Abstract

In terms of macroeconomics, national debt and unemployment are two of the most prominent indicators of a country's economic performance. But whether those two indicators always have a clear relationship is left open as a question, though multiple articles such as "The National Debt Is Still a Problem" from the New York Times have already suggested that budget deficit (and thus more debt) is accompanied with unemployment and a badly performing economy.

This paper takes the two of the most prominent economic indices available to the public, national unemployment rate and national debt growth rate. Covering about 50 years' worth of data, each of the data sets were initially run through the Gumbel distribution and examined in relation to the frequency of the values using return periods as a key element. Then, each of the data sets underwent the Holt's Linear Trend method to allow the forecasting of each data set's trend.

National Unemployment Rate Trend Analysis

Instructions:

1) Label column A as 'time' and column B as the 'National Unemployment Rate' and enter them in appropriate values in each cell, which is shown in Table I below.

TABLE I. TIME AND NATIONAL UNEMPLOYMENT RATE

Time	Rate	Time	Rate
01/12/1967	3.8	01/12/2009	9.9
01/12/1968	3.4	01/12/2010	9.3
01/12/1969	3.5	01/12/2011	8.5
01/12/1970	6.1	01/12/2012	7.9
01/12/1971	6	01/12/2013	6.7
01/12/1972	5.2	01/12/2014	5.6

01/12/1973	4.9	01/12/2015	5
01/12/1974	7.2	01/12/2016	4.7
01/12/1975	8.2	01/12/2017	4.1
01/12/1976	7.8	01/12/2018	3.9

2) Select the National Unemployment Rate values in column B then sort the values from smallest to largest by clicking on the ‘Sort and Filter’ tool. Allow expanding the selection. An example can be seen in Table II below.

TABLE II. SORTED NATIONAL UNEMPLOYMENT RATE

Rate	continued...
3.4	7.2
3.5	7.3
3.8	7.3
3.9	7.3
3.9	7.4
4	7.8
4.1	7.9
4.4	8.2
4.4	8.3
4.7	8.5
4.7	8.5
4.9	9.3
4.9	9.9
5	10.8

3) Label column C as ‘Rank (i)’ and rank the data in decreasing order (from N to 1).

4) Create a fourth column called q_i. Gringorten’s plotting position formula will be used to calculate the estimated exceedance probabilities relevant to past observations. An example can be seen in Table III below.

$$q_i = \frac{i-a}{N+1-2a} \qquad (1)$$

q_i = exceedance probability associated with a specific observation

$N =$ number of annual maxima observations

$i =$ Rank of specific observation (i=1 is the largest and i=N is the smallest)

$a =$ constant for estimation (0.44)

TABLE III. q_i OF NATIONAL UNEMPLOYMENT RATE VALUES

Time	Rate	rank	q_i
01/12/1968	3.4	52	0.9640045
01/12/1969	3.5	51	0.94994376
01/12/1967	3.8	50	0.92182227
01/12/2000	3.9	49	0.90776153
01/12/2018	3.9	48	0.89370079
01/12/1999	4	47	0.85151856
01/12/2017	4.1	46	0.83745782
01/12/1998	4.4	45	0.78121485
01/12/2006	4.4	44	0.76715411
01/12/1997	4.7	43	0.73903262
01/12/2016	4.7	42	0.72497188
01/12/1973	4.9	41	0.71091114
01/12/2005	4.9	40	0.69685039
01/12/2007	5	39	0.65466817
01/12/2015	5	38	0.64060742

5) Make another column and label it p_i. Then make it equal to $1 - q_i$. p_i refers to the non-exceedance probability.

B. Statistical Definition of Return Period

- If X is a random variable with a cumulative distribution function $F_x(x)$, the probability that X is less than equal (not exceeding) to a given event x_p is:

$$F_x(x) = P(X \leq x_p) = p \qquad (2)$$

- The probability that this event will be exceeded is now $1 - p$, and the percent exceedance would be $100(1 - p)\%$.
- For an event x_p, the return period corresponding to this exceedance probability is denoted by T.

$$T = \frac{1}{(1-p)} \qquad (3)$$

- For example, a 100-year return period is an event with a probability of exceedance $1 - p = 0.01$ or a non-exceedance probability $p = 0.99$. There is a 99% chance that this event will not be exceeded within a given year.

6) Create one more column and label it 'T_p estimated' and evaluate the values in p_i using the equation for the return period. An example can be seen in Table IV below.

TABLE IV. q_i, p_i, T_p OF NATIONAL UNEMPLOYMENT RATE VALUES

Time	Rate	Rank	q_i	p_i	T_p estimated
01/12/1968	3.4	52	0.9640045	0.0359955	1.03733956
01/12/1969	3.5	51	0.94994376	0.05005624	1.0526939
01/12/1967	3.8	50	0.92182227	0.07817773	1.08480781
01/12/2000	3.9	49	0.90776153	0.09223847	1.10161091
01/12/2018	3.9	48	0.89370079	0.10629921	1.11894273
01/12/1999	4	47	0.85151856	0.14848144	1.17437252
01/12/2017	4.1	46	0.83745782	0.16254218	1.19408999
01/12/1998	4.4	45	0.78121485	0.21878515	1.2800576
01/12/2006	4.4	44	0.76715411	0.23284589	1.30351906

continued...

Time	Rate	Rank	q_i	p_i	T_p estimated
01/12/2008	7.3	11	0.14848144	0.85151856	6.73484849
01/12/1992	7.4	10	0.1344207	0.8655793	7.43933054
01/12/1976	7.8	9	0.12035996	0.87964005	8.30841122

01/12/2012	7.9	8	0.10629921	0.89370079	9.40740741
01/12/1975	8.2	7	0.09223847	0.90776153	10.8414634
01/12/1983	8.3	6	0.07817773	0.92182227	12.7913669
01/12/1981	8.5	5	0.06411699	0.93588302	15.5964912
01/12/2011	8.5	4	0.05005624	0.94994376	19.9775281
01/12/2010	9.3	3	0.0359955	0.9640045	27.78125
01/12/2009	9.9	2	0.02193476	0.97806524	45.5897436
01/12/1982	10.8	1	0.00787402	0.99212598	127

'T_p estimated' is the estimated distribution of 52 years of data. We assume that the data follows a specific distribution to estimate the parameters.

