ARE STOCK RETURNS AND OUTPUT GROWTH HIGHER UNDER DEMOCRATS?

By

Ray C. Fair

March 2025

COWLES FOUNDATION DISCUSSION PAPER NO. 2277R1

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Are Stock Returns and Output Growth Higher Under Democrats?

Ray C. Fair*

March 2025

Abstract

Recent literature suggests that both stock returns and economic growth are significantly higher under Democratic presidential administrations. This is a puzzle in that persistent differences in stock returns seem unlikely in efficient markets, and it is not obvious why Democrats should do better. Often these kinds of results go away upon further analysis or more data, and this appears to be true in the present case. In this paper the sample is extended to 28 administrations, from Wilson-1 through Biden. While the mean stock return under the Democrats is higher, none of the differences in means is significant at conventional significance levels. There is considerable variation in the mean return across administrations, which results in lack of significance. Similarly, while the mean output growth rate under the Democrats is larger, the difference is not significant. Again, there is considerable variation in output growth across administrations. Results are also presented with the nine administrations between Hayes and Taft added, a total of 37 administrations. While the added data are likely not as good, the conclusion is the same—no significant differences. .

*Cowles Foundation, Department of Economics, Yale University, New Haven, CT 06520-8281. e-mail: ray.fair@yale.edu; website: *fairmodel.econ.yale.edu*. I am indebted to Alan Blinder, Jesse Shapiro, and Mark Watson for helpful comments.

1 Introduction

Recent literature suggests that both stock returns and economic growth are significantly higher under Democratic presidential administrations than under Republican presidential administrations in the United States. Regarding stock returns, Santa-Clara and Valkanov (2003) examined the period 1927–1998 and found significant differences. They found this a puzzle, since persistent differences seem unlikely in efficient markets. Pástor and Veronesi (2020) extended the sample period to 1927–2015 and also found significant differences. They argue that Democrats are more likely to get elected when risk aversion is high, which then mean-reverts during the administration.

Regarding output growth, Blinder and Watson (2016) examined the 16 administrations between Truman-2 and Obama-1. They found that output growth and other measures of economic activity are significantly higher under the Democrats. Pástor and Veronesi (2020) got similar results for the 1930–2015 period. The Blinder-Watson result was cited in the media—Leonhardt (2021)—at the time of the switch of administrations in 2021. After interviewing a number of economists, Leonhardt concluded that "much of the partisan gap remains mysterious." Cohan and Potrafke (2021) examine the 1949–2017 period and find significant output growth differences between Democrats and Republicans, including state governments. They also find the result puzzling.

Often these kinds of results go away upon further analysis or more data. This appears to be true in the present case. For the results in this paper the sample is extended to 28 administrations, from Wilson-1 through Biden. This period includes 15 Democratic administrations and 13 Republican. While mean returns across the administrations are higher under the Democrats, none of the differences in means is significant at conventional significance levels. There is considerable variation in the mean returns across administrations, which results in lack of significance. Similarly, while the mean output growth rate under the Democrats is

larger, the difference is not significant. Again, there is considerable variation in output growth across administrations. Section 6 contains results going back nine more administrations to Hayes, with the same conclusion.

The reason for the different conclusion in this paper is because of the use of the extended sample period. The results in Section 5 confirm the significant results of the previous studies using their sample periods.

2 The Tests

Consider stock returns first. Let R_t denote a measure of stock returns during administration t, where t runs from 1 to 28. The various measures of R_t are discussed in the next section. Let m_D and m_R denote the means of R_t over the Democratic (D) and Republican (R) administrations, respectively. Assuming that R_t is drawn from a normal distribution, it is straightforward to test the hypothesis that the means are equal, assuming either a common variance between the D and R observations or separate variances. For output growth, let G_t denote a performance measure of the economy during administration t, where again t runs from 1 to 28. Then treat G_t as R_t above and run the tests.

Another way to test for the equality of means is to regress R_t or G_t on a constant and D_t, where D_t is 1 if the administration is Democratic and 0 if Republican. The test of the equality of means is that the coefficient of D_t is zero, which is just a t-test. This is the same test as the equality of means test assuming common variance. If in this regression White's (1980) correction for heteroskedasticity is used, the test is the same as the equality of means test assuming separate variances.

In Section 4 both t-statistics are reported, one assuming common variance and one assuming separate variances. A third t-statistic is also reported, which uses the Newey-West (1987) correction for heteroskedasticity and autocorrelation with a lag of 2. Both Pástor and Veronesi (2020) and Blinder and Watson (2016) correct for heteroskedasticity and autocorrelation. As will be seen, the results are not

sensitive to which correction is used.

