Are Stock Returns and Output Growth Higher Under Democrats?

Ray C. Fair*

September 2021

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 27 administrations, from Wilson-1 through Trump. While the mean stock return under the Democrats is generally higher, none of the differences in means are 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 ten administrations between Grant-2 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. .

1 Introduction

Recent literature suggests that both stock returns and economic growth are significantly higher under Democratic presidential administrations than under Republican

*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.

presidential administrations in the United States. Regarding stock returns, Santa-Clara and Valkanov (2003) (SV) examined the period 1927–1998 and found significant differences. Pástor and Veronesi (2020) (PV) extended the sample period to 1927–2015 and also found significant differences. SV found this a puzzle, since persistent differences seem unlikely in efficient markets. PV 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) (BW) 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. PV got similar results for the 1930–2015 period. The BW 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 main results in this paper the sample is extended to 27 administrations, from Wilson-1 through Trump. This period includes 14 Democratic administrations and 13 Republican. While the mean stock return under the Democrats is generally higher than the mean stock return under the Republicans, depending on the particular stock return used, none of the differences in means are 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 than the mean output growth under the Republicans, the difference is not significant. Again, there is considerable variation in output growth across administrations. Section 5 contains results going back ten more administrations to Grant-2, with a similar conclusion.

The reason for the different conclusion in this paper is because of the use of the extended sample period. The results below confirm the significant results of PV and BW using their sample periods. It turns out that this significance is fragile to the choice of sample periods. Adding more observations tends to lower the mean differences rather than increasing the variances, as discussed at the end of Section 5.

2 The Test

Consider stock returns first. Let R_t denote a measure of stock returns during administration t, where t runs from 1 to 27. 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 27. Then treat G_t as R_t above and run the tests. Five measures are considered: real output growth, per capita real output growth, CPI inflation, GDP deflator inflation, and the three-month Treasury bill rate. The main concern is with output growth.

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 is of the hypothesis 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 reported, which uses the Newey-West

(1987) correction for heteroskedasticity and autocorrelation with a lag of 2. Both PV and BW 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 BW. As discussed below, 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 also be seen that this test gives very similar results to the others.

3 The Data

The data collection, which is somewhat tedious, is explained in the appendix. For the main results data were collected for the 1912–2020 period. Data were collected on the S&P 500 stock price index (SP), S&P 500 dividends (DIV), the three-month Treasury bill rate (RS), the consumer price index (CPI), the CRSP value-weighted return excluding dividends (VWX), the CRSP value-weighted return including distributions (VWD), real GDP (Y), the GDP deflator (P), and population (POP). The following is a discussion of the measures using these data.

$R1$: 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.[1]

[1] All logs in this study have been divided by 4 to put them at an annual rate and multiplied by

$R2$: Quarterly data were collected on DIV. $R2$ assumes that the dividends are invested in SP at the end of each quarter and cumulated. The value of SP used for each investment is SP 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.

$R3$: Monthly data on RS were collected, average for the month. Quarterly data were constructed by summing the three months. $R3$ assumes that the dividends are invested at the end of each quarter in three month Treasury bills and rolled over each quarter after that. The interest rate used for, say, DIV paid in the first quarter of the administration is the quarterly value of RS in the second quarter, and so on. No interest is accumulated on DIV paid in the last quarter of the administration. $R3$ is then the log of the value of SP at the end of December of the fourth year of the administration plus the value of all the dividend and interest income received during the administration minus the log of SP at the end of December of the fourth year of the previous administration.

ZRS: This measure is used in the construction of other measures. It uses quarterly data on RS. 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. When ZRS is subtracted from, say, $R1$, this is a measure of excess returns—returns over and above what could be achieved by rolling over three-month Treasury bills.

ZP and ZPP: These measures are also used in the construction of other measures. The CPI data are monthly. ZP is the log of CPI in December of the fourth year of the administration minus the log of CPI in December of the fourth year of the previous administration. The P data are quarterly. ZPP 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. When ZP

100 to put them in percentage points.

or ZPP is subtracted from, say, $R1$, this is a measure of real returns.

$R4$ and $R5$: 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, $R4$ is the log of the value at the end. The same procedure was followed for VWD, where $R5$ is the log of the value at the end. These two measures can then have ZRS subtracted from them to make them excess returns, and they can have ZP or ZPP subtracted from them to make them real returns.