7) We will follow the 'Gumbel' or 'Extreme Value Type 1' distribution. The CDF (Cumulative distribution of function) of the Gumbel distribution is the following:

$$F_x(x) = exp\left[-exp\left(-\frac{x-u}{\alpha}\right)\right] = p \qquad (4)$$

x is the observed National Unemployment Rate data; u and α are the calculated parameters of the distribution. This distribution will allow us to calculate the theoretical estimate of p.

8) Create two columns labeled '$(x-u)/\alpha$' and 'p-theoretical. Using the following equations, calculate $\bar{x}$, s_x, u and α. Table V (below) shows such values that result from the existing data.

$$\underline{x} = \sum_{i=1}^{n} \frac{x_i}{n} \qquad (5)$$

$$s_x{}^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \underline{x})^2 \qquad (6)$$

$$u = \underline{x} - 0.5772\alpha \qquad (7)$$

$$\alpha = \frac{\sqrt{6}s_x}{\pi} \qquad (8)$$

TABLE V. SPECIAL VALUES

$\underline{x}$
6.16078431
$s_x{}^2$
2.75562594
u
5.36062074
α

$\boxed{1.29430278}$

9) Use the National Unemployment Rate values (x) and populate the column $(x-u)/\alpha$ as shown in Table VI below:

TABLE VI. q_i, p_i, T_p ESTIMATED, AND $(x-u)/\alpha$ OF NATIONAL UNEMPLOYMENT RATE VALUES

time	Rate	rank	T_p estimated	$(x-u)/\alpha$
01/12/1968	3.4	52	1.03733956	-1.5148084
01/12/1969	3.5	51	1.0526939	-1.4375467
01/12/1967	3.8	50	1.08480781	-1.2057617
01/12/2000	3.9	49	1.10161091	-1.1285
01/12/2018	3.9	48	1.11894273	-1.1285
01/12/1999	4	47	1.17437252	-1.0512384
01/12/2017	4.1	46	1.19408999	-0.9739767
01/12/1998	4.4	45	1.2800576	-0.7421917
01/12/2006	4.4	44	1.30351906	-0.7421917
01/12/1997	4.7	43	1.35312024	-0.5104066
01/12/2016	4.7	42	1.37936385	-0.5104066
01/12/1973	4.9	41	1.40664557	-0.3558833
01/12/2005	4.9	40	1.43502825	-0.3558833
01/12/2007	5	39	1.52749141	-0.2786216
01/12/1980	7.2	14	5.24483776	1.42113521
01/12/1984	7.3	13	5.66242038	1.49839689
01/12/1991	7.3	12	6.15224914	1.49839689
01/12/2008	7.3	11	6.73484849	1.49839689
01/12/1992	7.4	10	7.43933054	1.57565856
01/12/1976	7.8	9	8.30841122	1.88470526
01/12/2012	7.9	8	9.40740741	1.96196693
01/12/1975	8.2	7	10.8414634	2.19375195
01/12/1983	8.3	6	12.7913669	2.27101363

01/12/1981	8.5	5	15.5964912	2.42553698
01/12/2011	8.5	4	19.9775281	2.42553698
01/12/2010	9.3	3	27.78125	3.04363037
01/12/2009	9.9	2	45.5897436	3.50720042
01/12/1982	10.8	1	127	4.20255549

10) Use the CDF equation from step 7 to calculate the value of p-theoretical.

11) Use the equation used to calculate $'T_p$ estimated' and use it to calculate $'T_p$ theoretical' using the p theoretical values, as seen in Table VII below.

TABLE VII. $(x - u)/\alpha$, P THEORETICAL, AND T_p THEORETICAL OF NATIONAL UNEMPLOYMENT RATE VALUES

(x-u)/a	p theoretical	T_p theoretical
-1.5148084	0.01058254	1.01069573
-1.4375467	0.01484111	1.01506469
-1.2057617	0.03546171	1.03676548
-1.1285	0.04545572	1.04762034
-1.1285	0.04545572	1.04762034
-1.0512384	0.05720053	1.06067094
-0.9739767	0.07076041	1.07614872
-0.7421917	0.12239104	1.13945965
-0.7421917	0.12239104	1.13945965
-0.5104066	0.18900751	1.23305704
-0.5104066	0.18900751	1.23305704
-0.3558833	0.23992211	1.31565464
-0.3558833	0.23992211	1.31565464
-0.2786216	0.26678631	1.36385887

continued...

1.42113521	0.78549611	4.66192007

1.49839689	0.79972437	4.99311869
1.49839689	0.79972437	4.99311869
1.49839689	0.79972437	4.99311869
1.57565856	0.81312432	5.35115099
1.88470526	0.85909668	7.09706473
1.96196693	0.86885273	7.62501597
2.19375195	0.89449351	9.47809014
2.27101363	0.90193979	10.1978162
2.42553698	0.91536662	11.8156683
2.42553698	0.91536662	11.8156683
3.04363037	0.95345643	21.4852466
3.50720042	0.97046423	33.8572559
4.20255549	0.985154	67.3582113

C. Graphing the National Unemployment Rate Frequency Curve

6) Go to ‘insert’ tab and select charts. Plot $'T_p$ estimated’ vs Rate Value. On the same graph, also plot $'T_p$ theoretical’ vs Rate Value. Label the chart title and the axes on the obtained graph. Figure 2 (below) shows how the graph may look like.

