Predicting the Incumbent Party Vote Share in U.S. Presidential Elections

Masoud Moghaddam and Hallie Elich

Once every four years, it has become an American ritual to have the opportunity to make history and change a major part of the world by electing or reelecting a president. Certainly, presidential election years feature not only the symbolic exercising of a fundamental American right, but also the fruits and labors of an extraordinarily complex political process. However, even the firmly rooted structure of this process, the majority of which is some 240 years old, and the solidarity exuded by "dyed in the wool" Republicans and Democrats, has been transformed into what often appears to be a chaotic frenzy. Clouding the matter even further, recent advancements in information technology have increased mass media attention surrounding campaigns. The publicizing of debates, primaries, polls, and political mudslinging is now embedded in the political system and can be transmitted around the world in the blink of an eye. Unquestionably, the process of electing a president has long involved campaigns, primaries and caucuses, debates, polls, and pundits. Yet, in a modern way, our political preoccupation often appears unnecessarily self–created, mundane, phlegmatic, and to some extent capricious. The impetus of election-year popularity and commercialization are intertwined with an established ritual so profoundly rooted in American history. Indeed, as the modern process unwinds and more often than not meanders, recent history has revealed that, deep–

down, more is involved in choosing America's next president than merely flipping a coin. It is abundantly clear that economic, historical, and personal issues influence and form the basis for many swing voters' decisions. Toward that end, in this article, we present a brief summary of the presidential election literature in the United States, describe the empirical model used to predict the 2008 presidential election, and discuss policy implications of the estimated model.

The Voting Theory

Ray Fair (1978) developed an empirical model to predict the outcome of U.S. presidential elections. The underlying theory behind Fair's vote equation is predominantly that of Anthony Downs (1957), Gerald Kramer (1971), and George Stigler (1973) among others. Accordingly, although well informed voters look back more than a year, voting behavior depends heavily on economic events in the year of the election. Correspondingly, William Niskanen (1979) demonstrated that the Logit version of the incumbent's popular vote share (I) depicted by ln (I/1 − I, where ln = natural logarithm) in election years during the 1896–72 period, is determined by economic variables (four-year percentage changes in real per capita net national product, the employment rate, the consumer price index, the stock price index, and the corporate bond rate), fiscal policy tools (four-year percentage changes in the real per capita federal government expenditure or revenue), two political variables (ln [I(t − 4)/1 − I(t − 4)] and a dummy variable depicting whether or not there is an incumbent candidate), along with another dummy variable capturing major wars. In the context of an overparameterized election model (as is the case in other empirical studies), his overall findings tend to suggest that an incumbent president is most likely reelected if economic growth has been preserved, federal spending is under control, and a major war has been avoided. Most interestingly, changes in both real per capita government expenditures and tax income (in two separate regression equations) inversely and significantly determine ln (I/1 − I), while the difference between the two is a measure of "debt illusion."[1] By the same token, Fair (1996a, 1996b) contends that voters react to the rate of growth of real per capita GDP in the

---

[1]For a complete description of the data, see Niskanen (1979: 119–20).

election year and that they compare the past performance of rival parties. Consequently, they cast their votes for the party that appears to be able to maximize their expected utility (satisfaction).[2]

Fair (1988, 1996a, 1996b, and 2006) re-estimated his vote equation once every four years in the sample period 1916–2004 using ordinary least squares (OLS), and made predictions about whether the Democratic or Republican nominee would win the presidency.[3] The most recent re–estimation (the 2004 update), predicts the incumbent party share of the two–party popular presidential vote in 2008 and features both incumbency and economic explanatory variables.[4] The incumbency variables include three dummies that track whether the current administration is Democratic or Republican (*PARTY*), whether or not an incumbent is running for reelection (*PERSON*), and how long the incumbent party has been in office (*DURATION*). The economic variables include the growth rate of real per capita GDP in the first three quarters of the election year at an annualized rate (*GROWTH*), inflation depicted by the absolute value of the growth rate of the GDP deflator in the first 15 quarters of the administration at an annualized rate (*INFLATION*), good news proxied by the number of quarters in the first 15 quarters of the administration during which the growth rate of real per capita GDP (PGDP) is greater than 3.2 percent at an annualized rate (*GOODNEWS*), and a war dummy variable that is 1 in 1920, 1944, and 1948 and 0 otherwise (*WAR*). These three election years are either after or during the two World Wars, and are included to examine the effects of these two major wars on the incumbent party vote share (rallying around the flag). Indeed, Fair emphasizes the dominance of these wars by zeroing out *INFLATION* and *GOODNEWS* variables in 1920, 1944, and 1948.

