Replicating and Extending the Fair Presidential Vote Equation (1916–2024)

Omolara Oni

1.INTRODUCTION

The relationship between economic conditions and presidential election outcomes has been one of the most persistent and debated topics in political economy. A large body of literature argues that voters rely on recent economic performance when evaluating the incumbent party, a behavior commonly referred to as retrospective economic voting. The idea is that citizens reward the governing party when economic conditions have been favorable and penalize it when economic conditions have deteriorated. This perspective aligns with rational choice theory, where voters are modeled as agents who form judgments based on observable outcomes that affect their personal welfare. Because economic statistics provide a widely shared metric of national well being, they often serve as a focal point for voter evaluation.

Ray Fair's voting equation represents one of the earliest and most influential attempts to formalize this connection. Fair's approach combines political factors such as incumbency and party control with macroeconomic fundamentals such as inflation, real per capita income growth, and the number of favorable economic quarters during a presidential term. His model has been widely cited both in academic research and in media discussions of election forecasting. The simplicity of the framework has allowed researchers to test how faithfully economic fundamentals alone can account for variation in voter behavior across nearly a century of elections.

This replication project has two primary goals. The first goal is to reproduce Fair's original estimation results using the historical sample for the period 1916 through 1992. This involves reconstructing economic variables using modern time series from the Federal Reserve Economic Data system, aligning quarterly observations to Fair's definitions, and verifying that the estimated coefficients match those reported in the published work. The second goal is to evaluate the model's predictive performance in a more modern political environment. Using data for the elections from 1996 through 2024, the study applies the original estimated coefficients to generate out of sample forecasts and compares these predictions to actual election outcomes.

Carrying out this replication requires several methodological decisions and adjustments. Quarterly GDP data before 1947 are not available in modern datasets, which creates challenges in matching Fair's original values. Some political variables for modern elections must be hand coded because they are not included in the original dataset. In addition, definitions of inflation,

real income, and population growth in modern datasets require careful alignment to replicate Fair's construction of the economic variables g3, p15, and n. These details matter because small differences in measurement can lead to meaningful changes in the regression estimates.

Despite these challenges, replicating Fair's model provides a valuable opportunity to examine how sensitive the voting equation is to data construction and to assess whether the structural relationships identified in early twentieth century elections still hold in a political landscape shaped by polarization, media fragmentation, and changing voter coalitions. This paper shows the full replication process, presents the estimated results, and evaluates the model's performance using modern elections.

2.ECONOMETRIC MODEL

The empirical framework used in this replication is built around the voting equation proposed by Ray Fair. The equation is designed to explain the Democratic share of the two party presidential vote using a combination of political and economic predictors. The model treats the vote share as a linear function of incumbency, party control of the White House, the duration of the incumbent party in office, and several measures of recent economic performance. Although the structure is simple, the model captures the central theoretical idea that voters respond to both partisan expectations and macroeconomic conditions.

In its standard form, the model is written as:

$$V_t = \beta_0 + \beta_1 I_t + \beta_2 DPER_t + \beta_3 g_t + \beta_4 p_{15} + \beta_5 n_t + \beta_6 DUR_t + \varepsilon_t$$

Each component of the equation plays a distinct role. The variable $I_t$indicates whether the incumbent administration is Democratic or Republican. The variable $DPER_t$identifies whether the sitting president is the Democratic candidate, while DUR_tmeasures the length of time the incumbent party has already controlled the presidency. These political variables capture expectations that long periods of single party control may reduce voter support and that incumbency advantages vary by party.

The economic variables are designed to reflect short run conditions that voters may associate with presidential performance. The variable $g3_t$measures the sum of real per capita income growth over the three quarters prior to the election. The variable $p15_t$reflects inflation over the fifteen quarters prior to the election, following Fair's argument that voters respond more strongly to price instability when it accumulates over longer periods. The variable $n_t$counts the number of quarters during the presidential term in which real per capita GDP growth exceeded two point nine percent, which Fair considered a threshold for meaningful economic improvement.

The replication follows Fair's original methodology whenever possible, but several practical adjustments were required. The first challenge concerns the construction of the economic variables. The Federal Reserve Economic Data system provides quarterly time series for GDP, inflation, and population, but quarterly GDP data before 1947 are not available in the modern series. Fair appears to have used historical reconstructions that are no longer distributed through contemporary sources. As a result, this replication uses the earliest available quarterly observations and fits the model using the data that can be constructed consistently across time. This approach preserves the structure of the model but introduces unavoidable differences in the numerical values of g3, p15, and n for the early elections.

