Econometrics and Elections

Justus Hummels

November 24, 2025

Contents

1 Introduction

This document provides a summary of an independent research project based on Professor Fair's work on US presidential election prediction. It is not intended to be a full research paper, but rather a research proposal that outlines my initial attempts to understand, replicate, and extend Professor Fair's work.

I was first introduced to Professor Fair's election prediction model in an econometrics course with Professor Altonji. I returned to the model this year in light of the U.S. presidential election last year and the German parliamentary election this year, which renewed my interest in the question of what drives voting behavior. I was struck by how well the Fair model performs with relatively few and purely economic explanatory variables, which motivated me to study it more carefully and explore whether it could be applied to other political contexts. While this project primarily satisfies my intellectual curiosity, it is also intended as the starting point for a larger project that could eventually develop into a senior thesis or a more formal research paper.

In this write-up, I summarize my methodology and results from applying the Fair model first to the New York City mayoral election and then to Germany's federal elections. In both cases, I modified the model to account for limited data availability, differing political systems, and context-specific economic indicators. Consistent with the ideas of the original model, I focused on whether macroeconomic conditions can predict vote shares. The choice of the two specific contexts (NYC and Germany) is grounded in my personal interest and it's relevance in the news.

My models, unfortunately, have relatively weak explanatory and predictive power. This is likely due to bad data, and the institutional differences across contexts. Although the models I estimate are not strong and the results are not groundbreaking, the exercise has helped clarify the opportunities and obstacles involved in extending the Fair model to new settings. I hope these preliminary results may serve as a starting point for further research and possibly a conversation about how best to adapt the model beyond U.S. presidential elections.

2 Methodology

Across both contexts, my approach follows the core structure of Professor Fair's election prediction model: estimating the vote share of one candidate (whether a political party or a single individual) as a function of macroeconomic indicators and incumbency effects. For each case, I constructed a simplified version of the model based on the available economic indicators and the specific institutional setting.

Because both datasets include relatively few elections which leaves me with little data points for the dependent variable, I restricted the specifications to a very small set of predictors to reduce the risk of overfitting. All models were estimated using OLS, and I used heteroskedasticity- and autocorrelation-consistent (HAC) standard errors to account for serial correlation and non-constant variance in the error term, which I found to be both reasonable concerns in short, historical political time series. For each application, I describe below the data and my modeling choices.

3 New York Mayoral Elections

3.1 The Data

For New York City, I assembled election results beginning in 1897, yielding approximately 35 observations through 2025 (not including 2025, as I started this project before the recent election with the intention to predict it). While this provides a long time horizon for municipal-level analysis, the availability of historical economic data is much more limited. Reliable local macroeconomic indicators prior to the postwar period are scarce, and constructing a long-run series would require either deeper research or a retrospective reconstruction model. These remain possible directions for future work.

A central conceptual question is whether local or national economic conditions matter more for municipal elections. Both approaches have merit. Using local economic indicators aligns closely with the logic of the Fair model, which assumes that voters evaluate an administration based on whether it has tangibly improved economic conditions. Using federal indicators, by contrast, sheds light on whether national trends influence local voting behavior—i.e., whether New Yorkers vote in step with federal sentiment or independently of it.

Due to the lack of historical local GDP or income data, I relied on federal indicators for economic growth. In a next stage of the project, I could examine the correlation between federal and local economic performance or seek out better local historical data. For inflation, however, I used changes in local CPI. Initially, I was hoping to use rent inflation but I had to discard the idea due to limited data availability. Nevertheless, housing costs are part of the CPI as well. Inflation and rent levels have been central to New York City politics in recent years, with rent constituting a large share of household expenditures, and having the ability to be directly influenced by local policy (rent control, zoning, urban planning).

In total, after restricting the dataset to years for which both voting and economic data were available, I obtained 21 usable observations. I defined V_t as the vote share of the Democratic candidate, who has effectively functioned as the dominant incumbent party in modern NYC politics, that is similarly bipartisan like the US presidential election system. A more serious research study should think about the role of strong independent candidates and changing party affiliations when constructing V_t.

To construct the economic growth variable G, I relied on quarterly real GDP data from Professor Fair's website. In the original Fair model, G is defined as the growth rate of real GDP per capita over the first three quarters of the election year. The logic is that voters form their economic assessments based on conditions leading up to the election, not on data released afterward.