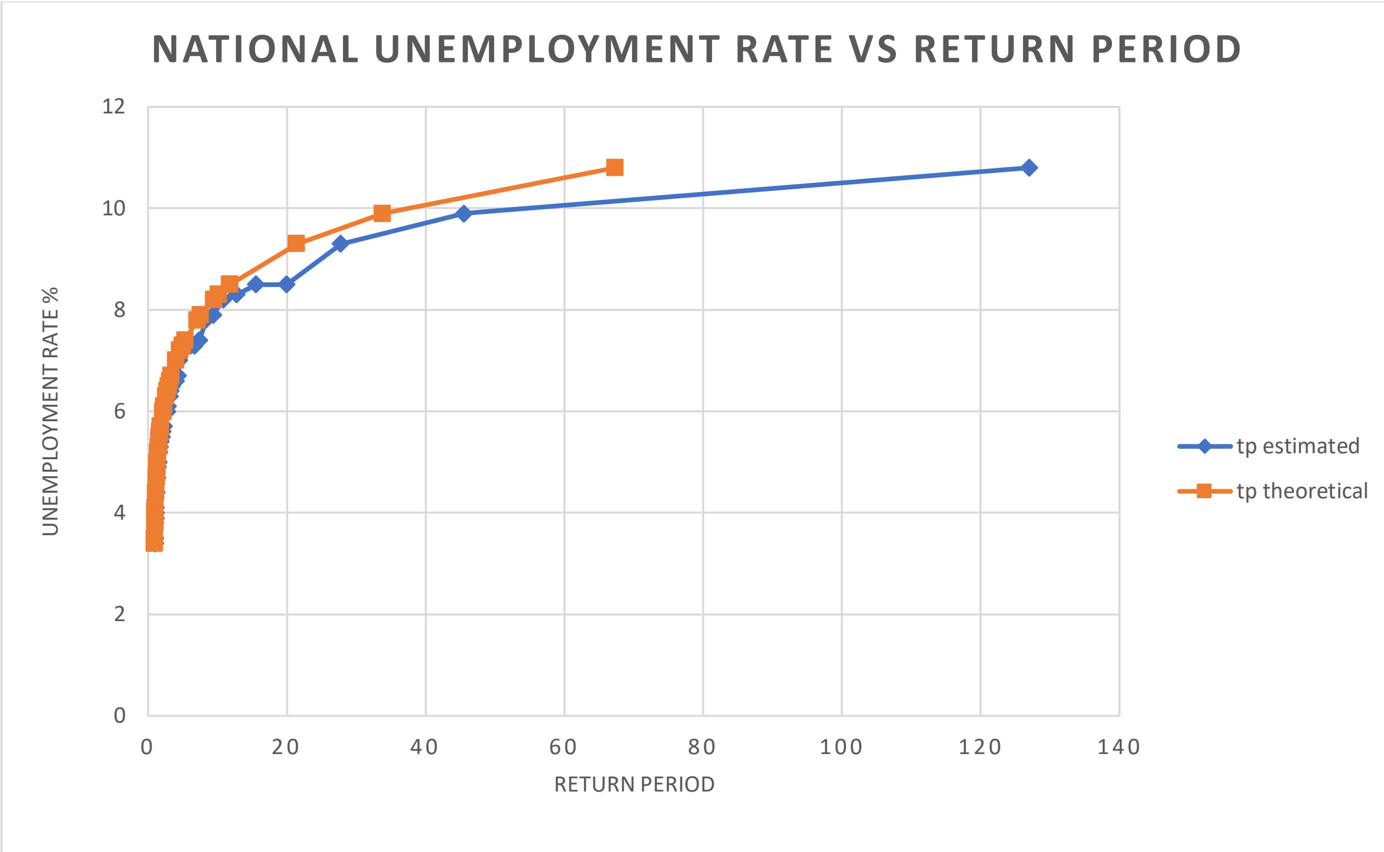

Fig. 1. T_p estimated, T_p theoretical vs National Unemployment Rate Value

7) Right click on the curve and select 'change chart type' and click on 'scatter with smooth lines' for theoretical and 'scatter' for estimated.

8) Right click again and select 'format chart area' to use the axis options command for the X-axis and select logarithmic scale. It will then contain return periods displayed from 1 to 100 in log scale. Figure 3 (below) shows what the graph may look like.