To compute the economic variables, I merged quarterly real GDP, population, and the GDP deflator. Real per capita GDP was computed directly from these series. The growth variable g3 was derived from quarterly growth rates expressed in annualized form using Fair's definitions. Inflation was calculated using quarterly changes in the GDP deflator. The structure of the variables mirrors Fair's original formulas, but minor differences arise because the modern series occasionally differ in composition or coverage from the historical estimates he used.

The political variables for the original sample are taken directly from the Fair dataset. For the elections from 1996 onward, these variables are not supplied by Fair, so I constructed them manually based on the definitions in the original work. This includes identifying the incumbent party, coding whether the Democratic candidate was the sitting president, and calculating the party duration variable. These decisions required consistency checks across multiple historical sources to ensure that each variable aligned with the definitions used in the earlier sample.

The model is estimated using ordinary least squares over the original sample period of 1916 through 1992. Fair used the same estimator and did not include corrections for heteroskedasticity or autocorrelation, so this replication follows that approach to maintain comparability. After estimating the coefficients for the original sample, the model is used to generate predictions for the elections from 1996 through 2024. These out of sample forecasts allow for a direct evaluation of how well the historical relationships extend into the modern political era.

Throughout the analysis, I place emphasis on transparency in variable construction and careful documentation of any assumptions required due to data availability constraints. The methodological goal is not only to reproduce Fair's numerical results, but to understand how the structure of the model behaves when updated with contemporary data and applied to elections shaped by very different political environments. To evaluate how well the Fair model performs, I next turn to the data used to construct its key economic and political variables. Because the model relies on specific quarterly series and historically defined political indicators, careful data construction is essential for both replication and extension

3.DATA AND VARIABLE CONSTRUCTION

This project relies on two sets of data. The first is the original dataset created by Ray Fair, which contains the variables used in his 1996 voting equation for the elections from 1916 through 1992. The second is an extended dataset that I created by reconstructing Fair's economic variables for the modern elections from 1996 through 2024 using publicly available macroeconomic series from the Federal Reserve Economic Data system. Bringing these two components together required a significant amount of data cleaning, reconstruction, and verification. This section explains the decisions behind each step and documents the final structure of the dataset used in the replication.

3.1.Original Fair Data: 1916 through 1992

The Fair dataset includes the Democratic two party vote share, the incumbency indicator, whether the Democratic candidate was the sitting president, the duration of the incumbent party in office, and Fair's economic variables. These economic variables consist of g3, p15, and n. The variable g3 is the sum of real per capita GDP growth over the three quarters preceding the election. The variable p15 is the sum of fifteen quarters of inflation. The variable n counts the number of quarters in which real per capita income growth exceeded two point nine percent during the incumbent president's term. Fair published these values in tabular form and the original dataset reproduces them exactly.

A unique issue arises with early elections. Several values for g3, p15, and n in the earliest elections cannot be derived directly from modern FRED data because the necessary quarterly GDP estimates do not exist before 1947. Fair appears to have used historical reconstructions that were produced by earlier research and are no longer available in a form that can be merged with contemporary sources. Because this replication does not attempt to recreate those unavailable historical reconstructions, the original Fair values for the pre 1947 elections are accepted as given.

3.2.Constructing Economic Variables for the Modern Elections

The second component of the dataset involves producing new values for g3, p15, and n for the elections from 1996 through 2024. To do this, I gathered quarterly real GDP, the GDP deflator, and population series from FRED. These series are complete from 1947 to the present, so they allow for consistent construction of the variables for all modern elections.

The process begins by converting the raw GDP data into real per capita GDP, which is central to Fair's model. I computed this by dividing real GDP by the quarterly population measure. The next step is to calculate quarterly growth rates of real per capita GDP expressed in annualized terms. This aligns with Fair's definition of economic growth. The variable g3 is then constructed by summing these growth values for the quarter of the election and the two quarters immediately preceding it.

Inflation is taken from the quarterly changes in the GDP deflator. Fair defines p15 as the sum of inflation across the fifteen quarters prior to the election. Although this definition is simple, it is sensitive to data availability and especially sensitive to the choice of price index. The GDP deflator is the closest available series to the one Fair used, so it forms the basis for the inflation calculations in this project.