Another significance test is a non parametric test due to Blinder and Watson (2016). As discussed in Section 5, they analyzed 16 administrations, 9 R and 7 D, 16 observations on a variable G_t. There are 11,440 different ways in which 9 observations can be assigned to R and 7 to D. For each assignment compute the mean for R and the mean for D and take the difference. Compare the absolute value of this difference to the actual difference (the observed difference in the data). Then count up the number of times out of 11,440 that the absolute value of the computed difference is greater than the actual difference. The percent of times is the p-value for the null hypothesis that the R and D means are equal. It will be seen that this non parametric test gives very similar results to the others.

3 The Data

It will be seen that the results are not sensitive to different measures of stock returns. In the next section results are presented for six measures. The results are also not sensitive to the use of real GDP or per capita real GDP, although results for both measures are presented in the next section.

The observations are either monthly or quarterly. They begin in 1912 and go through 2024 for 28 administrations. Table 1 lists the raw-data variables for which observations were collected. The collection is discussed in the Appensix.

Table 1
The Raw-Data Variables: 1912–2024

Variable	Definition
DIV	S&P 500 Dividends (quarterly)
P	GDP Deflator (quarterly)
POP	Total U.S. Population (quarterly)
RS	Three-Month Treasury Bill Rate (monthly)
SP	S&P Stock Price Index (monthly)
VWD	CRSP Value-Weighted Return Including Dividends (monthly)
VWX	CRSP Value-Weighted Return Excluding Dividends (monthly)
Y	Real GDP (quarterly)

See the Appensix for discussion of the data collection.

The six measures of stock returns and two measures of output growth are discussed below. In addition, three other measures are discussed, which are needed in the construction of some of the other variables. All the (four-year) log differences have been divided by 4 to put them at an annual rate and multiplied by 100 to put them in percentage points.

- $R1$: From 1926 on the monthly data on SP are the prices on the last trading day of the month. Between 1912 and 1925 the prices are the average for the month. $R1$ is the log of SP at the end of December (or the average in December) of the fourth year of the administration minus the log of SP at the end of December (or the average in December) of the fourth year of the previous administration.

- $R2$: S&P 500 dividends, DIV, are assumed to have been invested in SP at the end of each quarter and cumulated. The value of SP used for each investment is the value at the end of the third month of the quarter. $R2$ is then the log of the value of the stock holdings at the end of December of the fourth year of the administration minus the log of SP at the end of December of the fourth year of the previous administration.

- ZRS: For this measure the monthly data on RS were converted to quarterly data by averaging the three months. It measures the return of investing each quarter in three-month Treasury bills and rolling them over throughout the administration. Using a value of 1.0 at the beginning, ZRS is the log of the value at the end.

- $R3$: $R2 - ZRS$. This is a measure of excess returns, where the risk free rate is taken to be the three-month Treasury bill rate.

- ZP: This measure is the log of P in the fourth quarter of the fourth year of the administration minus the log of P in the fourth quarter of the fourth year of the previous administration. It is a measure of the inflation rate over the four years of the administration.

- $R4$: $R3 - ZP$. This is a measure of real excess returns.

- $R5$: This is the return on the CRSP value-weighted stock portfolio excluding dividends in excess of the return on the three-month Treasury bill rate. The data on VWX are monthly returns. The values were accumulated over the 48 months of an administration. Using a value of 1.0 at the beginning, the log of the value at the end was taken. Then ZRS was subtracted from this value to get the excess return.

- $R6$: This is the return on the CRSP value-weighted stock portfolio including dividends in excess of the return on the three-month Treasury bill rate. The same procedure was followed for $R6$ as for $R5$ using VWD instead of VWX.

- $G1$: The data on Y are quarterly. $G1$ is the log of Y in the first quarter of the next administration minus the log of Y in the first quarter of the current administration. This is the main specification of Blinder and Watson (2016), which assumes that an administration is responsible for the first quarter

of the next administration. This modification was not made for the Biden administration because data for the first quarter of 2025 were not available at the time of this writing.

- $ZPOP$: This measure is the log of POP in the fourth quarter of the fourth year of the administration minus the log of POP in the fourth quarter of the fourth year of the previous administration.

- $G2$: $G1 - ZPOP$. This is the growth rate of per capital real GDP.