In general, ignoring the duration of a party's occupation of the White House presupposes that incumbents running for president have an advantage over non–incumbents. However, the longer a party dominates the presidency, the greater the tendency of voters to subsequently elect the opposition party candidate (considering only Democrats and Republicans). Next, the greater the PGDP in

[2]Care should be exercised in interpreting Fair's results because the estimated election models suffer considerably from small sample biases.
[3]For a similar work, see Campbell and Wink (1990).
[4]Fair (1996) states that if there is a third party involved in the presidential election, on average, it takes away an equal number of votes from the two dominant parties.

the first three quarters of the election year, the greater the expected incumbent party vote share. Voter's opinion on the state of the economy is directly related to PGDP and should quite possibly affect the choice of a candidate. Witnessing a healthy rate of growth for PGDP during the election year is likely to increase the appeal of the incumbent party candidate (whether or not he or she is actually an incumbent). Similarly, the more quarters that the growth rate of PGDP exceeds 3.2 percent, the greater the benefit to the incumbent party vote share. However, the majority of swing voters vote their pocketbook, so an increase in inflation should favor the non–incumbent party.

Fair's updated results are generally supportive of the aforementioned political and economic intuition, though he emphasizes that "a voting equation should be judged according to the size of its errors, not according to how many winners it correctly predicts" (Fair 2006). He also acknowledges the danger in data mining considering the limited number of observations, as well as dealing with a model that may suffer from overparameterization. Moreover, in spite of the robust estimation within a stable model, he points out that the findings are very sensitive to the dummy variable depicting the World Wars. For example, his 2004 update features a 4.33 percent overstatement of the Republican Party vote share along with an incorrect prediction that George H.W. Bush would unquestionably win the presidential election in 1992. The considerable overprediction in 1992 is particularly concerning because the only issue negatively impacting Bush's vote share would be that Republicans had held the presidency for the past three consecutive terms.[5] As such, the prediction's anomalies in 1992 might be attributable to the model's misspecification, which is addressed in the next section.

Modifications of Fair's Model

In the modified Fair's model, the dependent and the first three explanatory variables are the same as those defined by Fair, but the rest are different as described below. The empirical model pursued in this article is as follows:

---

[5]For more information, see Haynes and Stone (1994).

(1) $VOTE = \beta_0 + \beta_1\ PARTY + \beta_2\ PERSON + \beta_3\ DURATION$
$+ \beta_4\ WARD + \beta_5\ GROWTHDIFF + \beta_6\ INFLATIOND$
$+ \beta_7\ GOODNEWSD + \beta_8\ HEIGHT + \beta_9\ HOUSE.$

Where:

$VOTE$ = Incumbent party share of the two party-popular vote.

β_0 and $\beta_1 - \beta_9$ = the intercept and regression coefficients, respectively.

$PARTY = \begin{cases} 1 \text{ if there is Democratic incumbent at the time of the election} \\ -1 \text{ if there is a Republican incumbent at the time of the election} \end{cases}$

$PERSON^6 = \begin{cases} 1 & \text{if the incumbent is running for reelection} \\ 0 & \text{otherwise} \end{cases}$

[6]Incumbents include elected presidents running for reelection and also elected vice presidents who are completing their first term in office and are running for president. Vice presidents who have served two terms and then, upon completing their second term as vice president, run for president are not considered incumbents. Also, note that there is a distinction between the incumbent party candidate and incumbents running for reelection. There will always be an incumbent party candidate, assuming Democratic and Republican constraints, but there is not necessarily always an incumbent running for reelection (as is the case in 2008). Lyndon B. Johnson was elected vice president in 1960 and assumed the presidency in 1963 following Kennedy's assassination. Thus, when President Johnson ran for a second term in 1964, with Hubert Humphrey as his running mate, he was considered an incumbent. Yet, when Humphrey ran for president in 1968, he was not considered an incumbent even though he was elected vice president in 1964. Also, in 1976, President Ford ran against Jimmy Carter but was not considered an incumbent since he was never elected president.