$G1$, $G2$, $GP1$, $GP2$, and $ZPOP$: The data on Y are quarterly. $G1$ is the log of Y in the fourth quarter of the fourth year of the administration minus the log of Y in the fourth quarter of the fourth year of the previous administration. $GP1$ is the same for per capita GDP, which is Y/POP. $ZPOP$ 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. The main specification of BW gives the first quarter of the new administration to the previous administration. $G2$ is the same as $G1$ but with this modification. Similarly, $GP2$ is the same as $GP1$ with this modification. This modification was not made for the Trump administration because data for 2021.1 were not available at the time of this analysis.

4 The Results

To avoid clutter, results for only a subset of the measures are presented in this section. More detailed results are in the appendix. Two measures of inflation are available, one using the CPI and one using the GDP deflator. These measures are highly correlated, and it makes little difference which is used in computing real returns. In this section the CPI measure is used.

Table 1
Eight Measures of Stock Returns for 27 Administrations
Percentage Points at Annual Rates

	Last Year	D	$R1$	$R2$	$R3$	$R1$ -ZRS	$R1$ -ZP	$R4$	$R5$	$R4$ -ZRS
1	1916.	1	1.10	6.08	5.67	-1.46	-3.38	1.10	1.10	-1.46
2	1920.	1	-9.10	-2.81	-1.19	-13.13	-21.96	-9.10	-9.10	-13.13
3	1924.	0	10.00	15.57	14.88	6.29	12.87	10.00	10.00	6.29
4	1928.	0	21.85	26.32	24.83	18.54	22.14	19.32	22.81	16.01
5	1932.	0	-31.45	-25.98	-21.44	-33.59	-24.79	-34.72	-29.36	-36.86
6	1936.	1	22.73	27.02	25.61	22.51	21.07	24.30	28.49	24.08
7	1940.	1	-12.12	-7.09	-6.62	-12.23	-12.30	-10.96	-5.86	-11.06
8	1944.	1	5.68	11.14	10.10	5.37	-0.14	6.26	12.23	5.95
9	1948.	1	3.38	7.94	8.13	2.78	-4.20	2.93	7.92	2.33
10	1952.	1	13.96	19.96	18.74	12.58	11.40	12.90	19.11	11.51
11	1956.	0	14.08	18.34	17.40	12.29	13.25	13.43	18.02	11.64
12	1960.	0	5.48	8.86	8.61	2.68	3.56	6.11	9.80	3.30
13	1964.	1	9.43	12.44	12.09	6.49	8.29	8.64	11.76	5.69
14	1968.	1	5.08	8.09	7.98	0.50	1.86	7.62	10.68	3.03
15	1972.	0	3.20	6.24	5.97	-2.13	-1.30	1.85	4.90	-3.48
16	1976.	0	-2.35	1.36	1.22	-8.70	-10.21	-3.39	0.48	-9.74
17	1980.	1	5.84	10.54	10.20	-2.55	-4.00	9.12	13.93	0.72
18	1984.	0	5.21	9.78	9.81	-5.33	0.24	4.97	9.51	-5.57
19	1988.	0	12.68	16.01	15.98	6.26	9.31	11.37	14.94	4.94
20	1992.	0	11.26	14.34	14.12	5.21	7.17	10.89	14.11	4.84
21	1996.	1	13.27	15.72	15.21	8.86	10.49	12.61	14.98	8.20
22	2000.	1	14.45	15.81	15.77	9.41	12.13	12.94	14.32	7.90
23	2004.	0	-2.14	-0.55	-0.73	-3.98	-4.38	-0.29	1.38	-2.12
24	2008.	0	-7.35	-5.41	-4.49	-10.73	-9.84	-6.75	-4.77	-10.13
25	2012.	1	11.42	13.50	13.16	11.31	9.22	11.99	14.28	11.88
26	2016.	1	11.27	13.28	13.04	11.16	10.02	9.55	11.71	9.44
27	2020.	0	12.94	14.81	14.39	11.61	11.04	11.71	13.27	10.39

$R1$ = S&P 500 stock return, $R2$ = stock return reinvested dividends,
$R3$ = stock return dividends invested in T-bills, $R1$-ZRS = excess return,
$R1$-ZP = real return, $R4$ = VWX return, $R5$ = VWD return,
$R4$-ZRS = excess return, ZRS = T-bill return, ZP = inflation,
D = 1 if Democratic, 0 if Republican.