The final economic variable n counts the number of quarters in which real per capita GDP growth exceeds two point nine percent. This threshold is taken directly from Fair's work, where he considered such growth to reflect meaningful economic improvement that voters might reward. For each presidential term, I identified the fifteen quarters that correspond to the incumbent term and counted how many of them surpassed this threshold.

The resulting values of g3, p15, and n for the modern elections differ somewhat from those reported in earlier versions of Fair's model. These differences arise entirely from updated data sources and from the fact that modern GDP and inflation series reflect methodological revisions that have accumulated over several decades. Rather than attempting to reverse engineer Fair's historical adjustments, I use these modern values directly and document the differences.

3.3.Constructing Political Variables for 1996 through 2024

The political variables for the modern period are not included in the original Fair dataset, so I constructed them manually. This required determining the party of the incumbent president, identifying whether the Democratic candidate was the sitting president, and computing the duration of the incumbent party in office. Each of these variables follows the definitions used in Fair's published work. The duration variable increases by one each time the incumbent party retains the presidency. If the presidency changes parties, the duration resets to zero.

These variables required careful consistency checks. I verified incumbency and party control using the historical election records provided by the American Presidency Project. I then matched these political variables to their corresponding economic variables in the final dataset.

3.4.Merging the Original and Extended Data

The last step was to merge the original Fair dataset with the newly constructed data for the modern elections. To do this, I aligned the variables by election year and replaced missing

values in the original dataset with the newly computed values for the extended period. I also created adjusted versions of p15 and n for the three war elections identified by Fair, in which he intentionally set these variables to zero.

This merge produced a complete dataset covering the period from 1916 through 2024. The dataset includes the original Fair values where available and the reconstructed values for the modern era. The final result is a transparent and reproducible version of the voting equation dataset suitable for both replication and extension. With the dataset assembled and all variables constructed, the next step is to replicate Fair's original regression results for the period from 1916 through 1992. This allows me to verify whether my reconstructed dataset produces the same qualitative patterns found in Fair's published work.

4.REPLICATION RESULTS

This section presents the results of my replication of the Fair voting equation using the original estimation period from 1916 through 1992. The goal is to reproduce Fair's published coefficient estimates as closely as possible and to examine how well the model captures historical presidential election outcomes. I also evaluate the accuracy of the model's fitted values and prediction errors within the original sample period. These steps are essential for assessing whether the Fair equation performs as expected before extending the model to modern elections.

4.1 Regression Estimates for 1916 through 1992

The replicated regression produces coefficient estimates that match the general structure and direction of Fair's published results. The signs of the coefficients are consistent with the economic and political interpretation of the model. In particular, the incumbency term enters with a positive coefficient, which reflects the historical advantage of Democratic incumbents during much of the twentieth century. The economic variables behave in the expected manner. Higher recent growth increases the predicted Democratic vote share. Higher inflation reduces it. A greater number of strong growth quarters increases support for the incumbent party.

Despite these intuitively reasonable signs, the standard errors remain large. Most coefficients are not statistically significant at conventional levels. This is not a shortcoming of the replication. It is a feature of the original Fair model and is something Fair explicitly discusses in his own work. The small sample size and the reliance on presidential level national data limit statistical precision. As a result, the model focuses more on capturing broad relationships than on identifying tightly estimated structural parameters.

The constant term carries most of the explanatory power. The R squared value for the replication is approximately forty one percent, which is comparable to Fair's reported value. This indicates that the voting equation explains general patterns but does not account for

unexplained shocks or idiosyncratic events that influence election outcomes. The replication confirms that the model performs as Fair intended and that the data structure and computations are correct.

Regression 1916–1992

VARIABLE	COEFFICIENT	STD ERROR	T-STAT
CONST	0.4916	0.0381	12.8943
I	0.0621	0.0478	1.2996
DPER	0.0064	0.0448	0.1436
g3	0.0017	0.0033	0.5067
p15_adj	-0.0018	0.0079	-0.2298
n_adj	0.0006	0.0054	0.1113
DUR	-0.0650	0.0351	-1.8535