$$
DURATION = \begin{cases} 0 & \text{if the incumbent party has been in power} \\ & \text{for one term} \\ 1 & \text{if the incumbent party has been in power} \\ & \text{for two terms} \\ 1.25 & \text{if the incumbent party has been in power} \\ & \text{for three terms} \\ 1.5 & \text{if the incumbent party has been in power} \\ & \text{for four terms} \\ \vdots \end{cases}
$$

$$
WARD^{7} \begin{cases} 1 & \text{in 1920, 1944, 1948, 1952, and 1972} \\ 0 & \text{otherwise} \end{cases}
$$

GROWTDIFF = the growth rate of real per capita GDP differential between the first three quarters of the election year (annualized) and the first three quarters of the second year of the current administration.

INFLATIOND = absolute value of the growth rate of the GDP deflator in the first 15 quarters of the administration (annualized), except for 1920, 1944, 1948, 1952, and 1972, where the values are zero.

---

[7]The war dummy *WARD* = 1 in 1952 and 1972 to reflect the Korean and Vietnam Wars. The 1952 election, between Dwight Eisenhower and Adlai Stevenson, occurred during the Korean War (1950–53). The 1972 election between Richard Nixon and George McGovern was waged during the Vietnam War. The impact of the Korean and Vietnam Wars was likely not as profound on U.S. presidential elections as were the two World Wars. Nevertheless, these wars are represented uniformly by the *WARD* dummy variable. As such, *INFLATIOND* and *GOODNEWSD* are zeroed in 1920, 1944, 1948, 1952, and 1972 to maintain consistency with Fair's treatment of *INFLATION* and *GOODNEWS* in his vote equation.

GOODNEWSD = number of quarters in the first fifteen quarters of the administration in which the growth rate of real per capita GDP is greater than 3.2 percent at an annual rate, except for 1920, 1944, 1948, 1952, and 1972, where the values are zero.

HEIGHT = incumbent party candidate's height (inches) minus challenger's height (inches).

$$HOUSE^8 = \begin{cases} 1 & \text{if the incumbent party has majority control over the House in the most recent Congress} \\ 0 & \text{otherwise} \end{cases}$$

The *GROWTHDIFF* variable is generated to not only study the per capita GDP rate of growth in the first three quarters of the election year (Fair's *GROWTH*), but also to capture the existence of possible political business cycles. The incumbent party has an incentive to impress voters by demonstrating economic strength in the election year. One way to achieve this is to slow down the economy in the second year of a presidential term and then stimulate the economy in time for the next election. While an administration does not have supreme power over the status of the economy, it certainly has a way of exerting significant influence. An administration of each party wants to remain in power either directly or via other party members—even if neither the president nor the vice president actually runs for reelection. Thus, there is always an incentive to create a political business cycle—as long as such a cycle yields positive returns for the incumbent party vote share. Without making a

---

[8]Note that the party in majority control of the House corresponds to party division immediately following the midterm election and does not remain constant in any given Congress—for example, due to members changing parties or dying. To illustrate *HOUSE* in the 1916 presidential election, *HOUSE* = 1 because the incumbent party was Democrat and the Congress elected in 1914 featured a Democratic majority. Majority control over the U.S. Senate can also be examined. However, House and Senate majority parties are often identical. Also, the structure of the six–year senatorial terms compared to the two-year representative terms is more complicated—every two years (even-numbered years) one-third of the Senate seats are up for election, whereas every House seat is up for election.

distinction in the case when *PERSON* = 1, political business cycles can be investigated via *GROWTHDIFF*.[9]

Although *WARD* depicts more wars than the same variable (*WAR*) in Fair's model, its impact on the incumbent vote share is the same as that of Fair's. Voters typically do not desire presidential party changes during times of war, as this might interfere with the war effort. However, if one party is particularly poor at managing a war or its aftermath, the opposition party may be favored. In like manner, there is a widely held belief that presidential candidates should "look presidential" in that they are tall, physically fit, and thus publicly appealing. Since most U.S. presidents have historically been tall, a height differential (*HEIGHT*) is included as an independent variable. This variable is simply the height difference between the incumbent party candidate and the challenger—the greater the *HEIGHT* is, the better the *VOTE* should be. While it may seem trivial to include the variable *HEIGHT*, the reality is that the image (in all its expressions) of presidential candidates likely matters to most voters. Finally, if the incumbent party has had a majority presence in the U.S. House of Representatives in the last two years of a presidential term, the incumbent party vote share is expected to increase. This type of administration is more likely to sustain incumbent party power than a mixed administration encompassing different presidential and House majority parties.