Table 2
Mean Results for Eight Measures of Stock Returns
27 Administrations: 1916–2020
Percentage Points at Annual Rates

	m_D	m_R	m_D -m_R	σ	σ_D	σ_R	t_1	t_2	t_3
$R1$	6.89	4.11	2.78	11.39	9.26	13.33	0.62	0.63	0.68
$R2$	10.83	7.67	3.16	11.07	8.57	13.25	0.73	0.74	0.80
$R3$	10.56	7.74	2.83	10.08	7.98	11.95	0.72	0.73	0.78
$R1$-ZRS	4.40	-0.12	4.52	11.55	9.70	13.26	1.01	1.02	1.12
$R1$-ZP	2.75	2.24	0.51	11.89	11.30	12.50	0.11	0.11	0.11
$R4$	7.14	3.42	3.71	11.47	9.09	13.58	0.83	0.84	0.89
$R5$	10.40	6.55	3.85	11.48	9.65	13.18	0.86	0.87	0.90
$R4$-ZRS	4.65	-0.81	5.46	11.56	9.40	13.52	1.21	1.23	1.31

$R1$ = S&P 500 stock return, $R2$ = stock return reinvested dividends,
$R3$ = stock return dividends invested in T-bills, $R1$-ZRS = excess return,
$R1$-ZP = real return, $R4$ = VWX return, $R5$ = VWD return,
$R4$-ZRS = excess return, ZRS = T-bill return, ZP = inflation,
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.

Stock Returns

Table 1 presents 8 measures of stock returns for each of the 27 administrations. These measures are for the four years of the administration, but they are at annual rates. An administration will be 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). From 1988 through 2020 the returns have been large except for 2004 and 2008 (G.W. Bush). Five of these nine administrations were Republican, including 2004 and 2008. The volatility across administrations is large. As reported below the standard deviation of $R1$ across the 27 administrations is 11.39 percent.

A similar story holds for the other measures.

The main results for stock returns are in Table 2. Presented for each of the 8 measures 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.

For $R1$ the D mean is 6.89 and the R mean is 4.11, for a difference of 2.78. This may seem large, but not from a statistical standpoint, where the t-statistics are only 0.62, 0.63, and 0.68. For the excess return, $R1 - ZRS$, the D mean is 4.40, the R mean is -0.12, with a difference of 4.52, with t-statistics of 1.01, 1.02, and 1.12. As shown below, the T-bill return is lower under D, with a D mean of 2.49 and an R mean of 4.23. Thus, less is subtracted for D, which means higher excess returns, although not significant. Also, as shown below, the inflation rate is higher under D, with a D mean of 4.14 and an R mean of 1.87. Thus, more is subtracted for D, which means lower real returns. For $R1 - ZP$ the difference for D is only 0.51, with t-statistics of 0.11, 0.11, and 0.11. The CRSP stock returns, which PV use, have somewhat larger differences favoring D, but the differences are still not significant.

The non parametric test discussed in Section 2 was performed for $R1$. There are 27 administrations, 14 D and 13 R. The total number of different assignments is 20,058,300. The number of cases where the difference in means was greater in absolute value than 2.78, the mean differnec in Table 2, was 11,084,369, a ratio of 0.553, which is the p-value. For the t-statistic t_2 in Table 2, 0.63, the p-value is 0.534 (25 degrees of freedom), so the non parametric test gives almost identical results.

How sensitive are these results to the choice of sample periods? PV's sample period was 1927–2015. They got significant results by running a monthly regression of monthly excess returns 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 return was the log return on the CRSP value-weighted stock portfolio (VWX) in excess of the log return on the 3-month T-bill. 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 conclusion is, however, the same. For this period and this measure of returns, there is a significant difference between the D and R means.

Regarding the administrations, PV's sample period includes 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 3. For the measure $R4$-ZRS, 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 3 are significant at conventional levels, but the t-statistics are all larger than they are in Table 2. For this sample period one would say that the results are mixed.

Another way of looking at the sample period question is to examine the five administrations omitted from Table 3. Table 4 contains rusults for these five: the first four and the last, the Trump administration. As expected, the results are quite positive for the Republicans. For $R1$, the simple stock return, the D mean is -4.00, the R mean is 14.93, with a difference of -18.93. The t-statistics are -3.04 and

Table 3
Mean Results for Eight Measures of Stock Returns
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
$R3$	11.95	4.65	7.31	9.67	7.64	11.69	1.70	1.76	1.64
$R1$-ZRS	6.35	-3.80	10.15	10.68	8.72	12.68	2.14	2.22	2.14
$R1$-ZP	5.32	-1.70	7.02	10.17	9.21	11.23	1.58	1.61	1.50
$R4$	8.99	0.35	8.65	11.15	8.16	13.96	1.73	1.81	1.69
$R5$	12.80	3.90	8.90	10.88	7.84	13.71	1.82	1.91	1.78
$R4$-ZRS	6.64	-4.32	10.96	10.86	8.25	13.38	2.26	2.36	2.22

See notes to Table 2.