Table 2 presents observations for nine of the measures discussed above for the 28 administrations. These measures are for the four years of the administration, but they are at annual rates. An administration is denoted by its last year.

The simple nominal stock return, $R1$, varies from -31.45 percent for 1932 (Hoover) to 23.73 percent for 1936 (Roosevelt). For the administrations between 1988 and 2024 the return has been large except for 2004 and 2008 (G.W. Bush). Five of these ten administrations were Republican, including 2004 and 2008. The variability across administrations is large. As reported in Table 3 below, the standard deviation of $R1$ across the 28 administrations is 11.20 percent. A similar story holds for the other stock-return measures. Note that the real excess return, $R4$, is low for 2024 (Biden) because inflation, ZP, is high.

The growth rate, $G1$ varies from -8.94 percent for 1932 (Hoover) to 13.35 percent for 1944 (Roosevelt). $G2$ follows a similar pattern. Table 3 below shows that the standard deviation of $G1$ across the 28 administrations is 4.09 percent, although Table 2 shows that the variation is less in the post World War II period.

Table 2
Nine Measurs for 28 Administrations: 1916–2024
Percentage Points at Annual Rates

	Last Year	D	$R1$	$R2$	$R3$	$R4$	$G1$	$G2$	ZRS	ZP	$ZPOP$
1	1916.	1	1.10	6.08	3.53	-0.94	-0.09	-1.73	2.55	5.92	1.66
2	1920.	1	-9.10	-2.81	-6.83	-19.69	-1.48	-2.63	4.03	11.89	1.12
3	1924.	0	10.00	15.57	11.85	14.72	8.44	6.78	3.72	-4.63	1.70
4	1928.	0	21.85	26.31	22.99	23.28	3.49	2.21	3.32	-0.39	1.31
5	1932.	0	-31.45	-25.98	-28.12	-21.45	-8.94	-9.75	2.14	-7.26	0.84
6	1936.	1	22.73	27.02	26.80	25.14	10.85	10.21	0.22	2.56	0.64
7	1940.	1	-12.12	-7.09	-7.19	-7.37	5.35	4.54	0.11	0.34	0.79
8	1944.	1	5.68	11.14	10.83	5.00	13.35	12.16	0.31	5.02	1.18
9	1948.	1	3.38	7.94	7.34	-0.24	-4.05	-5.59	0.60	7.91	1.51
10	1952.	1	13.96	19.96	18.57	16.01	6.20	4.52	1.39	2.26	1.70
11	1956.	0	14.08	18.34	16.55	15.72	2.59	0.83	1.79	1.86	1.75
12	1960.	0	5.48	8.86	6.05	4.13	2.09	0.33	2.81	1.91	1.79
13	1964.	1	9.43	12.44	9.49	8.35	5.55	4.11	2.95	1.26	1.46
14	1968.	1	5.08	8.09	3.51	0.28	4.79	3.72	4.59	3.20	1.09
15	1972.	0	3.20	6.24	0.92	-3.58	3.42	2.30	5.33	4.72	1.12
16	1976.	0	-2.35	1.36	-4.99	-12.85	1.86	0.91	6.35	7.20	0.95
17	1980.	1	5.84	10.54	2.15	-7.70	3.33	2.25	8.39	7.71	1.09
18	1984.	0	5.21	9.78	-0.76	-5.74	2.97	2.06	10.55	5.01	0.92
19	1988.	0	12.68	16.01	9.58	6.21	3.78	2.87	6.42	2.82	0.90
20	1992.	0	11.26	14.34	8.29	4.20	1.98	0.73	6.05	3.13	1.22
21	1996.	1	13.27	15.72	11.31	8.53	3.40	2.20	4.41	2.03	1.20
22	2000.	1	14.45	15.81	10.77	8.45	3.94	2.82	5.04	1.66	1.13
23	2004.	0	-2.14	-0.55	-2.39	-4.63	2.77	1.84	1.84	2.16	0.94
24	2008.	0	-7.35	-5.41	-8.79	-11.28	0.71	-0.23	3.38	2.58	0.94
25	2012.	1	11.42	13.50	13.39	11.19	2.01	1.21	0.11	1.44	0.81
26	2016.	1	11.27	13.28	13.17	11.91	2.34	1.58	0.12	1.30	0.77
27	2020.	0	12.94	14.81	13.49	11.59	2.05	1.60	1.32	1.79	0.50
28	2024.	1	11.21	12.68	9.67	4.87	3.12	2.45	3.01	4.31	0.67

$R1$ = S&P 500 simple stock return, $R2$ = stock return reinvested dividends,
$R3$ = excess return = $R2$-ZRS, $R4$ = real excess return = $R3$-ZP,
$G1$ = real GDP growth, $G2$ = real per capita GDP growth,
ZRS = T-bill return, ZP = inflation, $ZPOP$ = population growth,
D = 1 if Democratic, 0 if Republican.