TABLE 2A

Regression 1916–1988

VARIABLE	COEFFICIENT	STD ERROR	T-STAT
CONST	0.4635	0.0525	8.8280
I	0.0336	0.0604	0.5567
DPER	0.0375	0.0601	0.6240
g3	-0.0001	0.0040	-0.0327
p15_adj	-0.0005	0.0082	-0.0585
n_adj	0.0040	0.0069	0.5778
DUR	-0.0423	0.0457	-0.9241

TABLE 2B

Regression 1916–1960

VARIABLE	COEFFICIENT	STD ERROR	T-STAT
CONST	0.4281	0.0700	6.1155
I	-0.0311	0.0836	-0.3723
DPER	0.1010	0.0980	1.0309
g3	0.0002	0.0069	0.0288
p15_adj	0.0262	0.0153	1.7111
n_adj	-0.0018	0.0104	-0.1773

DUR	0.0060	0.0560	0.1072

TABLE 2C

4.2 In Sample Fit and Prediction Errors

To evaluate how well the model fits the historical sample, I computed fitted values and residuals for each election. The prediction error for a given year is the difference between the model's predicted Democratic two party vote share and the actual outcome. These errors highlight the strengths and weaknesses of the model during the estimation period.

The results align closely with the findings presented by Fair. The model performs reasonably well in stable political and economic environments, but it struggles during elections marked by severe disruptions. For example, the model substantially underpredicts the Democratic vote in 1932, when the Great Depression dominated every aspect of political life. Similarly, the model misses the dramatic shift in 1980, when rising inflation and dissatisfaction with the incumbent party led to a large forecasting error. These years expose the limits of a purely linear economic voting model when confronted with rare or extreme events.

At the same time, the errors for more typical elections remain modest. Several predicted values fall within two or three percentage points of the actual results. This pattern reinforces Fair's argument that although the model is imperfect, it captures underlying tendencies in the relationship between macroeconomic conditions and presidential vote share.

Insert Table 3 here.

4.3 Interpretation of the Replication Results

The replication confirms that the Fair model behaves consistently with its published version. The direction of each coefficient, the magnitude of the standard errors, and the pattern of in sample errors all match the expected structure of the voting equation. Importantly, the replication did not uncover any computational mistakes or irregularities in the original sample period. This provides confidence that the model extension performed later in the paper begins from a correct and faithful reconstruction of Fair's work.

The results also highlight a central challenge of presidential forecasting. Even when the model is correctly specified, economic variables alone cannot explain all electoral outcomes. Many historical elections were shaped by wars, charismatic candidates, shifts in political coalitions, or major national crises. These factors create variation that a simple economic model cannot fully capture. The replication results therefore serve as a benchmark against which the model's performance in modern elections can be evaluated. While the replication results provide a baseline for comparison, the central question of this project is whether the Fair model can explain modern presidential elections. To address this question, I extend the dataset forward to 2024 and evaluate the model out of sample

YEAR	ACTUAL V	PREDICTED V	ERROR
1916	0.5168	0.5586	0.0418
1920	0.3612	0.4697	0.1085
1924	0.4176	0.4133	-0.0043
1928	0.4118	0.5060	0.0942
1932	0.5916	0.4682	-0.1234
1936	0.6246	0.5803	-0.0443
1940	0.5500	0.5029	-0.0471
1944	0.5377	0.4837	-0.0540
1948	0.5237	0.4678	-0.0559
1952	0.4460	0.4416	-0.0044
1956	0.4224	0.4204	-0.0020
1960	0.5009	0.4934	-0.0075
1964	0.6134	0.5727	-0.0407
1968	0.4960	0.4962	0.0002
1972	0.3821	0.4275	0.0454
1976	0.5105	0.4900	-0.0205
1980	0.4470	0.5429	0.0959
1984	0.4083	0.4272	0.0189
1988	0.4610	0.4949	0.0339
1992	0.5345	0.4998	-0.0347

RMSE: 0.056557803627257 23

TABLE 3

5.EXTENSION TO MODERN ELECTIONS (1996 - 2024)

After replicating Fair's original model, the next step was to evaluate how well the same voting equation performs when applied to elections after 1992. Extending the model required careful construction of economic variables that mirror Fair's definitions from the early sample period. Because Fair's model uses quarterly measures of real output, inflation, and population, I reconstructed the variables using modern data from FRED and applied the same formulas used in Fair's work. Extending the dataset required several methodological decisions, since the availability and precision of postwar macroeconomic data differ from the historical series that Fair used. This section explains these decisions, presents the extended dataset, and interprets the out of sample forecasts for the elections from 1996 through 2024.

