

Problem Set 7: Understanding the 2024 Election I

In this problem set, voting statistics from the 2020 and 2024 presidential elections are analyzed. In Part 1, voting statistics from the 2024 election are summarized and swing voters are discussed. Next, Part 2 conducts regression analysis of Biden and Harris voter statistics, Trump 2020 and 2024 statistics, and 3rd party 2020 and 2024 statistics, as well as scatterplots to further illustrate the relationships. Part 3 summarizes this analysis, which is supported by STATA code in the Appendix.

Part 1: Summary Statistics and Vote Swings

2.

Table 1: Voter Statistics by U.S. State and the District of Columbia

<i>Variable</i>	<i>Mean</i>	<i>SD</i>	<i>Median</i>
Harris_Pct	46.42	11.59	47.49
Trump_2024_Pct	51.49	11.87	50.92
Other_2024_Pct	2.09	0.79	1.92

4. Overall, the Republican party saw within-party swing votes in every state and Washington D.C. While this generally appeared as a 1-3 percent decrease in Democrat party swing votes and a similar increase in Republican votes, Texas (4.01%) New York (5.62%), New Jersey (5.45%), Texas (4.69%), saw the largest decrease in Democrat within-party swing votes. Utah (0.64%) was the only state that saw an increase in Democrat swing votes. The Republican trend roughly mirrored the Democrat totals, with no states reporting a decline in Republican swing votes. Overall, the majority of states reported a less than 1 percent change in swing votes from 2020 to 2024, except for Utah, which reported a 2.5% decline in votes and Nebraska, with a 1.32% decline.

7. California (-8.41%), Florida (-9.71%), Massachusetts (-8.69%), New Jersey (-10.21%), New York (-11.53%), Texas (-8.26%) experienced the biggest shift in votes for the Democrat party between 2020 and 2024. No state reported a positive change in Democratic vote margin between the two elections, with Washington state (-0.70%) coming the closest to achieving parity.

Bailey Williams
 Problem Set 7
 Due Date: August 4, 2025
 Word Count: 1164
 Submitted: August 4, 2025

Part 2: Regression Analysis

1.

Table 2: Regression Analysis of the Percentage of Biden Votes on the Function of the Percentage of Harris Votes

<i>Source</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>Number of obs</i> = 51
Model	6642.80	1	6642.80	<i>F</i> (1, 49) = 4122.88
Residual	78.95	49	1.61	<i>Prob > F</i> = 0.00
Total	6721.74	50	134.43	<i>R-squared</i> = 0.99
				<i>Adj R-squared</i> = 0.99
				<i>Root MSE</i> = 1.27

<i>Harris_Pct</i>	<i>Coefficient</i>	<i>Std. err.</i>	<i>t</i>	<i>P>t</i>	<i>P>t</i>	<i>[95% conf. interval]</i>
BidenPct	0.96	0.01	64.21	0.00	0.93	0.99
_cons	-0.06	0.75	-0.08	0.94	-1.56	1.44

Table 3: Regression Analysis of the Percentage of Trump 2020 Votes on the Function of the Percentage of Trump 2024 Votes

<i>Source</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>Number of obs</i> = 51
Model	6969.14	1	6969.14	<i>F</i> (1, 49) = 4889.77
Residual	69.84	49	1.43	<i>Prob > F</i> = 0.00
Total	7038.98	50	140.78	<i>R-squared</i> = 0.99
				<i>Adj R-squared</i> = 0.99
				<i>Root MSE</i> = 1.19

Bailey Williams
 Problem Set 7
 Due Date: August 4, 2025
 Word Count: 1164
 Submitted: August 4, 2025

<i>Trump_2024~t</i>	<i>Coefficient</i>	<i>Std. err.</i>	<i>t</i>	<i>P>t</i>	<i>[95% conf. interval]</i>	
TrumpPct2020	0.99	0.01	69.93	0.00	0.96	1.01
_cons	3.02	0.71	4.24	0.00	1.59	4.45