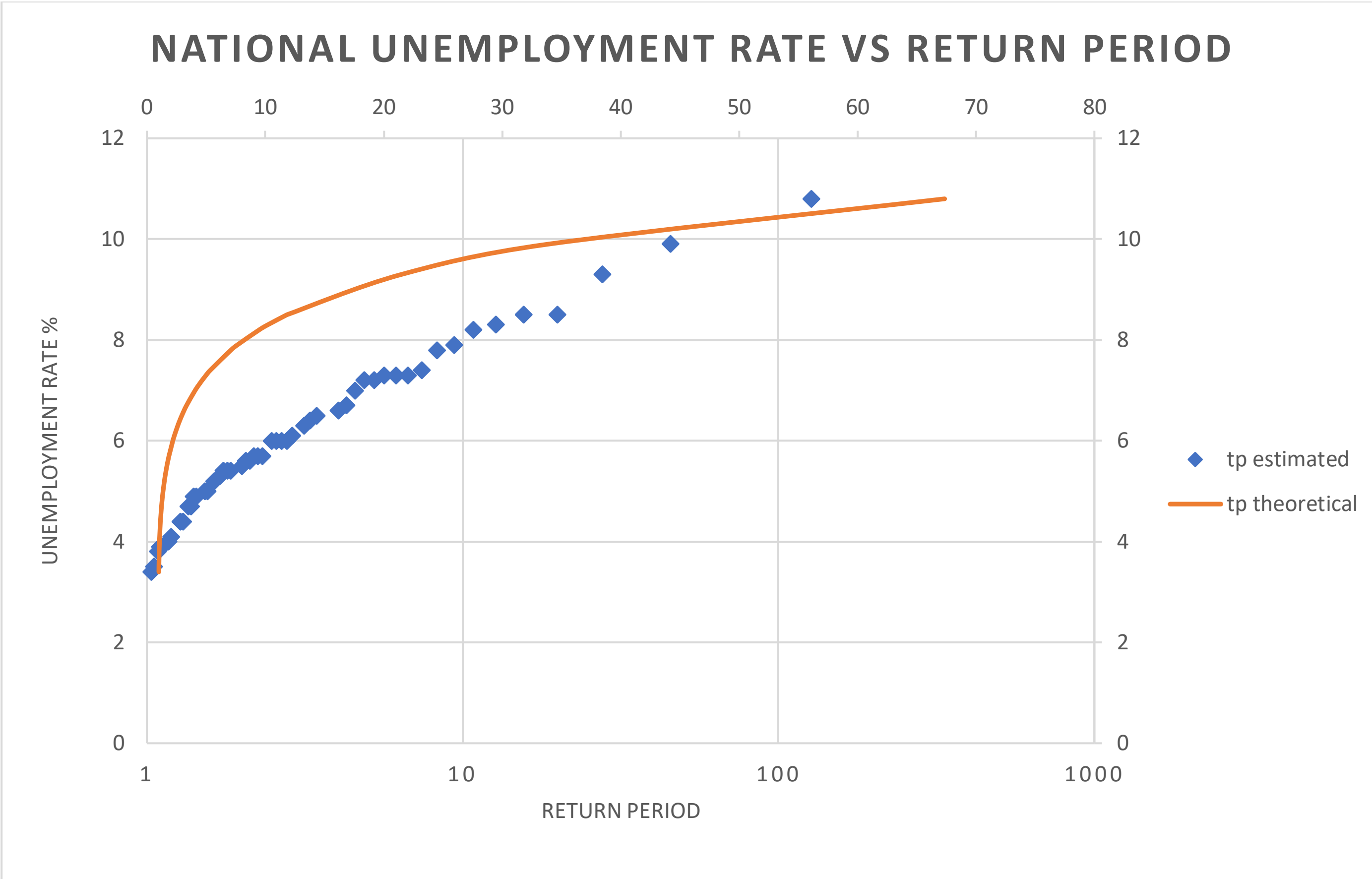

Fig. 2. T_p estimated, T_p theoretical vs National Unemployment Rate Value, Logarithmic Scale

National Debt Growth Rate Trend Analysis

Instructions:

9) Label column A as ‘time’ and column B as the ‘National Debt Growth Rate’ and enter them in appropriate values in each cell, which is shown in Table VIII below.

TABLE VIII. TIME AND NATIONAL DEBT GROWTH RATE

Time	Rate	Time	Rate
01/12/1967	4.66	01/12/2009	15.06
01/12/1968	3.88	01/12/2010	13.92
01/12/1969	2.85	01/12/2011	8.54
01/12/1970	5.68	01/12/2012	7.95
01/12/1971	8.99	01/12/2013	4.4
01/12/1972	5.74	01/12/2014	5.74
01/12/1973	4.59	01/12/2015	4.3
01/12/1974	5.03	01/12/2016	5.57
01/12/1975	17.05	01/12/2017	2.58
01/12/1976	13.33	01/12/2018	7.23

10) Select the National Unemployment Rate values in column B then sort the values from smallest to largest by clicking on the ‘Sort and Filter’ tool. Allow expanding the selection. An example can be seen in Table IX below.

TABLE IX. SORTED NATIONAL DEBT GROWTH RATE

Rate	continued...
-1.97	10.39
2.03	10.59
2.58	12.98
2.85	13.33
2.88	13.82
3.37	13.92

3.88	13.95
3.93	15.06
4.3	15.93
4.4	16.36
4.59	17.02
4.66	17.05
4.97	17.85
5.03	17.88

11) Label column C as 'Rank (i)' and rank the data in decreasing order (from N to 1).

12) Create a fourth column called q_i. Gringorten's plotting position formula will be used to calculate the estimated exceedance probabilities relevant to past observations. An example can be seen in Table X below.

$$q_i = \frac{i-a}{N+1-2a} \tag{1}$$

q_i = exceedance probability associated with a specific observation

$N =$ number of annual maxima observations

$i =$ Rank of specific observation (i=1 is the largest and i=N is the smallest)

$a =$ constant for estimation (0.44)

TABLE X. q_i OF NATIONAL DEBT GROWTH VALUES

Time	Rate	rank	q_i
01/12/2000	-1.97	52	0.98925556
01/12/1998	2.03	51	0.97006907
01/12/2017	2.58	50	0.95088258
01/12/1969	2.85	49	0.93169609
01/12/1999	2.88	48	0.91250959
01/12/1997	3.37	47	0.8933231
01/12/1968	3.88	46	0.87413661

01/12/1995	3.93	45	0.85495012
01/12/2015	4.3	44	0.83576362
01/12/2013	4.4	43	0.81657713
01/12/1973	4.59	42	0.79739064
01/12/1967	4.66	41	0.77820414
01/12/2001	4.97	40	0.75901765
01/12/1974	5.03	39	0.73983116
01/12/2016	5.57	38	0.72064467

13) Make another column and label it p_i. Then make it equal to $1 - q_i$. p_i refers to the non-exceedance probability.

D. Statistical Definition of Return Period

- If X is a random variable with a cumulative distribution function $F_x(x)$, the probability that X is less than equal (not exceeding) to a given event x_p is:

$$F_x(x) = P(X \leq x_p) = p \qquad (2)$$

- The probability that this event will be exceeded is now $1 - p$, and the percent exceedance would be $100(1 - p)\%$.
- For an event x_p, the return period corresponding to this exceedance probability is denoted by T.

$$T = \frac{1}{(1-p)} \qquad (3)$$

- For example, a 100-year return period is an event with a probability of exceedance $1 - p = 0.01$ or a non-exceedance probability $p = 0.99$. There is a 99% chance that this event will not be exceeded within a given year.

14) Create one more column and label it 'T_p estimated' and evaluate the values in p_i using the equation for the return period. An example can be seen in Table XI below.

TABLE XI. q_i, p_i, T_p OF NATIONAL DEBT GROWTH RATE VALUES

Time	**Rate**	**Rank**	q_i	p_i	T_p **estimated**
01/12/2000	-1.97	52	0.98925556	0.01074444	1.01086113
01/12/1998	2.03	51	0.97006907	0.02993093	1.03085443

01/12/2017	2.58	50	0.95088258	0.04911742	1.05165456
01/12/1969	2.85	49	0.93169609	0.06830391	1.07331137
01/12/1999	2.88	48	0.91250959	0.08749041	1.09587889
01/12/1997	3.37	47	0.8933231	0.1066769	1.11941581
01/12/1968	3.88	46	0.87413661	0.12586339	1.14398595
01/12/1995	3.93	45	0.85495012	0.14504988	1.16965889
01/12/2015	4.3	44	0.83576362	0.16423638	1.19651056

continued...