New York City mayoral elections are always held in November, which aligns cleanly with this definition. Therefore, for each election year, I computed G using real GDP growth between the first quarter (Q1) and the third quarter (Q3) of that same year. This preserves consistency with the timing assumptions in Professor Fair's framework.

To construct P and R, I computed the annual average of CPI and rent inflation of the first 15 quarters of a legislation, following Fair's methodology.

3.2 The Model

Given the small sample size and limited set of economic indicators, I estimated a model using only two predictors: real income (GDP) growth (G_t) and rent inflation (P_t). The model is specified as

$$V_t = \alpha + \beta_1 G_t + \beta_2 P_t + u_t,$$

where V_t is the vote share of the Democratic candidate, G_t captures economic growth, and P_{It} measures local inflation pressure. I used HAC standard errors to correct for potential serial correlation and heteroskedasticity in the residuals.

4 German Federal Elections

4.1 The Data

For Germany, usable data begins in the early 1950s due to the country's postwar political reconstruction. While the first federal election took place in 1949, I begin the analysis in 1953 because my dependent variable is the vote share of the governing (chancellor's) party. Constructing this variable requires information from the previous election period, which makes 1949 unusable as a starting point.

The German political system is not bipartisan, which makes a direct analogue to the U.S. system impossible. To adapt the Fair framework, I defined

V_t as the vote share of the governing party, allowing an assessment of whether voters reward or punish incumbents for economic performance.

The economic data available are generally reliable but not always quarterly, complicating the construction of the growth variable G and the inflation variable P in the precise manner used by Professor Fair. Further research could likely uncover higher-frequency or archival datasets that would permit a closer replication of Fair's original methodology.

Structural breaks pose additional challenges: the division between East and West Germany, the economic dynamics of reunification in 1990, and subsequent transformations in party competition all introduce discontinuities that could meaningfully affect both voting behavior and macroeconomic indicators. Future work might incorporate those heterogeneity and discontinuities into the model. Another potential direction is to explore alternative dependent variables more analogous to a two-party system, such as vote shares of the governing coalition or broader ideological blocs.

4.2 The Model

As in the NYC application, I used a deliberately simple specification with two predictors: real GDP growth (G_t) and inflation (P_t). With only nineteen elections since 1953, adding more predictors would almost certainly result in overfitting.

The model is

$$V_t = \alpha + \beta_1 G_t + \beta_2 P_t + u_t.$$

HAC standard errors were again used to address potential autocorrelation and heteroskedasticity in the error term. Given the limited number of observations and the institutional differences between German and U.S. elections, the results should be interpreted cautiously and viewed primarily as exploratory.

5 Results and Discussion

5.1 New York City Mayor Elections

This section presents the results of the regressions using New York City mayoral election data. Additionally to the specified model earlier, I present additional models that use the incumbancy dummy variable and its interaction with the inflation and growth variable. I also construct a model with rent data. Unfortunately, my models have very low explanatory power. Given the limited number of observations and the small set of predictors available, this is unsurprising. With a larger dataset, more covariates capturing local political and demographic conditions could likely improve explanatory power. If with better data the results remain insignificant, it can also be concluded that the chosen indicators (macroeconomic indicators) don't have an effect on local elections.

Regression with G and P. The baseline regression including GDP growth (G) and inflation (P) shows almost no predictive power. The adjusted R^2 is negative and neither coefficient is statistically significant. The results are summarized below:

Dep. Variable:	democratic_pct	**R-squared:**	0.037
Model:	OLS	**Adj. R-squared:**	-0.070
Method:	Least Squares	**F-statistic:**	0.8468
Date:	Sun, 23 Nov 2025	**Prob (F-statistic):**	0.445
Time:	16:30:08	**Log-Likelihood:**	-82.813
No. Observations:	21	**AIC:**	171.6
Df Residuals:	18	**BIC:**	174.8
Df Model:	2		
Covariance Type:	HAC		

	coef	**std err**	**z**	**P>|z|**	**[0.025**	**0.975]**
const	52.2635	6.580	7.942	0.000	39.366	65.161
G	-0.5731	0.532	-1.077	0.281	-1.616	0.470
P	0.4997	1.634	0.306	0.760	-2.702	3.702

Regression with G and R (13 observations). Restricting the regression to the sample for which rent inflation (R) is available reduces the dataset to only 13 observations. This specification again yields no significant coefficients and a negative adjusted R^2, underscoring the difficulty of achieving reliable inference with so few data points. This specification of the model could make sense if you assume that economic indicators only affect voting behavior when a candidate runs again, i.e. is an incumbent.