Table 4: Regression Analysis of the Percentage of 3rd Party 2020 Votes on the Function of the Percentage of 3rd Party 2024 Votes

<i>Source</i>	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>Number of obs</i>	<i>=</i>	<i>51</i>
Model	17.10	1	17.10	<i>F(1, 49)</i>	<i>=</i>	<i>58.47</i>
Residual	14.33	49	0.29	<i>Prob > F</i>	<i>=</i>	<i>0.00</i>
Total	31.43	50	0.63	<i>R-squared</i>	<i>=</i>	<i>0.54</i>
				<i>Adj R-squared</i>	<i>=</i>	<i>0.53</i>
				<i>Root MSE</i>	<i>=</i>	<i>0.54</i>

<i>Other_2024~t</i>	<i>Coefficient</i>	<i>Std. err.</i>	<i>t</i>	<i>P>t</i>	<i>[95% conf. interval]</i>	
OtherPct2020	0.72	0.09	7.65	0.00	0.53	0.91
_cons	0.50	0.22	2.23	0.03	0.05	0.94

2.

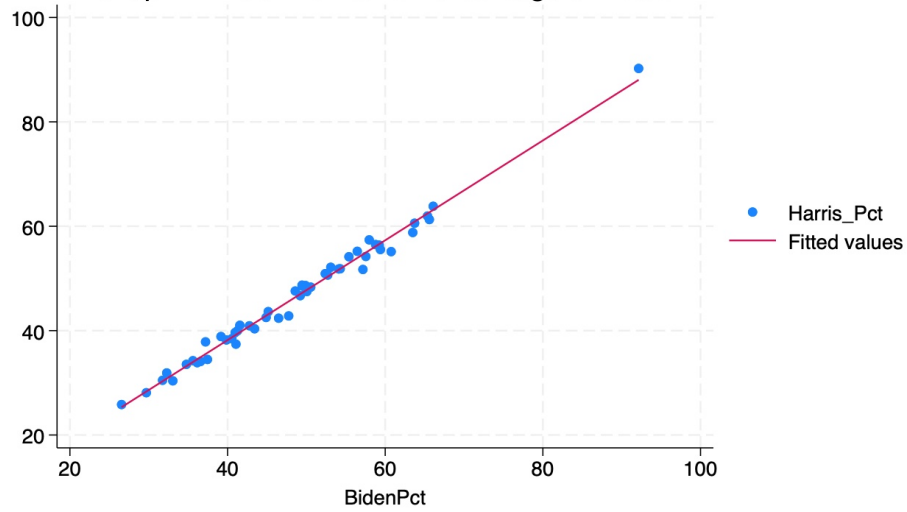
The direction in effect between the percentage of votes for Biden and percentage of votes for Harris is positive, meaning that voters that supported Biden in 2020 were likely to support Harris in 2024. The magnitude of effect is 0.96, meaning that for every 1 unit increase in Biden percentage, Harris percentage increased by 0.96 units. The R-squared value is 0.99, meaning that this is a very strong relationship, with 99 percent of the variation in Harris' vote percentage being explained as a function of Biden's vote percentage. The p-value is 0.00, leading me to reject the null hypothesis that there is no relationship between these percentages.

The direction in effect between the percentage of votes for Trump in 2020 and percentage of votes for Trump in 2024 is positive, meaning that voters that supported Trump in 2020 were likely to support Trump in 2024. The magnitude of effect is 0.99, meaning that for every 1 unit increase in Trump 2020 vote percentage, Trump's 2024 vote percentage increased by 0.99 units. The R-squared value is 0.99, meaning that this is a very strong relationship, with 99 percent of the variation in Trumps 2024 vote percentage being explained as a function of his 2020 vote percentage. The p-value is 0.00, leading me to reject the null hypothesis that there is no relationship between these percentages.