5.1 Construction of Economic Variables for Modern Years

To extend the model, I computed three economic indicators for each modern election year. These include the sum of real per capita GDP growth over the last three quarters before the election, the sum of inflation over the last fifteen quarters, and the number of strong growth quarters during the term of the incumbent administration.

The variable for growth, labeled g three, was constructed by taking the growth rate of real per capita GDP and summing the values from the last three quarters prior to November of each election year. This required identifying the third quarter of each election year, locating its index within the quarterly FRED data, and summing backward for two quarters. I applied a straightforward growth formula that expresses change in real per capita output in annualized percentage points, which keeps the construction consistent with Fair's published approach.

Inflation, labeled p fifteen, was constructed by summing fifteen consecutive quarterly inflation rates leading up to the election quarter. I calculated each quarterly inflation rate as the annualized percentage change in the GDP deflator, reflecting Fair's preference for a broad price index. Although the deflator is a modern construct, its properties are similar enough to Fair's earlier inflation series to allow for meaningful comparison.

Finally, the measure of good news quarters, labeled n, counts the number of quarters during the incumbent president's term in which real per capita GDP growth exceeds two point nine percent. This threshold is taken directly from Fair's model. Because modern output growth tends to be lower on average than early twentieth century growth, the interpretation of this threshold may be less straightforward today. However, I retained Fair's definition to preserve the structure of the original model.

All three variables were computed for the years 1996 through 2024, producing a complete economic dataset aligned with Fair's definitions.

5.2 Political Variables for Modern Elections

In addition to economic conditions, the Fair model incorporates three political indicators. These include the identity of the incumbent party, whether the incumbent party's nominee is the sitting president, and the duration of time the party has held the White House. I constructed these variables for each modern election based on historical records.

The variable I equals positive one when the Democratic Party holds the presidency and negative one when the Republican Party holds it. The variable DPER equals one when the Democratic nominee is also the sitting president. The duration variable DUR increases with the number of consecutive terms the incumbent party has held office. Because these variables have clear public definitions, they were straightforward to compute for the modern elections.

5.3 Merging and Preparing the Extended Dataset

After constructing both the economic and political variables, I merged them with the original dataset covering 1916 through 1992. The result is a unified dataset that spans more than a century of presidential elections. During this process I encountered missing early GDP data, particularly before 1947. I addressed this issue by retaining the historical values for the original sample period and applying the extended FRED based computations only to the modern years. This preserves the integrity of the original data while allowing the extended model to perform meaningful out of sample forecasts.

VARIABLE	DEFINITION
V	Democratic two-party vote share
I	Indicator for whether the incumbent party is Democratic (+1) or Republican (−1).
DPER	Indicates whether the Democratic candidate is the sitting president.
DUR	Incumbent party duration effect used by Fair to penalize long-tenure parties
g3	Sum of real per-capita GDP growth over the last three quarters before the election (Q4 of t−1 through Q2 of election year t).
p15	Sum of inflation over the 15 quarters leading up to the election.
n	Number of "good-news" quarters in the president's term, defined as quarters where real per-capita GDP growth exceeds 2.9%.

TABLE 1

YEAR	V	V_HAT	ERROR	g3	p15	n	I	DPER	DUR
1996	0.5474	0.4993	-0.0481	3.3005	0.0401	4	1	1	1
2000	0.5026	0.4283	-0.0743	2.0148	1.6449	7	1	0	2
2004	0.4876	0.4664	-0.0212	2.1833	2.1147	2	-1	1	1
2008	0.5368	0.2941	-0.2427	-1.3881	2.7189	3	-1	0	2
2012	0.5196	0.4962	-0.0234	1.1357	1.4718	2	1	1	1
2016	0.5110	0.4616	-0.0494	1.4083	1.2652	4	1	0	2
2020	0.5227	0.5022	-0.0205	-2.9862	1.7600	3	-1	1	1
2024	0.4925	0.4932	0.0007	1.6697	4.5891	6	1	1	1

TABLE 4

5.4 Out of Sample Forecast Results

To evaluate the predictive performance of Fair's model in the modern era, I applied the estimated coefficients from the 1916 through 1992 regression to the extended dataset covering 1996 through 2024. This produced predicted Democratic vote shares for each election and allowed me to compute the associated forecast errors.