Empirical Findings

The modified model has been estimated using OLS in the sample period 1916–2004 and the findings are reported as follows:[10]

$$(2)\ VOTE = 50.05 - 2.06\ PARTY + 1.85\ PERSON - 5.44\ DURATION$$
$$\qquad\quad (33.9)\ (-5.63)\qquad (2.12)\qquad\quad (-7.54)$$

[9]GROWTHDIFF only takes into account the difference in the growth rates of real per capita GDP in the election year and that of two years prior. As such, GROWTHDIFF does not completely capture the purported slowing down of the macroeconomy in the middle of an administration's term. Moreover, GROWTHDIFF can represent the difference between two growth rates of the same sign or between growth rates of opposite signs. Nonetheless, GROWTHDIFF reflects changes in the growth rates of real per capita GDP in these two years and when positive, it indicates an economic expansion.

[10]The data sources are Fair (2006), Office of the Clerk (2008), U. S. Senate (2005), and Wikipedia (2008a, 2008b).

$$+ 7.34 \; WARD + 0.226 \; GROWTHDIFF$$
$$(4.92) \qquad (5.16)$$
$$- 0.976 \; INFLATIOND + 1.06 \; GOODNEWSD$$
$$(-5.99) \qquad (6.55)$$
$$+ 0.699 \; HEIGHT + 1.52 \; HOUSE.$$
$$(6.43) \qquad (2.09)$$

R^2	= 0.97
Adjusted R^2	= 0.95
Standard Error of regression	= 1.49
Durbin-Watson statistic	= 2.54

The t–statistics are in the parentheses below each estimated regression coefficient.

All the estimated regression coefficients have the correct sign and are significantly different from zero at about the 5 percent level. Since $PARTY = 1$ or -1, depending on the incumbent party, with all other variables equal to zero, $VOTE = 50.05 + 2.06 = 52.11$ if the incumbent party is Republican, and $VOTE = 50.05 - 2.06 = 47.99$ if the incumbent party is Democratic. Thus, ignoring all other explanatory variables, Republicans are more likely to stay in power than are Democrats. In fact, the presence of a Democratic administration alone does not benefit the Democratic nominee, but rather favors the Republican nominee—possibly suggesting more solidarity among Republicans and a greater ability on their part to maintain power.

The coefficient for *PERSON* implies that when an incumbent runs for president, the incumbent party vote share increases by 1.85 percent. The coefficient for *DURATION* states that the longer the incumbent party is in power, the lower (–5.44 percent) the incumbent party vote share. Although the incumbent party vote share is not adversely affected by that party's one–term presence, this rather large decrease supports the theory that voters desire party changes even after two consecutive terms of the same party rule. Next, when *WARD* = 1 (in 1920, 1944, 1948, 1952, and 1972), the incumbent party vote share increases by 7.34 percent, which signals that voters favor the incumbent party after and/or during the wars. This is consistent with the expectation that voters probably believe switching presidential parties would be detrimental to war and war-recovery

efforts. The coefficient of *GROWTHDIFF* shows that for every one percent increase in the difference between the PGDP in the first three quarters of the election year and that in the first three quarters of the second year of the administration, the incumbent party vote share increases by 0.226 percent.