Table 4
Mean Results for Eight Measures of Stock Returns
5 Administrations: 1916–1928 and 2020
Percentage Points at Annual Rates

	m_D	m_R	m_D -m_R	σ	σ_D	σ_R	t_1	t_2
$R1$	-4.00	14.93	-18.93	6.54	7.21	6.17	-3.04	-3.17
$R2$	1.64	18.90	-17.26	6.39	6.28	6.44	-2.98	-2.96
$R3$	2.24	18.03	-15.80	5.57	4.85	5.89	-3.27	-3.11
$R1$-ZRS	-7.29	12.15	-19.44	6.92	8.25	6.14	-2.85	-3.08
$R1$-ZP	-12.67	15.35	-28.02	9.01	13.14	5.95	-2.83	-3.41
$R4$	-4.00	13.68	-17.68	5.81	7.21	4.96	-3.02	-3.33
$R5$	-4.00	15.36	-19.36	6.85	7.21	6.66	-3.03	-3.10
$R4$-ZRS	-7.29	10.89	-18.19	6.21	8.25	4.88	-2.81	-3.21

See notes to Table 2.

-3.17.[2] The other measures have similar results, including $R4$-ZRS. This is, of course, the reason the observations added in this paper decrease the overall mean difference for the Democrats. The previous results are fragile to adding the four observations.

Economic Growth

Turning now to the economic measures, Table 5 presents 8 measures of the economy for each of the 27 administrations. These measures are for the four years of the administration, but as in Table 1, they are at annual rates. The growth rate, $G2$, which BW prefer, varies from -8.94 for 1932 (Hoover) to 13.35 for 1944 (Roosevelt). All four measures of growth are quite similar.

The main results for the economic measures are in Table 6. This table has the same format at Table 2. The overall standard deviation for $G2$ is 4.18. For D it is 4.50, and for R it is 3.80. The D mean is 3.96 and the R mean is 2.03, for a difference of 1.93. Again, this may seem large, but it is not statistically significant, with t-statistics of 1.21, 1.20, and 1.09. The other three growth measures give very similar results. This table also shows that the T-bill return is lower under D and inflation is higher, but none of the differences are significant. The t-statistics are, however, higher in absolute value than they are for the four growth measures. The mean growth rates of population between D and R are almost identical, which means that the per capita results are almost identical to the non per capita ones.

To check on the significance results, the non parametric test was run for $G2$. The p-value was 0.252. For the t-statistic t_2 in Table 6, 1.20, the p-value is 0.241 (25 degrees of freedom), so again the non parametric test gives almost identical results.

[2] In this case the number of observations is too small to get sensible Newey West results.

Table 5
Eight Measures of the Economy for 27 Administrations
Percentage Points at Annual Rates