Dep. Variable:	democratic_pct	**R-squared:**	0.059
Model:	OLS	**Adj. R-squared:**	-0.130
Method:	Least Squares	**F-statistic:**	0.7554
Date:	Sun, 23 Nov 2025	**Prob (F-statistic):**	0.495
Time:	16:31:00	**Log-Likelihood:**	-52.400
No. Observations:	13	**AIC:**	110.8
Df Residuals:	10	**BIC:**	112.5
Df Model:	2		
Covariance Type:	HAC		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	52.4971	15.085	3.480	0.001	22.930	82.064
G	-1.1005	1.174	-0.937	0.349	-3.402	1.201
R	1.1889	3.861	0.308	0.758	-6.378	8.756

Regression with only interaction terms G_I and P_I. The model using only the interaction terms G_I and P_I delivers a markedly higher adjusted R^2 of 0.371, and both coefficients are statistically significant. However, this specification is likely affected by omitted variable bias: by excluding the main effects of G, P, and incumbency (I), the interaction terms absorb their influence, preventing meaningful interpretation.

Dep. Variable:	democratic_pct	**R-squared:**	0.475
Model:	OLS	**Adj. R-squared:**	0.371
Method:	Least Squares	**F-statistic:**	12.90
Date:	Sun, 23 Nov 2025	**Prob (F-statistic):**	0.00170
Time:	16:32:10	**Log-Likelihood:**	-48.598
No. Observations:	13	**AIC:**	103.2
Df Residuals:	10	**BIC:**	104.9
Df Model:	2		
Covariance Type:	HAC		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	53.3217	3.113	17.128	0.000	47.220	59.423
G_I	3.6504	1.374	2.657	0.008	0.958	6.343
P_I	2.1476	0.486	4.421	0.000	1.196	3.100

Full interaction model. The full model including all main effects and interaction terms yields the highest unadjusted R^2, but once again the adjusted R^2 falls substantially (0.278), reflecting overfitting due to the small sample size relative to the number of predictors. Several coefficients appear statistically significant, but the model is unlikely to generalize.

Dep. Variable:	democratic_pct	**R-squared:**	0.579
Model:	OLS	**Adj. R-squared:**	0.278
Method:	Least Squares	**F-statistic:**	9.216
Date:	Sun, 23 Nov 2025	**Prob (F-statistic):**	0.00550
Time:	16:33:45	**Log-Likelihood:**	-47.171
No. Observations:	13	**AIC:**	106.3
Df Residuals:	7	**BIC:**	109.7
Df Model:	5		
Covariance Type:	HAC		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	74.0213	12.068	6.134	0.000	50.369	97.673
G	-1.2641	1.306	-0.968	0.333	-3.823	1.295
P	-6.7372	3.539	-1.904	0.057	-13.674	0.200
G_I	3.7964	2.230	1.703	0.089	-0.574	8.166
P_I	8.5592	2.948	2.903	0.004	2.780	14.338
I	-17.8191	8.689	-2.051	0.040	-34.849	-0.790

5.2 Leave-One-Out Cross-Validation Results

Given the low explanatory power of all model specifications, it is unsurprising that the Leave-One-Out Cross-Validation (LOOCV) results confirm the limited predictive ability of the regressions. Across specifications, the models perform poorly out-of-sample, with RMSE values well above a reasonable threshold for election forecasting and negative LOOCV R^2 values indicating that the models predict worse than a simple mean benchmark.

For the most extensive model, the LOOCV RMSE is 11.20 and the LOOCV R^2 is -0.156. The more restrictive model using only the interaction terms G_I and P_I performs similarly, with a LOOCV RMSE of 10.83 and a LOOCV R^2 of -0.079.