4 The Results

The main results in this paper are presented in Table 3. Presented for each measure are: D mean, R mean, the difference in means, the overall standard error, the D standard error, the R standard error, the t-statistic for the hypothesis that the means are equal using different estimated variances, the t-statistic using the overall estimated variance, and the t-statistic using the Newey West correction with lag of 2. Remember that the t-statistics are tests of the hypothesis that the coefficient of D_t is zero from a regression of the measure on a constant and D_t. The second t-statistic is for the standard OLS regression. The first uses White's correction for heteroskedasticity.

The results are easy to summarize. None of the six measures of stock returns and neither of the two growth rates is significantly different between D and R. The mean differences are higher for D, but the standard deviations are also high. The net effect is insignificant differences. This shows that the previous significant results in the literature are fragile.

The mean T-bill return, ZRS, is higher under R, but not quite significant. This is the reason for the larger positive difference (5.00) for D for excess returns, $R3$, than for $R1$ and $R2$, although it is still not significant.

Mean inflation is higher under D, but again not quite significant. This is the reason for the smaller positive difference (2.68) for D for real excess returns, $R4$, than for $R3$.

The non parametric test discussed in Section 2 was performed for $R1$. There are 28 administrations, 15 D and 13 R. The total number of different assignments is 37,442,160. The number of cases where the difference in means was greater in absolute value than 3.07, the mean difference in Table 3, was 19,517,395, a ratio of 0.516, which is the p-value. For the t-statistic t_2 in Table 3, 0.72, the p-value is 0.478 (26 degrees of freedom), so the non parametric test gives similar results.

Table 3
Mean Results for 28 Administrations: 1916–2024
Percentage Points at Annual Rates

	m_D	m_R	m_D -m_R	σ	σ_D	σ_R	t_1	t_2	t_3
$R1$	7.17	4.11	3.07	11.20	8.99	13.33	0.70	0.72	0.78
$R2$	10.95	7.67	3.29	10.86	8.27	13.25	0.77	0.80	0.87
$R3$	8.43	3.44	5.00	11.00	8.76	13.13	1.17	1.20	1.33
$R4$	4.51	1.83	2.68	11.72	10.61	12.90	0.60	0.60	0.58
$R5$	4.71	-0.77	5.48	11.36	9.06	13.55	1.24	1.27	1.38
$R6$	7.87	2.39	5.48	11.40	9.72	13.09	1.24	1.27	1.32
$G1$	3.91	2.09	1.81	4.09	4.34	3.79	1.18	1.17	1.05
$G2$	2.79	0.96	1.83	4.10	4.44	3.65	1.19	1.18	1.04
ZRS	2.52	4.23	-1.71	2.53	2.45	2.62	-1.78	-1.79	-1.87
ZP	3.92	1.61	2.31	3.54	3.23	3.87	1.70	1.73	1.38
$ZPOP$	1.12	1.14	-0.02	0.37	0.34	0.39	-0.16	-0.17	-0.22

See notes to Table 2.
$R5$ = VWX excess return, $R6$ = VWD excess return,
m_d = D mean, m_R = R mean, σ = oveall standard deviation,
σ_D = D standard deviation, σ_R = R standard deviation,
t_1 = t-statistic, different variances, t_2 = t-statistic, common variance,
t_3 = t-statistic, Newey West lag 2.

The non parametric test was also run for $G1$. The p-value was 0.297. For the t-statistic t_2 in Table 3, 1.17, the p-value is 0.253 (26 degrees of freedom), so again the non parametric test gives similar results.

5 Duplicating Previous Results

To check that the above methodology is consistent with the previous literature, the significant results of Pástor and Veronesi (2020) and Blinder and Watson (2016) have been duplicated. This shows that the lack of significance in this paper is due

to the use of an extended sample period rather than the use of a different procedure or different variables.