The results show that the model does not generalize well to the contemporary political environment. In most of the modern elections the model underpredicts the Democratic vote share. This pattern appears in 1996, 2000, 2004, 2012, 2016, and 2020. The consistent direction of these errors suggests that the relationship between economic fundamentals and voter behavior has changed since the early sample period. Fair's linear specification appears to be missing structural changes that affect modern elections.

The most striking failure occurs in 2008. The model predicts a Democratic vote share of only twenty nine point four percent, compared with the actual value of fifty three point seven percent. This error of approximately twenty four points is by far the largest in the extended sample. The financial crisis of 2008 created economic and political dynamics that do not resemble earlier recessions. Voters assigned responsibility to the incumbent Republican administration, producing an outcome that the Fair model is not equipped to capture. This illustrates how extraordinary events with strong partisan associations can disrupt models that rely primarily on aggregate economic conditions.

In contrast, the forecasts for 2012, 2020, and 2024 fall relatively close to the actual outcomes. However, this accuracy does not appear to signal structural stability. Instead, it seems coincidental given the poor performance in the surrounding years. Taken together, the extended results demonstrate that the Fair voting equation performs well within its historical sample but does not maintain that level of accuracy when applied to modern presidential elections.

5.5 Interpretation of the Extended Forecasts

The out of sample performance shows that the link between macroeconomic performance and presidential voting has weakened over time. Structural changes in voter behavior, increased political polarization, the rise of identity based voting, new media environments, and crisis driven elections have likely altered the mechanisms that once connected economic fundamentals to vote shares. The Fair model was developed in an era when economic performance carried greater political weight, and its predictive limitations in the modern period are consistent with the broader literature on economic voting. The extended results therefore provide not only a technical evaluation but also a substantive insight into how the determinants of presidential elections have evolved.

The results of the modern extension raise broader interpretive questions about stability, structural change, and the long-run usefulness of economic voting models. The following

section discusses these issues and places the results in a wider political and methodological context.

6.DISCUSSION

The results of the replication and extension exercise highlight both the strengths and weaknesses of the Fair voting equation as a tool for understanding presidential elections. Within the original estimation window of 1916 through 1992, the model performs reasonably well in capturing broad movements in Democratic vote share. The direction of the coefficients generally aligns with Fair's theoretical expectations: positive economic growth increases support for the incumbent party, inflation reduces support, and longer incumbent-party tenure introduces a penalty. Even though many coefficients are statistically insignificant, the model still reproduces the qualitative insights that Fair emphasized in his own work.

At the same time, the replication process made clear that the Fair model is very sensitive to the construction of its economic variables. The values of g_3, p_{15}, and n depend entirely on how quarterly growth and inflation are measured and on the availability of long historical data. Because FRED only provides post–1947 quarterly GDP and inflation, early values must either be approximated or left missing, which affects the comparability between the early and later parts of the sample. Fair himself used historical sources that are not easily reconstructed, and this likely explains why some of the early observations in my dataset produce slightly different patterns than in his original tables.

The most important finding comes from extending the model to elections after 1992. Once the model is applied to modern presidential elections, its predictive accuracy deteriorates sharply. The systematic underprediction of Democratic vote share across many post-1992 elections suggests a structural shift in how voters respond to economic conditions. The large forecast error in 2008 is especially revealing. The financial crisis represented an economic shock of a kind that the linear Fair model is not designed to capture. Rather than punishing the out-party, voters placed responsibility on the incumbent Republicans. This dynamic contradicts the simple form of retrospective voting embedded in Fair's equation.

The broader pattern suggests that the relationship between economics and voting behavior has changed since the early 1990s. Rising partisan polarization, increasing nationalization of elections, and greater influence of identity-driven political preferences all reduce the predictive power of economic fundamentals. In earlier decades, elections behaved more like referenda on economic performance, but more recent elections appear less tied to short-run macroeconomic indicators and more tied to partisan loyalty. The extension results reinforce this shift, as the model fails to capture the more complex way modern voters respond to economic and political information.

Finally, through this replication process, I learned that small differences in data definitions and coding choices accumulate in meaningful ways. Constructing macroeconomic variables, especially for early years, requires a degree of judgment that can alter the model's performance. These decisions show why replication matters: even a relatively simple equation like Fair's is sensitive to measurement decisions that are often left implicit in published work. My results therefore provide both a validation of the broad logic of Fair's model and a

demonstration of its limitations when applied outside its original historical context. Taken together, these findings point toward several conclusions about how the Fair model behaves across different historical periods and what lessons can be drawn for forecasting and political-economy research. I summarize these points in the conclusion.