Overall, the LOOCV results align with expectations: all specifications exhibit similarly poor predictive performance. The consistently high RMSE and negative LOOCV R^2 values reinforce the conclusion that the available data and model structure offer little forecasting value for the NYC mayoral election.

5.3 German Federal Elections

A key limitation of the German dataset is that macroeconomic indicators are available only at yearly frequency. To approximate short-run economic conditions, I constructed lagged values of GDP growth and inflation (up to three years). A correlation matrix of these variables provides an initial overview of potential relationships with the governing party's vote share.

Inflation does not display a meaningful correlation with vote share, even when lagged. In contrast, GDP growth with a one-year lag shows a relatively strong correlation of 0.669. Based on these correlations and the limited number of observations, I restrict the preferred specification to two predictors: lagged

GDP growth and lagged inflation. Including additional variables would risk overfitting.

	Jahr	G	P	vote_share	G_lag1	P_lag1	G_lag2	P_lag2	G_lag3	P_lag3	DUR	DUR_dummy
Jahr	1.000000	-0.711534	-0.097810	-0.840480	-0.789002	-0.324930	-0.618384	-0.227927	-0.607932	-0.042927	-0.041255	-0.055239
G	-0.711534	1.000000	0.061165	0.561216	0.578390	0.067694	0.385583	0.275050	0.295484	0.152115	0.276611	0.223378
P	-0.097810	0.061165	1.000000	0.110728	-0.222688	0.617834	-0.193371	0.050449	0.349128	0.446107	0.119663	0.167132
vote_share	-0.840480	0.561216	0.110728	1.000000	0.669436	0.264332	0.555958	-0.041249	0.602743	-0.267467	0.104460	0.259264
G_lag1	-0.789002	0.578390	-0.222688	0.669436	1.000000	-0.131453	0.712361	-0.117098	0.427597	-0.331664	0.045432	0.055769
P_lag1	-0.324930	0.067694	0.617834	0.264332	-0.131453	1.000000	0.015469	0.605093	0.432323	0.494749	-0.251049	-0.262794
G_lag2	-0.618384	0.385583	-0.193371	0.555958	0.712361	0.015469	1.000000	-0.014678	0.713968	-0.450959	-0.258165	-0.075907
P_lag2	-0.227927	0.275050	0.050449	-0.041249	-0.117098	0.605093	-0.014678	1.000000	0.174871	0.874698	-0.392371	-0.482400
G_lag3	-0.607932	0.295484	0.349128	0.602743	0.427597	0.432323	0.713968	0.174871	1.000000	-0.030228	-0.127173	-0.000303
P_lag3	-0.042927	0.152115	0.446107	-0.267467	-0.331664	0.494749	-0.450959	0.874698	-0.030228	1.000000	-0.115790	-0.205479
DUR	-0.041255	0.276611	0.119663	0.104460	0.045432	-0.251049	-0.258165	-0.392371	-0.127173	-0.115790	1.000000	0.742307
DUR_dummy	-0.055239	0.223378	0.167132	0.259264	0.055769	-0.262794	-0.075907	-0.482400	-0.000303	-0.205479	0.742307	1.000000

Baseline model with G_lag1 and P_lag1. The resulting model produces statistically significant coefficients for both predictors and an adjusted R^2 of 52.4%, indicating moderate explanatory power given the small sample. Both higher GDP growth and higher inflation in the year prior to the election are associated with increased vote shares for the incumbent government.

Dep. Variable:	vote_share	**R-squared:**	0.574
Model:	OLS	**Adj. R-squared:**	0.524
Method:	Least Squares	**F-statistic:**	15.48
Date:	Sun, 23 Nov 2025	**Prob (F-statistic):**	0.000148
Time:	17:47:22	**Log-Likelihood:**	-61.550
No. Observations:	20	**AIC:**	129.1
Df Residuals:	17	**BIC:**	132.1
Df Model:	2		
Covariance Type:	HAC		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	30.3911	3.061	9.929	0.000	24.392	36.390
G_lag1	1.6976	0.311	5.457	0.000	1.088	2.307
P_lag1	1.7917	0.663	2.704	0.007	0.493	3.091

Polynomial inflation model. Given the historically low-inflation environment in the Eurozone, I also estimated a specification in which inflation enters as a second-degree polynomial. This functional form allows for potential non-linearities, such as diminishing or increasing marginal effects of inflation on electoral support. The adjusted R^2 increases to 0.572, suggesting a somewhat improved fit, although the quadratic term is not statistically significant at conventional levels.