Pástor and Veronesi (2020) (PV)

PV's sample period was 1927–2015. They got significant results by running a monthly regression of monthly excess returns, $R5$, on a constant and a dummy variable that was 1 under Democrats and 0 under Republicans. Their monthly sample period was 1927.01–2015.12. Their dummy variable assumed that the new administration did not start until February of the first year. The standard errors were corrected for heteroskedasticity and autocorrelation. From this regression PV got an estimated difference of 10.90 with a t-statistic of 2.73. This same regression was run using the data in this study, and the results were similar. The estimated difference was 10.12 with a t-statistic of 2.31 using the Newey West correction with lag of 2. The failure to duplicate the PV result exactly is likely due to slightly different values of the T-bill rates. (The VWX data are the same.) The conclusion is, however, the same. For this period and this measure of returns, there is a significant difference between the D and R means.

Another way of looking at the sample-period question is to do the results as in Table 3 for PV's choice of administrations. Thier sample period indludes half of administration 1928 and three fourths of administration 2016. An approximation to this sample period is administrations 1932 through 2016. Results for these 22 administrations are presented in Table 4. This table has the same format as Table 3. For $R5$, which is the measure used in the monthly regression, the mean difference is 10.96, with t-statistics of 2.26, 2.36, and 2.22. Not all the measures in Table 4 are significant at conventional levels, but the t-statistics are all larger than they are in Table 3.

Table 4
Mean Results for 22 Administrations: 1932–2016
Percentage Points at Annual Rates

	m_D	m_R	m_D -m_R	σ	σ_D	σ_R	t_1	t_2	t_3
$R1$	8.70	0.86	7.84	10.92	8.45	13.33	1.61	1.68	1.57
$R2$	12.36	4.30	8.06	10.61	8.08	13.05	1.70	1.77	1.66
$R3$	10.01	-0.37	10.38	10.38	8.46	12.34	2.25	2.33	2.24
$R4$	6.95	-2.78	9.73	9.81	9.12	10.59	2.28	2.32	2.32
$R5$	6.64	-4.32	10.96	10.86	8.25	13.38	2.26	2.36	2.22
$R6$	10.45	-0.76	11.21	10.61	8.06	13.07	2.36	2.47	2.31
$G1$	4.76	1.32	3.43	4.08	4.36	3.71	1.99	1.96	1.79
$G2$	3.64	0.19	3.46	4.10	4.45	3.62	2.01	1.97	1.77
ZRS	2.35	4.66	-2.31	2.75	2.72	2.80	-1.96	-1.96	-2.00
ZP	3.06	2.41	0.65	3.16	2.51	3.80	0.46	0.48	0.54
$ZPOP$	1.11	1.14	-0.02	0.34	0.33	0.35	-0.16	-0.16	-0.18

See notes to Tables 2 and 3.
m_d = D mean, m_R = R mean, σ = oveall standard deviation,
σ_D = D standard deviation, σ_R = R standard deviation,
t_1 = t-statistic, different variances, t_2 = t-statistic, common variance,
t_3 = t-statistic, Newey West lag 2.

Regarding growth rates, the results for $G1$ are fairly close to those of PV. The D mean in Table 4 is 4.76 versus 4.86 for PV; the R mean is 1.32 versus 1.70; the difference is 3.43 versus 3.16,; and the t-statistic is 1.79 (Newey West) versus 2.40. The reason for these differences is because of the use of more recently revised data.

Blinder and Watson (2016) (BW)

BW use a considerably shorter sample period, 16 administrations, 1952–2012. Results are presented in Table 5 for the 1952–2012 sample period. The results in Table 5 are close to the BW results for $G1$. For $G1$ the D mean is 4.17 in Table 5 versus 4.09 for BW. The R mean is 2.46 versus 2.67. The difference is 1.71 versus 1.42. The t-statistics are 2.74, 2.48, and 3.21 versus 2.25 for BW. The failure to reproduce exactly is likely due to the use of later revised data here. The conclusion is, however, the same. The differences in mean growth rates are significant for this sample period.

The non parametric test was run for $G1$, and out of the 11,440 possible assignments 129 had the absolute value of the computed difference greater than 1.71, the mean difference in Table 5, for a p-value of 0.011. BW ran this test and got 146 cases, for a p-value of 0.013. So their result has been almost exactly reproduced here. For the t-statistic t_2 in Table 5, 2.89, the p-value is 0.012 (14 degrees of freedom), so again the non parametric test gives almost identical results.