01/12/1976	13.33	11	0.20260936	0.79739064	4.93560606
01/12/1986	13.82	10	0.18342287	0.81657713	5.45188285
01/12/2010	13.92	9	0.16423638	0.83576362	6.08878505
01/12/1990	13.95	8	0.14504988	0.85495012	6.89417989
01/12/2009	15.06	7	0.12586339	0.87413661	7.94512195
01/12/2008	15.93	6	0.1066769	0.8933231	9.37410072
01/12/1982	16.36	5	0.08749041	0.91250959	11.4298246
01/12/1985	17.02	4	0.06830391	0.93169609	14.6404494
01/12/1975	17.05	3	0.04911742	0.95088258	20.359375
01/12/1983	17.85	2	0.02993093	0.97006907	33.4102564
01/12/1984	17.88	1	0.01074444	0.98925556	93.0714286

‘T_p estimated’ is the estimated distribution of 52 years of data. We assume that the data follows a specific distribution to estimate the parameters.

15) We will follow the ‘Gumbel’ or ‘Extreme Value Type 1’ distribution. The CDF (Cumulative distribution of function) of the Gumbel distribution is the following:

$$F_x(x) = exp\left[-exp\left(-\frac{x-u}{\alpha}\right)\right] = p \qquad (4)$$

x is the observed National Debt Growth Rate data; u and α are the calculated parameters of the distribution. This distribution will allow us to calculate the theoretical estimate of p.

16) Create two columns labeled ‘$(x-u)/\alpha$’ and ‘p-theoretical. Using the following equations, calculate $\bar{x}$, s_x, u and α. Table XII (below) shows such values that result from the existing data.

$$\underline{x} = \sum_{i=1}^{n} \frac{x_i}{n} \qquad (5)$$

$$s_x{}^2 = \frac{1}{(n-1)} \sum_{i=1}^{n} (x_i - \underline{x})^2 \qquad (6)$$

$$u = \underline{x} - 0.5772\alpha \qquad (7)$$

$$\alpha = \frac{\sqrt{6} s_x}{\pi} \qquad (8)$$

TABLE XII. SPECIAL VALUES

$\underline{x}$
8.50961538
$s_x{}^2$
21.5234077
u
6.4217255
α
3.61727285

17) Use the National Unemployment Rate values (x) and populate the column $(x-u)/\alpha$ as shown in Table XIII below:

TABLE XIII. q_i, p_i, T_p ESTIMATED, AND $(x-u)/\alpha$ OF NATIONAL DEBT GROWTH RATE VALUES

time	Rate	rank	T_p estimated	$(x-u)/\alpha$
01/12/2000	-1.97	52	1.01086113	-2.3199039
01/12/1998	2.03	51	1.03085443	-1.2140985
01/12/2017	2.58	50	1.05165456	-1.0620502
01/12/1969	2.85	49	1.07331137	-0.9874084
01/12/1999	2.88	48	1.09587889	-0.9791148
01/12/1997	3.37	47	1.11941581	-0.8436537
01/12/1968	3.88	46	1.14398595	-0.7026635
01/12/1995	3.93	45	1.16965889	-0.6888409
01/12/2015	4.3	44	1.19651056	-0.5865539
01/12/2013	4.4	43	1.22462406	-0.5589088
01/12/1973	4.59	42	1.25409047	-0.506383
01/12/1967	4.66	41	1.28500986	-0.4870314

01/12/2001	4.97	40	1.31749242	-0.4013315
01/12/1974	5.03	39	1.35165975	-0.3847444
01/12/1988	10.39	14	3.84365782	1.09703489
01/12/1981	10.59	13	4.14968153	1.15232516
01/12/1991	12.98	12	4.50865052	1.81304391
01/12/1976	13.33	11	4.93560606	1.90980188
01/12/1986	13.82	10	5.45188285	2.04526305
01/12/2010	13.92	9	6.08878505	2.07290819
01/12/1990	13.95	8	6.89417989	2.08120173
01/12/2009	15.06	7	7.94512195	2.38806274
01/12/2008	15.93	6	9.37410072	2.62857542
01/12/1982	16.36	5	11.4298246	2.74744951
01/12/1985	17.02	4	14.6404494	2.9299074
01/12/1975	17.05	3	20.359375	2.93820095
01/12/1983	17.85	2	33.4102564	3.15936203
01/12/1984	17.88	1	93.0714286	3.16765557

18) Use the CDF equation from step 15 to calculate the value of p-theoretical.

19) Use the equation used to calculate 'T_p estimated' and use it to calculate 'T_p theoretical' using the p theoretical values, as seen in Table XIV below.

TABLE XIV. $(x-u)/\alpha$, P THEORETICAL, AND T_p THEORETICAL OF NATIONAL DEBT GROWTH RATE VALUES

(x-u)/a	p theoretical	T_p theoretical
-2.3199039	3.81228E-05	1.00003812
-1.2140985	0.034484095	1.03571572
-1.0620502	0.055448822	1.05870388
-0.9874084	0.068271094	1.07327356
-0.9791148	0.069801569	1.07503944
-0.8436537	0.097798535	1.10839988

-0.7026635	0.132771798	1.15309903
-0.6888409	0.136503372	1.15808211
-0.5865539	0.165665863	1.19856057
-0.5589088	0.17398901	1.21063764
-0.506383	0.190276172	1.23498897
-0.4870314	0.196424386	1.24443796
-0.4013315	0.224515088	1.28951574
-0.3847444	0.230100584	1.29887097

continued...