Dep. Variable:	vote_share	**R-squared:**	0.639
Model:	OLS	**Adj. R-squared:**	0.572
Method:	Least Squares	**F-statistic:**	14.32
Date:	Sun, 23 Nov 2025	**Prob (F-statistic):**	8.54e-05
Time:	17:46:41	**Log-Likelihood:**	-59.901
No. Observations:	20	**AIC:**	127.8
Df Residuals:	16	**BIC:**	131.8
Df Model:	3		
Covariance Type:	HAC		

	coef	**std err**	**z**	**P> |z|**	**[0.025**	**0.975]**
const	34.3967	2.720	12.648	0.000	29.066	39.727
G_lag1	1.9793	0.442	4.475	0.000	1.112	2.846
P_lag1	-2.8552	2.958	-0.965	0.335	-8.654	2.943
P_lag1_sq	0.7814	0.482	1.622	0.105	-0.163	1.726

Overall, the German results align more closely with theoretical expectations: economic growth appears to benefit incumbent parties, while inflation plays a secondary but potentially nonlinear role. Nevertheless, the limited number of observations requires cautious interpretation.

5.4 Leave-One-Out Cross-Validation Results

To evaluate the out-of-sample performance of the German election model, I applied leave-one-out cross-validation (LOOCV) to both specifications: the baseline model using G_{lag1} and P_{lag1}, and an extended specification including the squared term P^2_{lag1}.

For the baseline model, LOOCV produces a root mean squared error (RMSE) of approximately 6.066 percentage points and an R^2 of 0.43. An RMSE of more than six percentage points is clearly too large for an election-forecasting setting. These results therefore underline the limited predictive power of the specification given the small sample of post-war German federal elections.

Including the quadratic term does not improve predictive accuracy. For the extended model with P^2_{lag1}, the LOOCV RMSE increases slightly to 6.143, while the LOOCV R^2 decreases to 0.418. This indicates that introducing nonlinear structure adds noise relative to the available sample size, rather than capturing a meaningful curvature in how economic fundamentals relate to vote share.

Overall, the LOOCV results strengthen a key conclusion of this project: while the in-sample fit of these regressions may look acceptable, the out-of-sample performance remains weak. With fewer than twenty observations, limited economic variation across cycles, and potential omitted political variables, the models struggle to generalize. LOOCV thus highlights the need for richer datasets, additional predictors, or alternative modelling approaches to improve the reliability of election forecasting in this context.

6 Conclusion and Outlook

Overall, this project has been a valuable exercise in translating the logic of the Fair model into two distinct political contexts. Even though the empirical results themselves were not highly predictive, particularly in the NYC mayoral case, the process of specifying, estimating, and interpreting these models was good practice and provided good insights into both the strengths and the limits of econometric election forecasting.

By constructing models for the NYC mayoral elections and the German federal elections, I was able to test how well economic fundamentals correlate with voting results in settings characterized by very different data constraints. The German model showed somewhat stronger relationships, suggesting that the Fair model performs better on the national level. Moreover, local or regional elections require more modification. The sparse and heterogeneous data illustrated how small samples and missing variables quickly undermine predictive power.

Working through these challenges helped clarify several key lessons. First, the usefulness of such models depends critically on the availability, quality, and temporal depth of the data. More degrees of freedom would have allowed me to incorporate additional predictors and reduce the risk of omitted-variable bias. Second, I gained a clearer understanding of which types of economic indicators are theoretically relevant for election outcomes and how they should be used in the regressions. Third, I now have a practical sense of how sensitive these models are to specification choices.

Despite the limitations of the initial results, this document establishes a strong foundation for future work. I now know what kind of data I need to collect when applying the Fair model to other countries, subnational elections, or different time periods. I also have a replicable workflow for experimenting with alternative specifications. Going forward, I plan to further work on the German application and hope to improve my model by finding better data.

Overall, this project deepened my understanding of how political-economy models are built, why some models travel well across contexts while others do not, and how empirical limitations shape the reliability of forecasting exercises. These lessons will inform my future work as I refine the model and explore its application to other elections and datasets.