6 Adding Nine More Administrations

As discussed in the Appendix, monthly data from Robert Shiller's website on SP are available back to to 1871.01. Quarterly data on real and nominal GDP and on population are available back to 1877.1. It is thus possible to compute observations on $R1$, the simple stock return, and $G1$ and $G2$ back to the administration ending in 1880—Hayes. Table 2 shows that the results across the various measures of stock returns are fairly close, so $R1$ is a good proxy for all the measures. For these administrations the new administration did not begin until March, which $G1$ and $G2$ capture by assuming that the first quarter of the next administration is the responsibility of the previous one.

Table 5
Mean Results for 16 Administrations: 1952–2012
Percentage Points at Annual Rates

	m_D	m_R	m_D -m_R	σ	σ_D	σ_R	t_1	t_2	t_3
$R1$	10.49	4.45	6.04	6.14	3.83	7.41	2.11	1.95	1.73
$R2$	13.72	7.66	6.06	6.57	3.89	8.02	1.99	1.83	1.67
$R3$	9.88	2.72	7.17	7.10	5.64	8.01	2.10	2.00	2.14
$R4$	7.09	-0.77	7.86	8.33	7.36	8.99	1.92	1.87	2.22
$R5$	6.99	-0.70	7.69	6.20	4.15	7.37	2.64	2.46	2.54
$R6$	10.32	2.65	7.66	6.56	4.35	7.82	2.49	2.32	2.38
$G1$	4.17	2.46	1.71	1.17	1.44	0.92	2.74	2.89	3.21
$G2$	2.97	1.29	1.68	1.10	1.19	1.02	2.98	3.05	2.84
ZRS	3.84	4.95	-1.11	2.77	2.70	2.82	-0.80	-0.79	-1.05
ZP	2.79	3.49	-0.69	2.01	2.26	1.80	-0.66	-0.68	-1.22
$ZPOP$	1.21	1.17	0.04	0.33	0.29	0.36	0.26	0.25	0.14

See notes to Tables 2 and 3.
m_d = D mean, m_R = R mean, σ = oveall standard deviation,
σ_D = D standard deviation, σ_R = R standard deviation,
t_1 = t-statistic, different variances, t_2 = t-statistic, common variance,
t_3 = t-statistic, Newey West lag 2.

Table 6 contains observations for the nine administrations. It has the same format as Table 2 but with fewer variables. Only two of the nine administrations were Democratic, 1888 and 1896, one with fairly good returns and growth and one not. A number of the Republican administrations have good returns and growth.

The mean results for the 37 administrations are presented in Table 7. This table has the same format as Tables 3, 4, and 5. The t-statistics for $R1$ are about the same as they are in Table 3. For $R1$ the D mean is 6.19 versus 7.17 in Table 3; the R mean is 4.00 versus 4.11; and the difference is 2.19 versus 3.07. The t-statistics are 0.65, 0.64, and 0.68 versus 0.70, 0.72, and 0.75. The additional observations have not changed the story.

Table 6
Five Measurs for 9 Administrations 1880-1912
Percentage Points at Annual Rates

	Last Year	D	$R1$	$G1$	$G2$	ZP	$ZPOP$
1	1880.	0	12.23	8.09	5.89	-1.72	2.03
2	1884.	0	-7.42	1.88	-0.52	-1.79	2.42
3	1888.	1	4.23	2.21	0.03	0.11	2.19
4	1892.	0	1.74	5.59	3.56	-1.55	2.04
5	1896.	1	-6.67	1.30	-0.59	-2.36	1.90
6	1900.	0	12.18	6.27	4.48	2.20	1.77
7	1904.	0	4.58	2.95	1.01	1.89	1.93
8	1908.	0	2.26	3.86	1.95	1.90	1.91
9	1912.	0	0.95	3.96	2.16	2.06	1.79

$R1$ = S&P 500 simple stock return,
$G1$ = real GDP growth,
$G2$ = real per capita GDP growth,
ZP = inflation, $ZPOP$ = population growth,
D = 1 if Democratic, 0 if Republican.

Table 7
Mean Results for 37 Administrations: 1880–2024
Percentage Points at Annual Rates

	m_D	m_R	m_D -m_R	σ	σ_D	σ_R	t_1	t_2	t_3
$R1$	6.19	4.00	2.19	10.32	9.07	11.27	0.65	0.64	0.68
$G1$	3.65	2.99	0.66	3.78	4.12	3.47	0.52	0.53	0.46
$G2$	2.43	1.55	0.88	3.76	4.28	3.25	0.69	0.71	0.61
ZP	3.33	1.19	2.13	3.39	3.48	3.32	1.90	1.90	1.65
$ZPOP$	1.23	1.44	-0.21	0.50	0.45	0.53	-1.30	-1.28	-1.59

See notes to Table 3.