	Last Year	D	$G1$	$GP1$	$G2$	$GP2$	ZRS	ZP	ZPP	$ZPOP$
1	1916.	1	0.41	-1.25	-0.09	-1.73	2.55	4.47	5.92	1.66
2	1920.	1	-0.95	-2.08	-1.48	-2.63	4.03	12.86	11.89	1.12
3	1924.	0	6.71	5.00	8.44	6.78	3.72	-2.86	-4.63	1.70
4	1928.	0	3.56	2.26	3.49	2.21	3.32	-0.29	-0.39	1.31
5	1932.	0	-7.17	-8.01	-8.94	-9.75	2.14	-6.66	-7.26	0.84
6	1936.	1	9.63	8.99	10.84	10.21	0.22	1.66	2.56	0.64
7	1940.	1	4.50	3.71	5.35	4.54	0.11	0.18	0.34	0.79
8	1944.	1	13.50	12.32	13.35	12.16	0.31	5.83	5.02	1.18
9	1948.	1	-3.05	-4.55	-4.05	-5.59	0.60	7.58	7.91	1.51
10	1952.	1	5.39	3.69	6.20	4.51	1.39	2.56	2.26	1.70
11	1956.	0	2.89	1.14	2.59	0.82	1.79	0.83	1.86	1.75
12	1960.	0	2.09	0.29	2.10	0.33	2.81	1.92	1.90	1.79
13	1964.	1	5.12	3.66	5.55	4.11	2.95	1.15	1.26	1.46
14	1968.	1	5.00	3.91	4.79	3.72	4.59	3.23	3.20	1.09
15	1972.	0	3.20	2.08	3.42	2.31	5.33	4.50	4.72	1.12
16	1976.	0	2.18	1.23	1.86	0.91	6.35	7.86	7.20	0.95
17	1980.	1	3.14	2.05	3.33	2.25	8.39	9.85	7.71	1.09
18	1984.	0	3.22	2.30	2.97	2.06	10.55	4.97	5.01	0.92
19	1988.	0	3.77	2.86	3.78	2.87	6.42	3.37	2.82	0.90
20	1992.	0	2.19	0.97	1.98	0.73	6.05	4.09	3.13	1.22
21	1996.	1	3.28	2.08	3.40	2.20	4.41	2.78	2.03	1.20
22	2000.	1	4.19	3.06	3.96	2.84	5.04	2.32	1.66	1.13
23	2004.	0	2.42	1.49	2.77	1.84	1.84	2.24	2.13	0.94
24	2008.	0	1.20	0.26	0.64	-0.30	3.38	2.49	2.57	0.94
25	2012.	1	1.44	0.68	1.95	1.20	0.11	2.20	1.47	0.76
26	2016.	1	2.40	1.70	2.32	1.62	0.12	1.26	1.38	0.70
27	2020.	0	1.23	0.72	1.23	0.72	1.32	1.90	1.79	0.52

$G1$ = real GDP growth, $GP1$ = per capita growth,
$G2$ = real GDP growth, first quarter adjustment, $GP2$ = per capita growth,
ZRS = T-bill return, ZP = CPI inflation,
ZPP = GDP deflator inflation, $ZPOP$ = population growth
D = 1 if Democratic, 0 if Republican.

Table 6
Mean Results for Eight Measures of the Economy
27 Administrations: 1916–2020
Percentage Points at Annual Rates

	m_D	m_R	m_D -m_R	σ	σ_D	σ_R	t_1	t_2	t_3
$G1$	3.86	2.11	1.74	3.69	4.15	3.12	1.24	1.23	1.14
$GP1$	2.71	0.97	1.74	3.69	4.24	2.98	1.24	1.23	1.11
$G2$	3.96	2.03	1.93	4.18	4.50	3.80	1.21	1.20	1.09
$GP2$	2.82	0.89	1.93	4.18	4.61	3.65	1.21	1.20	1.07
ZRS	2.49	4.23	-1.75	2.58	2.54	2.62	-1.76	-1.76	-1.86
ZP	4.14	1.87	2.26	3.66	3.67	3.66	1.60	1.60	1.43
ZPP	3.90	1.60	2.30	3.60	3.34	3.87	1.65	1.66	1.33
$ZPOP$	1.14	1.15	0.00	0.37	0.34	0.39	-0.01	-0.01	-0.01

$G1$ = real GDP growth, $GP1$ = per capita growth,
$G2$ = real GDP growth, first quarter adjustment, $GP2$ = per capita growth,
ZRS = T-bill return, ZP = CPI inflation,
ZPP = GDP deflator inflation, $ZPOP$ = population growth
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.

As with stock returns, it is interesting to examine the growth differences for different sample periods. BW use a considerably shorter sample period, 16 administrations, 1952–2012. Results are presented in Table 7 for the 1952–2012 sample period. The results in Table 7 are close to the BW results for $G1$ and $G2$. (BW present results for $G1$ in an online appendix.) For $G1$ the D mean is 3.94 in Table 7 versus 4.09 for BW. The R mean is 2.57 versus 2.67. The difference is 1.37 versus 1.42. The t-statistics are 2.30, 2.48, and 2.61 versus 2.25 for BW. For $G2$ the D mean is 4.17 versus 4.33; the R mean is 2.46 versus 2.54; the difference is 1.71 versus 1.79; and t-statistics of 2.71, 2.86, and 2.45 versus 2.67 for BW. The failure to reproduce exactly is likely due to the use of later revised data here. The