While *GROWTHDIFF* must exceed approximately 4.425 to increase the incumbent party vote share by just 1 percent, this supports the incentive of the incumbent party to stimulate the macro-economy in the last two years of a term in order to prop up its vote share in the next election. Next, an incremental increase in the absolute value of the growth rate of the GDP deflator in the first 15 quarters of an administration (annual rate), results in a decrease in the incumbent party vote share by 0.976 percent. Also, for every quarter in the first 15 quarters of an administration in which the PGDP exceeds 3.2 percent at an annual rate, there is an associated 1.06 percent increase in the incumbent party vote share. This supports the incumbent party desiring a strong economy, especially near the end of its term to cater to the nearsighted and especially the swing voters. For every inch taller the incumbent party candidate is than the challenger, there is a 0.699 percent increase in the incumbent party vote share. For example, a positive four-inch height differential corresponds to an increase of 2.796 percent in the incumbent party vote share. Lastly, if the majority of the House is that of the incumbent president's party, the incumbent party vote share is likely to increase by 1.52 percent. The R^2, adjusted R^2, and the standard error of the regression are indicative of a good fit and are more desirable than those corresponding to Fair's 2004 update. The Durbin-Watson (D-W) statistic is lower than that of Fair's model, although it hints at a negligible first-order negative residual autocorrelation.

In examining the actual and predicted values of the incumbent party vote share (*VOTE*), compared to Fair's 2004 update, the modified model predicts 95 percent of the winners correctly and exhibits a 40 percent reduction in the standard error of regression.[11] Furthermore, different criteria for evaluating the predictability power of the two models are summarized in Table 1.

---

[11]In the 2000 controversial presidential election, the modified model predicts 50.15 percent of the vote share for the incumbent party (a marginal win) compared to Fair's 2004 update 49.63 percent (a marginal defeat). Moreover, Fair's 2004 update predicts the winner incorrectly in 1916, 1960, 1968, and 1992.

TABLE 1

Prediction Power of Fair's Model and
Its Modified Version

Evaluation Criteria	Modified Model	Fair's Model
Root Mean Squared Error (RMSE)	1.122	2.052
Bias Proportion	0.000	0.000
Variance Proportion	0.007	0.025
Covariance Proportion	0.993	0.975

In terms of evaluating the predicted values, the smaller the RMSE, the better the predictability power of the model and vice versa. Demonstrably, the modified model features approximately a 45 percent reduction in the RMSE compared to Fair's model. Bias and variance proportions measure how far the mean and variance of the forecast are from those of the actual data, respectively. The covariance bias is a measure of the remaining unsystematic forecasting errors. In particular, it is desirable for the bias and variance proportion to be small so that most of the bias is concentrated on the covariance proportion. The modified model apparently outperforms Fair's model in every aspect and parallels the bias proportion. In line with Fair's suggestion (see Fair 2002), to shed more light on the issue, the prediction errors (predicted values minus actual values) of both models are displayed in Figure 1.

As can be seen, the modified model on average displays much smaller prediction errors compared to Fair's 2004 update. Therefore, the modified model is likely to exhibit accuracy in predicting the winner of the 2008 election.

The 2008 Conditional Prediction

For the 2008 election year, *PARTY* = – 1, *PERSON* = 0, *DURA-TION* = 1, and *WARD* = 0. Fair made a January 31, 2008, prediction for the winner of the 2008 election using predicted values of *GROWTH, INFLATION*, and *GOODNEWS* of 1.8 percent, 3.1 percent, and 2, respectively. However, *INFLATIOND = INFLATION*

FIGURE 1
PREDICTION ERRORS IN FAIR'S AND MODIFIED MODELS
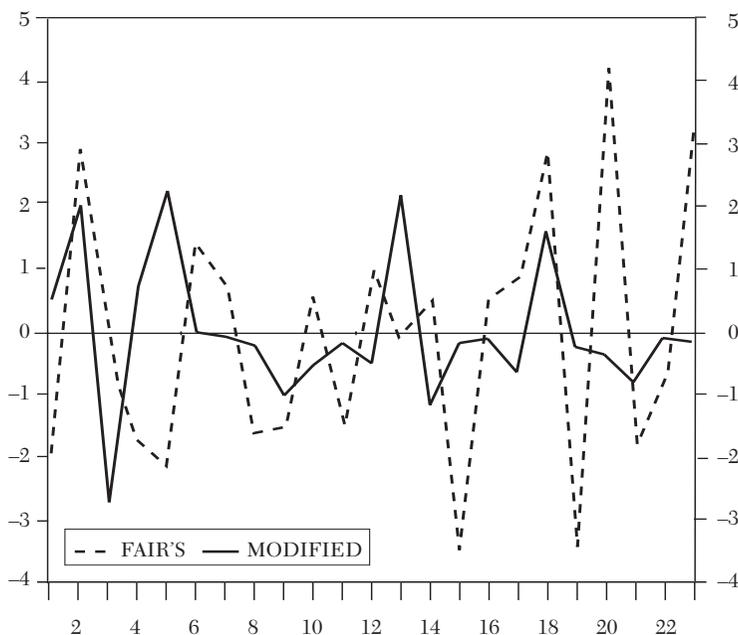
and *GOODNEWSD = GOODNEWS* in every year other than 1952 and 1972. Therefore, it is assumed that *INFLATIOND* = 3.1 percent, *GOODNEWSD* = 2, and *GROWTHDIFF* = −0.5 percent.[12] Currently, Democrats control the House and thus, *HOUSE* = 0.