The t-statistics for $G1$ and $G2$ in Table 7 are lower than they are in Table 3. For $G2$ the D mean is 2.43 versus 2.79 in Table 3; the R mean is 1.55 versus 0.96; and the difference is 0.88 versus 1.81. The t-statistics are 0.69, 0.71, and 0.61 versus 1.14, 1.18, and 1.04. The differences in growth means are clearly not significant.

The non parametric test was run for $G2$ in Table 7. There are 37 administrations, 20 R and 17 D. The number of possible different assignments is 15,905,368,710. Of these possibilities, 7,815,478,681 had the absolute value of the computed difference greater than 0.88, the mean difference in Table 7, for a p-value of 0.481. For the t-statistic t_2 in Table 7, 0.71, the p-value is 0.482 (35 degrees of freedom), so the non parametric test gives almost identical results.

7 Means versus Variances

In a series of papers in the mid 1980s—see, for example, Romer (1986)—Christina Romer argued that data before the Great Depression have measurement errors such that they show greater economic variation than actually existed. By adding earlier observations in this study it could be that the lack of significance is due to increased variation with no decrease in the mean differences, where the increased variation is due to measurement error. This is, however, not the case. The following chart gives results for $R1$ and $G2$.

			m_D -m_R	σ
$R1$				
Table 7	1880-2024	37 obs.	2.19	10.32
Table 3	1916-2024	28 obs.	3.07	11.20
Table 4	1932-2016	22 obs.	7.84	10.92
Table 5	1952-2012	16 obs.	6.04	6.14
$G2$				
Table 7	1880-2024	37 obs.	0.88	3.76
Table 3	1916-2024	28 obs.	1.83	4.10
Table 4	1932-2016	22 obs.	3.43	4.08
Table 5	1952-2012	16 obs.	1.68	1.17

For $R1$ the mean differences get smaller as earlier observations are added, and the overall standard deviation, σ, does not change much except for the short 1952-2012 period. This is also true for $G2$. except for the period 1952–2012, where the mean difference is smaller. The standard deviation is thus not sensitive to adding the earlier observations ignoring the short 16-observation case.

8 Conclusion

The results in this paper show that the view that stock returns and output growth are higher under Democrats is not robust to adding more observations. Using data on the past 27 administrations does not result in significant differences between Democrats and Republicans. This is also the case when the nine administrations between Hayes and Taft are added. In many cases the differences in means between the two parties look large, but there is considerable variation across administrations and the differences are not statistically significant. There is thus no puzzle, contrary to the conclusions of Leonhardt (2021) and Cahan and Potrafke (2021). As noted in the Introduction, the fact that there is no puzzle is not particularly surprising. In the case of significant stock return differences across administrations, theory suggests otherwise. In the case of significant growth differences, there is no compelling theory either way.

Appendix: Data Collection

Quarterly data on nominal GDP, real GDP, and population were collected for 1877–2024. For nominal GDP, annual data for 1929–1946 and quarterly data for 1947.1–2024.4 were obtained from the Bureau of Economic Analysis (BEA) website. The data are as of January 30, 2025. Quarterly data for 1877.1–1946.4 are available from Balke and Gordon (1986), pp. 789–795. The Balke and Gordon values for 1877.1–1928.4 were used exactly, but the values for 1929.1–1946.4 were adjusted to take account of the BEA annual data. For 1929.1–1946.4 each quarterly value for a given year was multiplied by a splicing factor for that year. The splicing factor is the ratio of the BEA value for that year to the respective yearly value in Balke and Gordon (1976), pp. 782–783.

The data on real GDP were obtained in a similar way. Annual data for 1929–1946 and quarterly data for 1947.1–2024.4 were obtained from the BEA website. The units are 2017 dollars. Quarterly data for 1877.1–1946.4 are available from Balke and Gordon (1986), pp. 789–795. The Balke and Gordon values were spliced to the BEA values. All the Balke and Gordon quarterly values for 1877.1–1929.4 were multiplied by the same number. This number is the ratio of the BEA value for 1929 to the 1929 value in Balke and Gordon (1976), p. 782. For 1930.1–1946.4 each Balke and Gordon quarterly value for a given year was multiplied by a splicing factor for that year. The splicing factor is the ratio of the BEA value for that year to the respective yearly value in Balke and Gordon (1976), pp. 782–783.