Table 7
Mean Results for Eight Measures of the Economy
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
$G1$	3.94	2.57	1.37	1.09	1.41	0.77	2.30	2.48	2.61
$GP1$	2.73	1.40	1.33	1.02	1.18	0.88	2.49	2.58	2.33
$G2$	4.17	2.46	1.71	1.19	1.45	0.94	2.71	2.86	2.45
$GP2$	2.98	1.29	1.69	1.10	1.19	1.03	2.98	3.04	3.20
ZRS	3.84	4.95	-1.11	2.77	2.70	2.82	-0.80	-0.79	-1.05
ZP	3.44	3.59	-0.14	2.47	2.90	2.08	-0.11	-0.12	-0.91
ZPP	2.80	3.48	-0.68	2.01	2.26	1.80	-0.66	-0.68	-1.21
$ZPOP$	1.20	1.17	0.03	0.33	0.30	0.36	0.21	0.21	0.28

See notes to Table 6.

Table 8
Mean Results for Eight Measures of the Economy
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
$G1$	4.55	1.60	2.95	3.70	4.08	3.17	1.91	1.86	1.74
$GP1$	3.44	0.46	2.98	3.71	4.14	3.09	1.93	1.88	1.73
$G2$	4.75	1.32	3.43	4.08	4.37	3.71	1.99	1.96	1.79
$GP2$	3.65	0.18	3.47	4.10	4.45	3.62	2.01	1.98	1.77
ZRS	2.35	4.66	-2.31	2.75	2.72	2.80	-1.96	-1.96	-2.00
ZP	3.38	2.56	0.82	3.32	2.89	3.79	0.56	0.58	0.73
ZPP	3.07	2.41	0.66	3.15	2.50	3.80	0.47	0.49	0.56
$ZPOP$	1.10	1.14	-0.03	0.34	0.34	0.35	-0.22	-0.22	-0.25

See notes to Table 6.

conclusion is, however, the same. The differences in mean growth rates are significant for this sample period.

The non parametric test was run for $G2$, and out of the 11,440 possible assignments 141 had the absolute value of the computed difference greater than 1.71, the mean difference in Table 7, for a p-value of 0.012. 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 7, 2.86, the p-value is 0.013 (14 degrees of freedom), so again the non parametric test gives almost identical results.

As noted in the Introduction, PV also examine the growth rate. Their sample period is longer than BW's: 1930–2015. For complete administrations, the closest period is 1932–2016, and results for this period are presented in Table 8. The results for $G1$ are fairly close to those of PV. A D mean of 4.55 in Table 8 versus 4.86 for PV; a R mean of 1.60 versus 1.70; a difference of 2.95 versus 3.16,; and a t-statistic of 1.74 (Newey West) versus 2.40.

To examine the exact 1930–2015 period, a quarterly regression was run for the 1930.1–2015.4 period, with the quarterly log growth rate regressed on a constant and a quarterly dummy variable that is 1 for Democrats and 0 for Republicans. The coefficient estimate for the dummy variable was 3.01, which compares to 3.10 for PV. PV report a t-statistic of 2.40. For the current results the t-statistic was 2.48 for the Newey West correction with a lag of 6. With a lag of 2 the t-statistic was 2.81. The results are thus close.

Finally, as with stock returns, it is interesting to examine the five administrations omitted from Table 8. Table 9 contains results for these five. The results are quite positive for the Republicans. For $G1$ the D mean is -0.27, the R mean is 3.83, with a difference of -4.11. The t-statistics are -2.38 and -1.95. Similar results are for $G2$. Again, the positive Republican result is the reason the observations added in this paper decrease the overall mean growth difference for the Democrats.

Table 9
Mean Results for Eight Measures of the Economy
5 Administrations: 1916–1928 and 2020
Percentage Points at Annual Rates

	m_D	m_R	m_D -m_R	σ	σ_D	σ_R	t_1	t_2
$G1$	-0.27	3.83	-4.11	2.31	0.96	2.75	-2.38	-1.95
$GP1$	-1.66	2.66	-4.32	1.81	0.58	2.17	-3.27	-2.62
$G2$	-0.79	4.39	-5.17	3.06	0.99	3.69	-2.31	-1.85
$GP2$	-2.18	3.24	-5.42	2.61	0.64	3.16	-2.88	-2.27
ZRS	3.29	2.78	0.50	1.21	1.04	1.28	0.48	0.46
ZP	8.66	-0.42	9.08	3.94	5.93	2.38	2.06	2.53
ZPP	8.91	-1.08	9.98	3.61	4.22	3.27	2.83	3.03
$ZPOP$	1.39	1.18	0.21	0.54	0.38	0.60	0.49	0.44

See notes to Table 6.

