that one point gain in votes for opposition Nominee costs 0.05% points to the incumbent.

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, **AR96** Arkansas in 1996, **HI08** and **HI12** Hawaii for Barack Obama in 2008 and 2012.

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

DLSC: dummy variable scored 1 in order 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 in District of Columbia when Democrats D or Republicans R are incumbents (0 otherwise).

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).

VT04 is outlier: dummy variable scored 1 (0 otherwise) for Vermont in 2004.