The data on population were obtained as follows. For 1877–1928 annual data were obtained from U.S. Department of Commerce (1973), pp. 200–201, A114 series. Each of these observations was multiplied by 1.000887, a splicing factor. The splicing factor is the ratio of the A114 value for 1929 in U.S. Department of Commerce (1973) to the value for 1929 in Table 8.2 in U.S. Department of Commerce (1992). For 1929–1945 annual data were obtained from U.S. Department of Commerce (1992), Table 8.2. Quarterly observations for 1877.1–1945.4

were obtained by interpolating the annual observations using the method presented in Fair (1994), Table B.6. For 1946.1–1946.4 quarterly data were obtained from the BEA website on October 27, 2006. For 1947.1–2024.4 quarterly data were obtained from the BEA website as of January 30, 2025.

Regarding the data used, the GDP deflator is nominal GDP divided by real GDP and per capita real GDP is real GDP divided by population.

Daily data on SP, the S&P 500 stock price index, were obtained from the Yahoo Finance website for 1928–2024. From these daily data a monthly series was constructed using the price on the last trading day of the month. End of month data were collected from CRSP for the 1926.01–1927.12 period. Monthly data for 1871.01–1925.12 were collected from Robert Shiller's website. These data are the average price for the month, not the price at the end of the month.

Quarterly data on S&P 500 dividends were obtained from Standard and Poors for the 1935.1–2024.4 period. For the period 1912.1-1934.4 data were taken from Shiller's website. The data on this site are monthly, and quarterly data were constructed by summing the three months.

Monthly data on VWX and VWD were obtained from CRSP for the 1926.01–2024.12 period. Both are monthly percent changes. For the 1912.01–1925.12 period, both VWX and VWD were taken to be the monthly percent change in SP. The correlation between VWX and the monthly percent change in SP for the 1926.01–2024.12 period is 0.980. For VWD it is 0.979. This procedure is thus likely to be a good approximation.

Monthly data on the three-month Treasury bill rate were obtained from the Board of Governors of the Federal Reserve System for the 1934.01–2024.12 period. Monthly data for the 1920.01–1933.13 period were taken from the FRED website, the three-month Treasury bill rate from the NBER Macroeconomic Database. Monthly data for the 1912.01–1918.12 period were also taken from the FRED website, the commercial paper rate for New York from the NBER Macroeconomic Database. RS for this paper was taken to be the commercial paper rate minus 1.75,

which splices it to the T-bill series.

References

[1] Balke, Nathan S., and Robert J. Gordon, 1986, "Appendix B Historical Data," in Robert J. Gordon (ed.), *The American Business Cycle: Continuity and Change*, Chicago: University of Chicago Press.

[2] Blinder, Alan S., and Mark W. Watson, 2016, "Presidents and the US Economy: An Econometric Exploration," *American Economic Review*, 106, 1015–1045.

[3] Cahan, Dodge, and Niklas Potrafke, 2021, "The Democrat-Republican Presidential Growth Gap and the Partisan Balance of the State Governments," *Public Choice*.

[4] Fair, Ray C., 1994, *Testing Macroeconometric Models*, Cambridge, MA: Harvard University Press.

[5] Leonhardt, David, 2021, "The Economy Does Much Better Under Democrats. Why?" *The New York Times*, February 7.

[6] Newey, Whitney K., and Kenneth D. West, 1987, "A Simple Positive Semi-Definite Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

[7] Pástor, Ľuboš, and Pietro Veronesi, 2020, "Political Cycles and Stock Returns," *Journal of Political Economy*, 128, 4011–4045.

[8] Romer, Christina D., 1986, "Is the Stabilization of the Postwar Economy a Figment of the Data," *The American Economic Review*, 76, 314–334.

[9] Santa-Clara, Pedro, and Rossen Valkanov, 2003, "The Presidential Puzzle: Political Cycles and the Stock Market," *The Journal of Finance*, 63, 1841–1872.

[10] U.S. Department of Commerce, 1973, *Long Term Economic Growth, 1860-1970*, Washington, DC: U.S. Government Printing Office.

[11] U.S. Department of Commerce, 1975, *Historical Statistics of the United States, Colonial Times to 1970*, Washington, DC: U.S. Government Printing Office.

[12] U.S. Department of Commerce, 1992, *National Income and Product Accounts of the United States, Volume 2, 1959–1988*, Washington, DC: U.S. Government Printing Office.

[13] White, Halbert, 1980, "A Heteroskedasticity Consistent Covariance Matrix Estimator and a Direct Test of Heteroskedasticity," *Econometrica*, 48, 817–838.