5 Adding Ten More Administrations

The Mean Results

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, back to the administration ending in 1876—Grant-2. 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. It is also possible to compute observations on $G1$, $GP1$, $G2$, $GP2$, ZPP, and $ZPOP$ back to the administration ending in 1880—Hayes. For all these administrations the new administration did not begin until March, and so the BW measure, $G2$, is likely more appropriate than $G1$.[3]. Given data on ZPP, GDP deflator inflation, the real return, $R1$-ZPP, can

[3]When creating observations on $G1$, $GP1$, ZPP, and $ZPOP$ for administration 1880, the base period for the log change was taken to be 1877.1 rather than 1876.4 because data prior to 1877 did

Table 10
Eight Measures for Administrations 1876–1912
Percentage Points at Annual Rates

	Last Year	D	$R1$	$R1$ -ZPP	$G1$	$GP1$	$G2$	$GP2$	ZPP	$ZPOP$
1	1876.	0	-8.70							
2	1880.	0	12.23	13.95	8.38	6.34	8.09	5.89	-1.72	2.03
3	1884.	0	-7.42	-5.64	1.60	-0.82	1.88	-0.52	-1.78	2.42
4	1888.	1	4.23	4.12	1.98	-0.21	2.21	0.03	0.11	2.19
5	1892.	0	1.74	3.29	5.85	3.81	5.59	3.56	-1.55	2.04
6	1896.	1	-6.67	-4.30	-0.29	-2.19	1.30	-0.59	-2.37	1.90
7	1900.	0	12.18	9.98	6.16	4.39	6.27	4.49	2.20	1.77
8	1904.	0	4.58	2.69	4.26	2.33	2.95	1.01	1.89	1.93
9	1908.	0	2.26	0.36	3.20	1.29	3.86	1.95	1.90	1.91
10	1912.	0	0.95	-1.11	5.10	3.31	3.96	2.16	2.06	1.79

See notes to Tables 1 and 5.

be computed beginning with administration 1880.

Table 10 contains observations on $R1$ for all 10 administrations and observations on the other measures for the 9 administrations beginning with 1880. It has the same format as Table 1. Only two of the ten 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 36 or 37 administrations are presented in Table 11. This table has the same format as Tables 2 and 6. The t-statistics for $R1$ are about the same as they are in Table 2, and the t-statistics for $R1$-ZPP are about the same as they are in Table A in the appendix. For $R1$ the D mean is 5.87 versus 6.89 in Table 2; the R mean is 3.39 versus 4.11; and the difference is 2.48 versus 2.78. The t-statistics are 0.73, 0.71, and 0.75 versus 0.62, 0.63, and 0.68. The additional observations have not changed the story.

not exist.

Table 11
Mean Results for Eight Measures
Percentage Points at Annual Rates

	m_D	m_R	m_D -m_R	σ	σ_D	σ_R	t_1	t_2	t_3
37 Administrations: 1876–2020									
$R1$	5.87	3.39	2.48	10.50	9.27	11.33	0.73	0.71	0.75
36 Administrations: 1880–2020									
$R1$	5.87	4.00	1.88	10.44	9.27	11.27	0.55	0.54	0.56
$R1$-ZPP	2.60	2.80	-0.20	10.59	10.56	10.62	-0.06	-0.06	-0.05
$G1$	3.48	3.10	0.38	3.53	4.02	3.09	0.31	0.32	0.28
$GP1$	2.22	1.66	0.56	3.51	4.18	2.87	0.46	0.48	0.41
$G2$	3.68	2.95	0.74	3.85	4.26	3.49	0.56	0.57	0.49
$GP2$	2.43	1.50	0.92	3.82	4.42	3.26	0.70	0.72	0.62
ZPP	3.27	1.19	2.08	3.44	3.58	3.32	1.79	1.80	1.57
$ZPOP$	1.26	1.44	-0.18	0.50	0.45	0.53	-1.12	-1.10	-1.34

See notes to Tables 2 and 6.

The t-statistics for the four growth measures in Table 11 are all lower than they are in Table 6. For $G2$ the D mean is 3.68 versus 3.96 in Table 6; the R mean is 2.95 versus 2.03; and the difference is 0.74 versus 1.93. The t-statistics are 0.56, 0.57, and 0.49 versus 1.21, 1.20, and 1.07. The differences in growth means are clearly not significant.