7.Conclusion

This replication project set out to reproduce Ray Fair's influential presidential vote equation and to evaluate whether the relationships he identified during the early and middle parts of the twentieth century continue to hold in the modern political era. The project required assembling historical data, reconstructing Fair's economic variables, implementing his econometric model, and extending the dataset to include elections from 1996 through 2024. The process revealed both the strengths and the limitations of relying on economic fundamentals as predictors of presidential election outcomes.

The replication of the 1916 through 1992 sample successfully reproduced the general structure of Fair's results. The signs of the key coefficients aligned with theoretical expectations and with Fair's published estimates. Economic growth entered positively, inflation entered negatively, and the duration variable imposed a penalty on parties that had controlled the presidency for multiple terms. These results support the retrospective economic voting framework that underlies Fair's model. Although statistical significance was limited, the main pattern of coefficients suggests that economic conditions once played a consistent role in shaping presidential outcomes.

The extension to modern elections, however, produced a very different picture. When the original estimated coefficients were applied to the elections after 1992, the model consistently underpredicted the Democratic vote share. This pattern occurred across a variety of economic environments and political contexts. The most striking failure occurred in 2008, when the model predicted an outcome that bore little resemblance to the actual election result. The dramatic forecasting errors indicate that the structural relationships embedded in Fair's model do not extend naturally into the contemporary era.

The diminished predictive power of the model likely reflects substantial changes in the political environment rather than simple measurement issues. The rise of political polarization, increasing partisan loyalty, changing media landscapes, and the emergence of crisis driven or identity driven voting patterns all suggest that modern elections are shaped by forces beyond aggregate economic conditions. Economic performance still matters, but it no longer appears to be the dominant factor that it once was. As a result, models that rely exclusively on macroeconomic variables cannot fully account for the outcomes of recent presidential elections.

The replication process also highlighted the importance of methodological transparency and the influence of data construction decisions. Computing Fair's variables using modern datasets

required careful alignment of definitions and assumptions. While these choices may produce small differences from Fair's original values, they cannot account for the magnitude of the differences observed in the out of sample period. Instead, the results point to genuine structural changes in voter behavior that any modern forecasting model must address directly.

Generally, this project demonstrates the value of replication as a research tool. Reproducing Fair's model required not only statistical work but also conceptual understanding, judgment in data construction, and critical evaluation of results. The replication confirms that Fair's model captured an important empirical regularity during much of the twentieth century. At the same time, the extended analysis shows that the forces governing presidential elections have evolved. This combination of replication and critique provides a deeper understanding of both historical voting behavior and the challenges of forecasting in a rapidly changing political environment.

APPENDIX

VARIABLE	DEFINITION
V	Democratic two-party vote share
I	Indicator for whether the incumbent party is Democratic (+1) or Republican (−1).
DPER	Indicates whether the Democratic candidate is the sitting president.
DUR	Incumbent party duration effect used by Fair to penalize long-tenure parties
g3	Sum of real per-capita GDP growth over the last three quarters before the election (Q4 of t−1 through Q2 of election year t).
p15	Sum of inflation over the 15 quarters leading up to the election.
n	Number of "good-news" quarters in the president's term, defined as quarters where real per-capita GDP growth exceeds 2.9%.

TABLE 1

VARIABLE	COEFFICIENT	STD ERROR	T-STAT
CONST	0.4916	0.0381	12.8943
I	0.0621	0.0478	1.2996
DPER	0.0064	0.0448	0.1436
g3	0.0017	0.0033	0.5067
p15_adj	-0.0018	0.0079	-0.2298
n_adj	0.0006	0.0054	0.1113
DUR	-0.0650	0.0351	-1.8535

TABLE 2A

VARIABLE	COEFFICIENT	STD ERROR	T-STAT
CONST	0.4635	0.0525	8.8280
I	0.0336	0.0604	0.5567
DPER	0.0375	0.0601	0.6240
g3	-0.0001	0.0040	-0.0327

p15_adj	-0.0005	0.0082	-0.0585
n_adj	0.0040	0.0069	0.5778
DUR	-0.0423	0.0457	-0.9241