1.09703489	0.716154359	3.5230416
1.15232516	0.729131343	3.69182618
1.81304391	0.849458029	6.64266578
1.90980188	0.862336489	7.26408901
2.04526305	0.878669752	8.24196781
2.07290819	0.881774131	8.45838571
2.08120173	0.882690923	8.52448954
2.38806274	0.912280868	11.4000217
2.62857542	0.930362283	14.3600343
2.74744951	0.937919535	16.1081269
2.9299074	0.947998857	19.2303464
2.93820095	0.948417073	19.3862595
3.15936203	0.958435687	24.0591011
3.16765557	0.958771801	24.2552433

E. Graphing the National Debt Growth Rate Frequency Curve

20) Go to 'insert' tab and select charts. Plot $'T_p$ estimated' vs Rate Value. On the same graph, also plot $'T_p$ theoretical' vs Rate Value. Label the chart title and the axes on the obtained graph. Figure 3 (below) shows how the graph may look like.

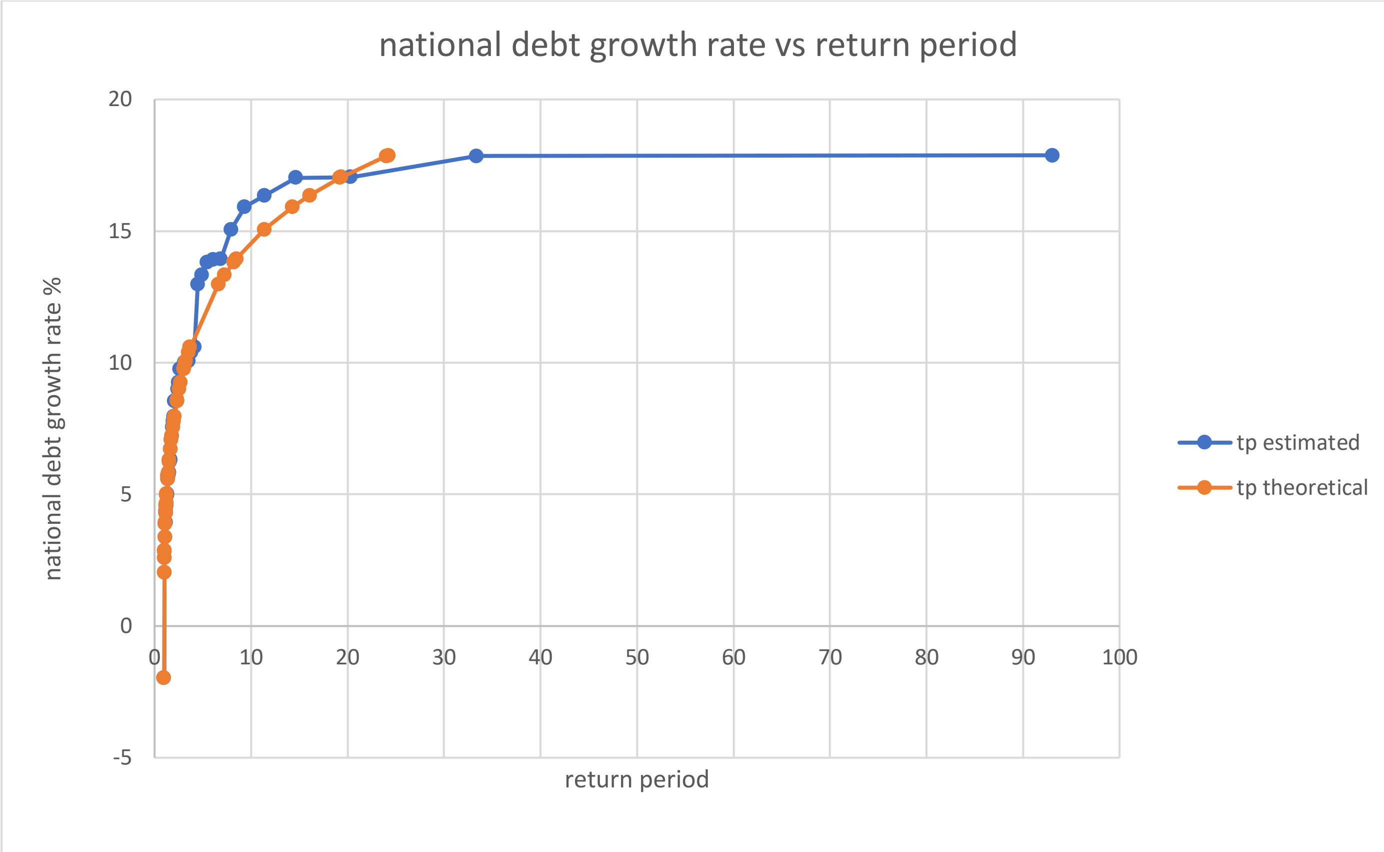

Fig. 3. T_p estimated, T_p theoretical vs National Debt Growth Rate Value

21) Right click on the curve and select 'change chart type' and click on 'scatter with smooth lines' for theoretical and 'scatter' for estimated.

22) Right click again and select 'format chart area' to use the axis options command for the X-axis and select logarithmic scale. It will then contain return periods displayed from 1 to 100 in log scale. Figure 4 (below) shows what the graph may look like.