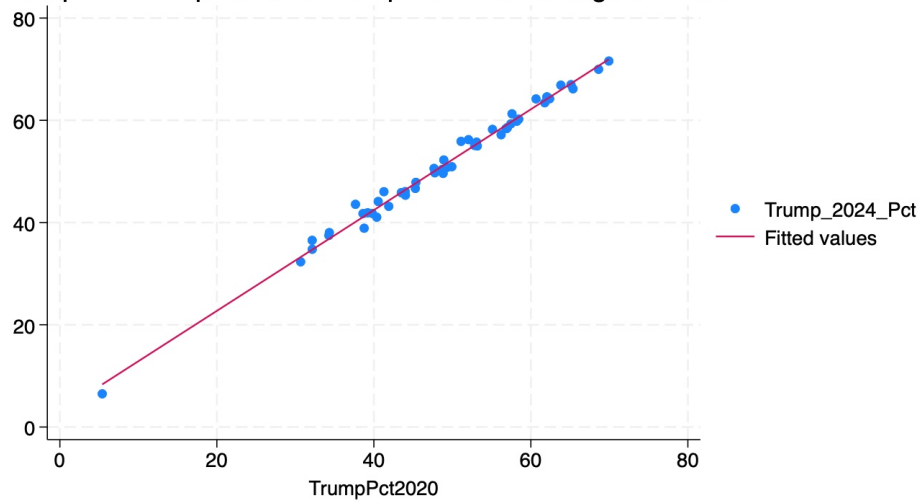
The direction in effect between the percentage of votes for 3rd parties in 2020 and percentage of votes for 3rd parties in 2024 is positive, meaning that voters that supported 3rd parties in 2020 were likely to support 3rd parties in 2024. The magnitude of effect is 0.72, meaning that for every 1 unit increase in 2020 3rd party vote percentages, 2024 3rd party vote percentages increased by 0.72 units. The R-squared value is 0.54, meaning that this is a medium-strength relationship, with 54 percent of the variation in 2024 3rd party vote percentages being explained as a function of 2020 3rd party vote percentages. The p-value is 0.00, leading me to reject the null hypothesis that there is no relationship between these percentages.

Bailey Williams
Problem Set 7
Due Date: August 4, 2025
Word Count: 1164
Submitted: August 4, 2025

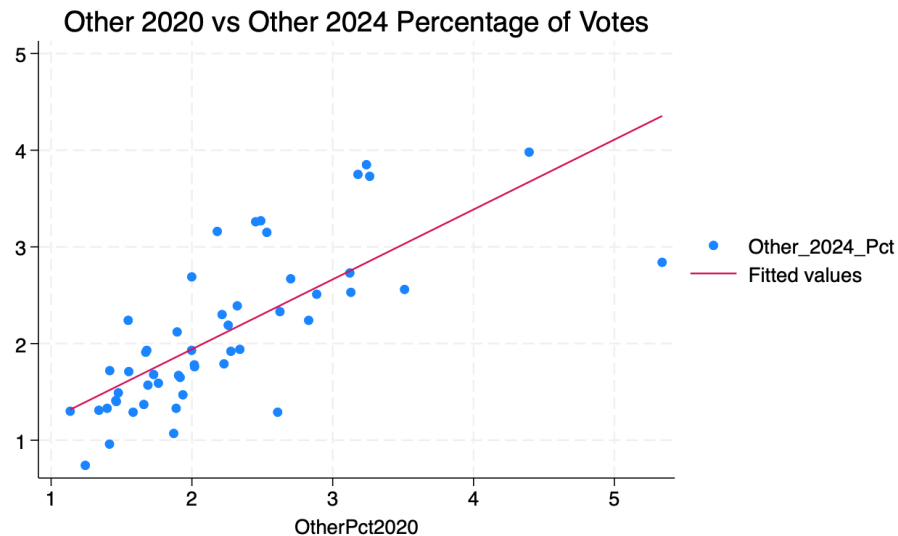
Graph 1: Biden vs Harris Percentage of Votes



Graph 2: Trump 2020 vs Trump 2024 Percentage of Votes



Bailey Williams
Problem Set 7
Due Date: August 4, 2025
Word Count: 1164
Submitted: August 4, 2025