Using the modified model, *VOTE* = 45.6514 + 0.699 *HEIGHT*. In regard to the height differential, since Barack Obama is 6 inches taller than John McCain, *VOTE* = 45.6514 + 0.699 (−6) ≈ 41.46. The modified Fair model correctly predicted a Democratic victory, despite the uncertainty surrounding the economy.

Conclusion

Economic, historical, and personal issues influence voters' decisions. Quantitative and qualitative economic, political, and personal variables have been examined and are significant in predicting the outcome of presidential elections. Although Fair's 2004 update is a

[12]The growth rate of real per capita GDP in the first three quarters of 2008 was estimated by Fair to be 1.8 percent. The growth rate of real per capita GDP in the first three quarters of 2006 was approximately 2.3 percent.

well-established vote model and appears to be stable with reasonable predictive power, it might be plagued with specification errors. Replacing *GROWTH* with *GROWTHDIFF* in Fair's model in order to examine the likelihood of political business cycles, redefining the *WAR* dummy variable to account for more wars than only world wars, introducing the *HEIGHT* and *HOUSE* variables lead to a model that exhibits considerably smaller prediction errors. Beyond the explanatory variables explored by Fair, the modified model demonstrates the existence of a mild political business cycle, the significance of height of the candidates, the importance of the incumbent party control of the U.S. House of Representatives, and the substance of wars in predicting the incumbent party share of the popular vote.

References

Campbell, J. E., and Wink, K. A. (1990) "Trial–Heat Forecasts of the Presidential Vote." *American Politics Quarterly* 18: 251–69.

Downs, A. (1957) An *Economic Theory of Democracy*. New York: Harper and Row.

Fair, R. C. (1978) "The Effect of Economic Events on Votes for President." *Review of Economics and Statistics* 60: 159–73.

_____ (1988) "The Effect of Economic Events on Votes for President: 1984 Update." *Political Behavior* 18: 119–39.

_____ (1996a) "The Effect of Economic Events on Votes for President: 1992 Update." *Political Behavior* 10: 168–77.

_____ (1996b) "Econometrics and Presidential Elections." *Journal of Economic Perspectives* 10 (3): 89–102.

_____ (2002) *Predicting Presidential Elections and Other Things*. Stanford, Calif.: Stanford University Press.

_____ (2006) "The Effect of Economic Events in Votes for President: 2004 Update." Yale University (http://fairmodel.econ.yale.edu).

Haynes, S. E., and Stone J. A. (1994) "Why Did Economic Models Falsely Predict a Bush Landslide in 1992?" *Contemporary Economic Policy* 12: 123–30.

Kramer, G. H. (1971) "Short-Term Fluctuations in U.S. Voting Behavior 1896–1964." *American Political Science Review* 65: 131–43.

Niskanen, W. A. (1979) "Economic and Fiscal Effects on the Popular Vote for the President." In D. W. Rae and T. J. Eismeier (eds.) *Public Policy and Public Choice,* 93–117. London: Sage.

Office of the Clerk (2008) "Party Division of the House of Representatives." Available at http://clerk.house.gov.

Stigler, G. J. (1973) "General Economic Conditions and National Elections." *American Economic Review* 63: 160–67.

U.S. Senate (2005) "Party Division in the Senate, 1789–Present." Available at http://www.senate.gov.

Wikipedia, the Free Encyclopedia (2008a) "United States Presidential Election" (http://en.wikipedia.org).

_____ (2008b) "Heights of United States Presidents and Presidential Candidates" (http://en.wikipedia.org).