TABLE 2B

VARIABLE	COEFFICIENT	STD ERROR	T-STAT
CONST	0.4281	0.0700	6.1155
I	-0.0311	0.0836	-0.3723
DPER	0.1010	0.0980	1.0309
g3	0.0002	0.0069	0.0288
p15_adj	0.0262	0.0153	1.7111
n_adj	-0.0018	0.0104	-0.1773
DUR	0.0060	0.0560	0.1072

TABLE 2C

YEAR	ACTUAL V	PREDICTED V	ERROR
1916	0.5168	0.5586	0.0418
1920	0.3612	0.4697	0.1085
1924	0.4176	0.4133	-0.0043
1928	0.4118	0.5060	0.0942
1932	0.5916	0.4682	-0.1234
1936	0.6246	0.5803	-0.0443
1940	0.5500	0.5029	-0.0471
1944	0.5377	0.4837	-0.0540
1948	0.5237	0.4678	-0.0559
1952	0.4460	0.4416	-0.0044
1956	0.4224	0.4204	-0.0020
1960	0.5009	0.4934	-0.0075
1964	0.6134	0.5727	-0.0407
1968	0.4960	0.4962	0.0002
1972	0.3821	0.4275	0.0454
1976	0.5105	0.4900	-0.0205
1980	0.4470	0.5429	0.0959
1984	0.4083	0.4272	0.0189
1988	0.4610	0.4949	0.0339
1992	0.5345	0.4998	-0.0347

RMSE: 0.05655780362725723

TABLE 3

YEAR	V	V_HAT	ERROR	G3	P15	N	I	DPER	DUR
1996	0.5474	0.4993	-0.0481	3.3005	0.0401	4	1	1	1
2000	0.5026	0.4283	-0.0743	2.0148	1.6449	7	1	0	2
2004	0.4876	0.4664	-0.0212	2.1833	2.1147	2	-1	1	1
2008	0.5368	0.2941	-0.2427	-1.3881	2.7189	3	-1	0	2
2012	0.5196	0.4962	-0.0234	1.1357	1.4718	2	1	1	1
2016	0.5110	0.4616	-0.0494	1.4083	1.2652	4	1	0	2
2020	0.5227	0.5022	-0.0205	-2.9862	1.7600	3	-1	1	1
2024	0.4925	0.4932	0.0007	1.6697	4.5891	6	1	1	1

TABLE 4

References

Fair, Ray C. *The Fair Model*. Yale University, https://fairmodel.econ.yale.edu.

American Presidency Project
(All have the same corporate author and publisher)

"Election of 2024." *The American Presidency Project*, University of California, Santa Barbara, https://www.presidency.ucsb.edu/statistics/elections/2024.

"Election of 2020." *The American Presidency Project*, University of California, Santa Barbara, https://www.presidency.ucsb.edu/statistics/elections/2020.

"Election of 2016." *The American Presidency Project*, University of California, Santa Barbara, https://www.presidency.ucsb.edu/statistics/elections/2016.

"Election of 2012." *The American Presidency Project*, University of California, Santa Barbara, https://www.presidency.ucsb.edu/statistics/elections/2012.

"Election of 2008." *The American Presidency Project*, University of California, Santa Barbara, https://www.presidency.ucsb.edu/statistics/elections/2008.

"Election of 2004." *The American Presidency Project*, University of California, Santa Barbara, https://www.presidency.ucsb.edu/statistics/elections/2004.

"Election of 2000." *The American Presidency Project*, University of California, Santa Barbara, https://www.presidency.ucsb.edu/statistics/elections/2000.

"Election of 1966." *The American Presidency Project*, University of California, Santa Barbara, https://www.presidency.ucsb.edu/statistics/elections/1966.

FRED Economic Data
"GDPDEF: Gross Domestic Product Deflator." *FRED, Federal Reserve Bank of St. Louis*, https://fred.stlouisfed.org/series/GDPDEF.

"B230RC0Q173SBEA: Civilian Population." *FRED, Federal Reserve Bank of St. Louis*, https://fred.stlouisfed.org/series/B230RC0Q173SBEA.

"GDPC1: Real Gross Domestic Product." *FRED, Federal Reserve Bank of St. Louis*, https://fred.stlouisfed.org/series/GDPC1.

Fair, Ray C. "The Effect of Economic Variables on Presidential Elections." 1996. PDF file. *1996B200.PDF*.