Part 3: Assessment

Overall, there is a strong relationship between the results of the 2020 election and the 2024 election. Biden supporters were likely to support Harris, and those that voted for Trump in 2020 were likely to do so again. 3rd party voters in 2020 were also likely to vote 3rd party in 2024, however this relationship is not as strong as that of the major parties. While many Harris voters previously supported Biden, overall she underperformed Biden in each state. This can be seen in her home state of California, where she received 58.8% of votes compared to Biden's 63.49%. Looking at swing states, in 2020 Biden received 49.22% of votes in Arizona compared to 46.70% of votes by Harris in 2024. This decline is reflected in the other swing states as well. Overall, the Democrat party saw a net-negative swing in every state between 2020 and 2024, with the Republican party reporting the opposite. The most pronounced case of this was New York, who experienced an 5.62% decrease in Democrat voters in 2024 and a 5.91% increase in Republican voters. The total turnout decreased by 5,384,254 votes from 2020 (158,506,576 votes) to 2024 (153,122,322), which is likely what lead to a different result in 2024. Based on the voter swing margins for Democrats and Republicans, along with the decline in voter turnout in 2024, it is likely that Trump was able to beat Harris and not Biden due to a reduction in Democrat voters turning out combined with a rise in Republican voters.

Bailey Williams

Problem Set 7

Due Date: August 4, 2025

Word Count: 1164

Submitted: August 4, 2025

Bailey Williams
Problem Set 7
Due Date: August 4, 2025
Word Count: 1164
Submitted: August 4, 2025

Appendix

```
use "/Users/baileywilliams/Downloads/PS 7 Data-1.dta"

generate Total2020 = Biden_Total + Trump_Total_2020 + Other_Total_2020

generate BidenPct = Biden_Total/Total2020

generate TrumpPct2020 = Trump_Total_2020/Total2020

generate OtherPct2020 = Other_Total_2020/Total2020

replace BidenPct = BidenPct * 100

replace TrumpPct2020 = TrumpPct2020 * 100

replace OtherPct2020 = OtherPct2020 * 100

fsum Harris_Pct Trump_2024_Pct Other_2024_Pct, stats(mean p50 sd)

generate Dem_Swing = Harris_Pct - BidenPct

generate Rep_Swing = Trump_2024_Pct - TrumpPct2020

generate Other_Swing = Other_2024_Pct - OtherPct2020

generate Dem_Margin_2024 = Harris_Pct - Trump_2024_Pct

generate Dem_Margin_2020 = BidenPct - TrumpPct2020

generate Margin_Swing = Dem_Margin_2024 - Dem_Margin_2020

regress Harris_Pct BidenPct

regress Trump_2024_Pct TrumpPct2020

regress Other_2024_Pct OtherPct2020

tway (scatter Harris_Pct BidenPct) (lfit Harris_Pct BidenPct)

graph save "Graph" "/Users/baileywilliams/Downloads/Scatter1.gph"

graph export "/Users/baileywilliams/Downloads/Scatter1.jpg", as(jpg) name("Graph")
quality(90)

tway (scatter Trump_2024_Pct TrumpPct2020) (lfit Trump_2024_Pct TrumpPct2020)
```


Bailey Williams

Problem Set 7

Due Date: August 4, 2025

Word Count: 1164

Submitted: August 4, 2025

```
graph export "/Users/baileywilliams/Downloads/Scatter2.jpg", as(jpg) name("Graph")  
quality(90)
```

```
graph save "Graph" "/Users/baileywilliams/Downloads/Scatter2.gph"
```

```
twoway (scatter Other_2024_Pct OtherPct2020) (lfit Other_2024_Pct OtherPct2020)
```

```
graph export "/Users/baileywilliams/Downloads/Scatter3.jpg", as(jpg) name("Graph")  
quality(100)
```

```
graph save "Graph" "/Users/baileywilliams/Downloads/Scatter3.gph"
```

```
save "/Users/baileywilliams/Downloads/PS 7 Data-1.dta", replace
```