For fun the non parametric test was run for $G2$ in Table 11. There are 36 administrations, 20 R and 16 D. The number of possible different assignments is 7,307,872,110. Of these possibilities, 4,229,352,415 had the absolute value of the computed difference greater than 0.74, the mean difference in Table 7, for a p-value of 0.579. For the t-statistic t_2 in Table 11, 0.57, the p-value is 0.572 (34 degrees of freedom), so again the non parametric test gives almost identical results.

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 11	1876-2020	37 obs.	2.48	10.50
Table 2	1916-2020	27 obs.	2.78	11.39
Table 3	1932-2016	22 obs.	7.84	10.92
$G2$				
Table 11	1880-2020	36 obs.	0.74	3.85
Table 6	1916-2020	27 obs.	1.93	4.18
Table 8	1932-2016	22 obs.	3.43	4.08
Table 7	1952-2012	16 obs.	1.71	1.19

For $R1$ the mean differences get smaller as earlier observations are added, and the overall standard deviation, σ, does not change much. This is also true for $G2$ except for the period 1952–2012, where the mean difference is smaller except for the mean difference for the longest period and the standard deviation is much smaller. This period is clearly an outlier, and the results are sensitive to even adding just the 6 observations in going from Table 7 to Table 8.

6 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 ten administrations between Grant-2 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

The Data Collection

Quarterly data on nominal GDP, real GDP, and population were collected for 1877–2020. For nominal GDP, annual data for 1929–1946 and quarterly data for 1947.1–2020.4 were obtained from the Bureau of Economic Analysis (BEA) website. The data are as of January 28, 2021. 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 1947.1–2020.4 were obtained from the BEA website. 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. Depart-

ment 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 1947.1–2020.4 quarterly data were obtained from the BEA website as of January 28, 2021.

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 1928–2020. 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 1935.1–2020.4 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 1926.01–2019.12 period. Both are monthly percent changes. For the 1912.01–1925.12 and 2020.01–2020.12 periods, 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 1926.01–2019.12 period is 0.957. For VWD it is 0.953. This procedure is thus likely to be a fairly 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 1934.01–2020.12 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.

Monthly data on the CPI for the entire 1912.01–2020.12 period were taken from Shiller's website.

Further Results

For sake of completeness, Table A contains results for other combinations of the stock return measures. This table has the same format as Table 2. The results contain no surprises relative to those in Table 2. The mean real returns are fairly close between D and R, and the t-statistics are all very low.

Table A
Mean Results for Twelve Measures of Stock Returns
27 Administrations: 1916–2020
Percentage Points at Annual Rates

	m_D	m_R	m_D -m_R	σ	σ_D	σ_R	t_1	t_2	t_3
$R2$-ZRS	8.34	3.44	4.91	11.21	9.09	13.13	1.12	1.14	1.25
$R3$-ZRS	8.08	3.50	4.57	10.24	8.48	11.85	1.15	1.16	1.27
$R2$-ZP	6.69	5.80	0.90	11.54	10.45	12.61	0.20	0.20	0.19
$R3$-ZP	6.43	5.86	0.57	10.65	9.84	11.47	0.14	0.14	0.13
$R1$-ZPP	2.98	2.50	0.48	11.82	11.16	12.49	0.10	0.11	0.10
$R2$-ZPP	6.93	6.06	0.86	11.47	10.30	12.63	0.19	0.20	0.18
$R3$-ZPP	6.66	6.13	0.53	10.59	9.68	11.50	0.13	0.13	0.12
$R4$-ZP	3.00	1.55	1.45	11.83	11.03	12.63	0.32	0.32	0.30
$R4$-ZPP	3.23	1.82	1.42	11.76	10.92	12.62	0.31	0.31	0.29
$R5$-ZP	6.26	4.67	1.59	11.88	11.60	12.18	0.35	0.35	0.32
$R5$-ZPP	6.50	4.94	1.55	11.81	11.51	12.13	0.34	0.34	0.31
$R5$-ZRS	7.91	2.32	5.60	11.59	10.08	13.03	1.24	1.25	1.29

$R1$ = S&P 500 stock return, $R2$ = stock return reinvested dividends,
$R3$ = stock return dividends invested in T-bills
$R4$ = VWX return, $R5$ = VWD return,
ZRS = T-bill return, ZP = CPI inflation,
ZPP = GDP Deflator inflation
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.

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.