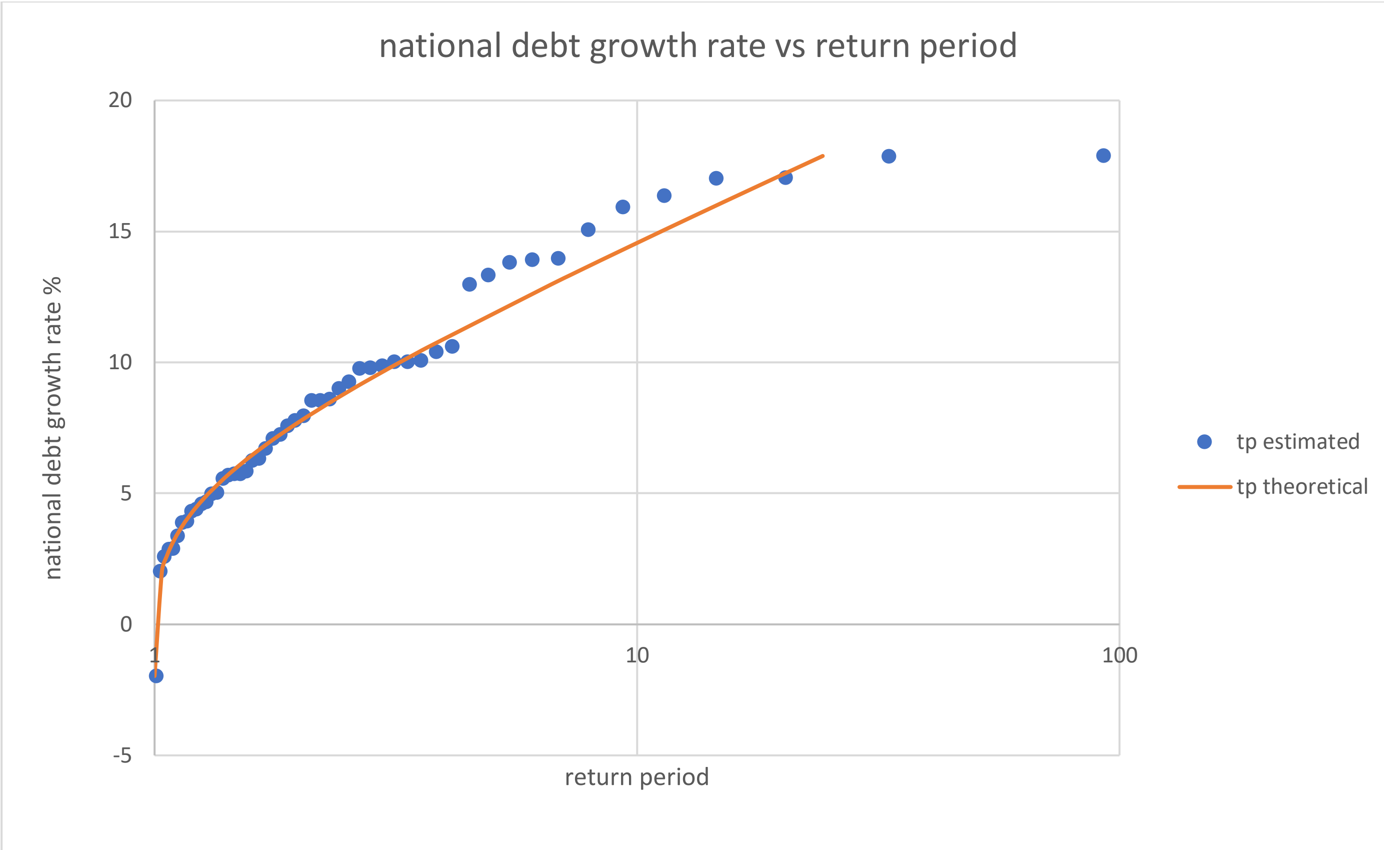

Fig. 4. T_p estimated, T_p theoretical vs National Debt Growth Rate Value, Logarithmic Scale

Holt's Linear Trend

Holt's Linear Trend uses the following equations:

$$u_1 = y_1$$

$$v_1 = 0$$

$$u_i = \alpha y_i + (1-\alpha)(u_{i-1} + v_{i-1})$$

$$v_i = \beta(u_i - u_{i-1}) + (1-\beta)v_{i-1}$$

$$\hat{y}_{i+1} = u_i + v_i$$

Where:

$$0 < \alpha \leq 1$$

$$0 < \beta \leq 1$$

Holt's Linear Trend for National Debt Growth Rate

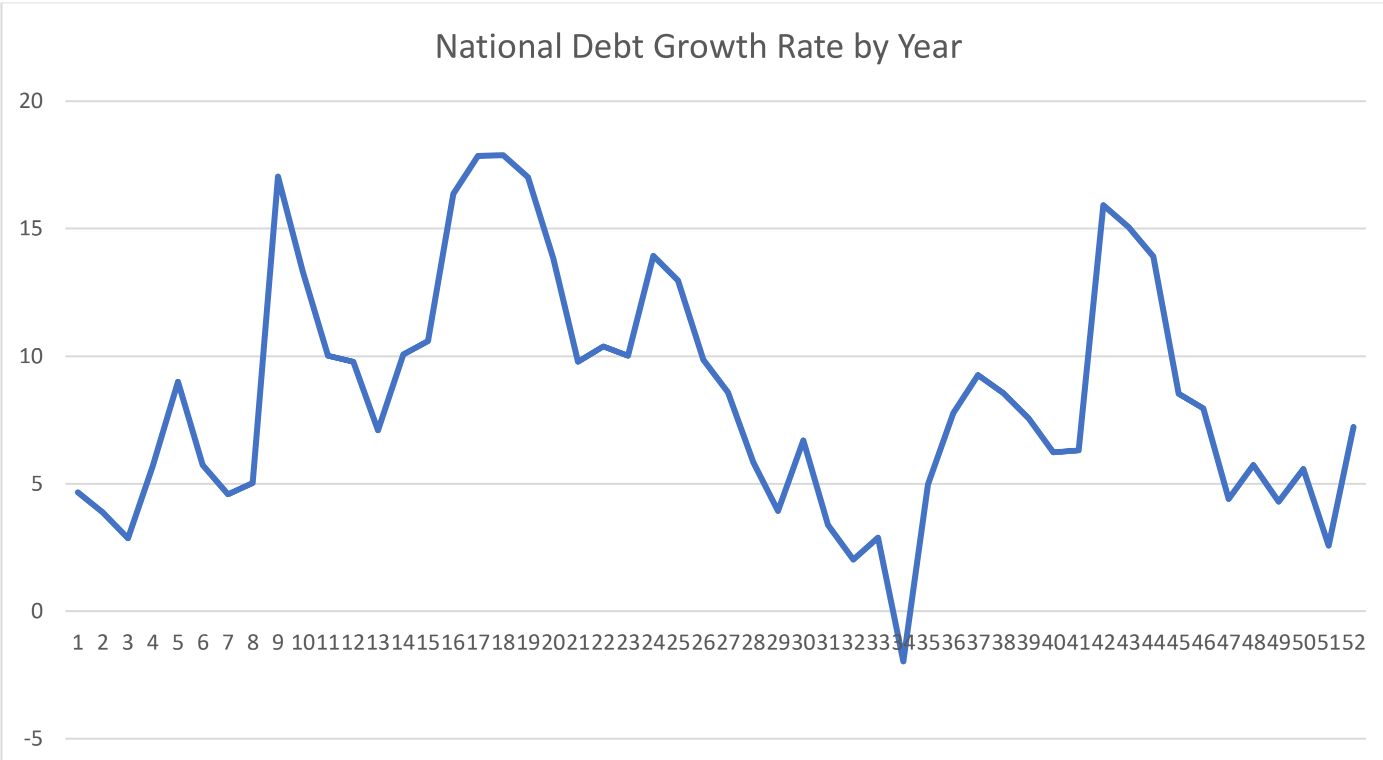

Fig. 5. National Debt Growth Rate (%) vs Years After 1966

1) Label column A as 't', column B as 'Date', and column C as 'Y' (National Debt Growth Rate) and enter them in appropriate values in each cell, which is shown in Table XV below.

TABLE XV. T AND NATIONAL DEBT GROWTH RATE

T	Date	Y (rate)
1	01/12/1967	4.66
2	01/12/1968	3.88
3	01/12/1969	2.85
4	01/12/1970	5.68
5	01/12/1971	8.99
6	01/12/1972	5.74
7	01/12/1973	4.59

8	01/12/1974	5.03
9	01/12/1975	17.05
10	01/12/1976	13.33

2) Create a fourth column called u and a fifth column called v. Then, create a sixth column called 'pred', which measures $\hat{y}_i$. Using the equations for Holt's Linear Trend, fill out the column values. Assume $\alpha = 0.4$ and $\beta = 0.7$.

TABLE XVI. $u_i, v_i, \hat{y}_i$ OF NATIONAL DEBT GROWTH RATE VALUES

t	*Date*	*y*	*u*	*v*	*pred*
1	01/12/1967	4.66	4.66	0	
2	01/12/1968	3.88	4.348	-0.2184	4.1296
3	01/12/1969	2.85	3.61776	-0.576688	3.041072
4	01/12/1970	5.68	4.0966432	0.16221184	4.25885504
5	01/12/1971	8.99	6.15131302	1.48693243	7.63824545
6	01/12/1972	5.74	6.87894727	0.9554237	7.83437097
7	01/12/1973	4.59	6.53662258	0.04699983	6.58362241
8	01/12/1974	5.03	5.96217345	-0.3880144	5.574159
9	01/12/1975	17.05	10.1644954	2.82522103	12.9897164

continued...

42	01/12/2008	15.93	10.5642799	1.44198947	12.0062694
43	01/12/2009	15.06	13.2277616	2.29703404	15.5247957
44	01/12/2010	13.92	14.8828774	1.84769125	16.7305687
45	01/12/2011	8.54	13.4543412	-0.445668	13.0086732
46	01/12/2012	7.95	10.9852039	-1.8620965	9.12310746
47	01/12/2013	4.4	7.23386448	-3.1845666	4.04929791
48	01/12/2014	5.74	4.72557875	-2.71117	2.01440877
49	01/12/2015	4.3	2.92864526	-2.0712044	0.85744083
50	01/12/2016	5.57	2.7424645	-0.7516879	1.99077663
51	01/12/2017	2.58	2.22646598	-0.5867053	1.63976066

52	01/12/2018	7.23	3.87585639	0.97856169	4.85441809

3) Using the fact that $|e| = |y - \hat{y}|$, calculate $|e|$ and e^2, as seen in Table XVII below.

TABLE XVII. $y, \hat{y}_i, |e|, e^2$ OF NATIONAL DEBT GROWTH RATE VALUES

y	***pred***	**|*e*|**	***e^2***
4.66			
3.88	4.1296	0.2496	0.06230016
2.85	3.041072	0.191072	0.03650851
5.68	4.25885504	1.42114496	2.019653
8.99	7.63824545	1.35175455	1.82724036
5.74	7.83437097	2.09437097	4.38638978
4.59	6.58362241	1.99362241	3.97453033
5.03	5.574159	0.544159	0.29610902
17.05	12.9897164	4.06028357	16.4859026
13.33	16.0463303	2.71633029	7.37845026
10.01	14.8621261	4.85212612	23.5431279
9.77	12.6298083	2.85980831	8.17850357
7.08	8.66047129	1.58047129	2.49788951
10.07	7.86953712	2.20046288	4.84203688

continued...

7.56	11.6016483	4.04164825	16.3349206
6.24	9.20191072	2.96191072	8.77291511
6.32	6.9871332	0.6671332	0.4450667
15.93	12.0062694	3.92373061	15.3956619
15.06	15.5247957	0.46479568	0.21603502
13.92	16.7305687	2.81056866	7.89929618
8.54	13.0086732	4.46867322	19.9690404
7.95	9.12310746	1.17310746	1.37618111
4.4	4.04929791	0.35070209	0.12299195

5.74	2.01440877	3.72559123	13.88003
4.3	0.85744083	3.44255917	11.8512137
5.57	1.99077663	3.57922337	12.8108399
2.58	1.63976066	0.94023934	0.88405002
7.23	4.85441809	2.37558191	5.64338942

4) MAE is the average of $|e|$ values, MSE is the average of e^2 values, and RMSE is the square root of MSE. Using these facts, calculate MAE, MSE, and RMSE values.

TABLE XVIII. SPECIAL VALUES

MAE
2.23822952
MSE
6.89800775
RMSE
2.62640586

Graphing Holt's Linear Trend

5) Go to 'insert' tab and select charts. Plot 't' vs Rate Value. On the same graph, also plot 't' vs $\hat{y}_i$ (pred). Label the chart title and the axes on the obtained graph. Figure 6 (below) shows how the graph may look like.

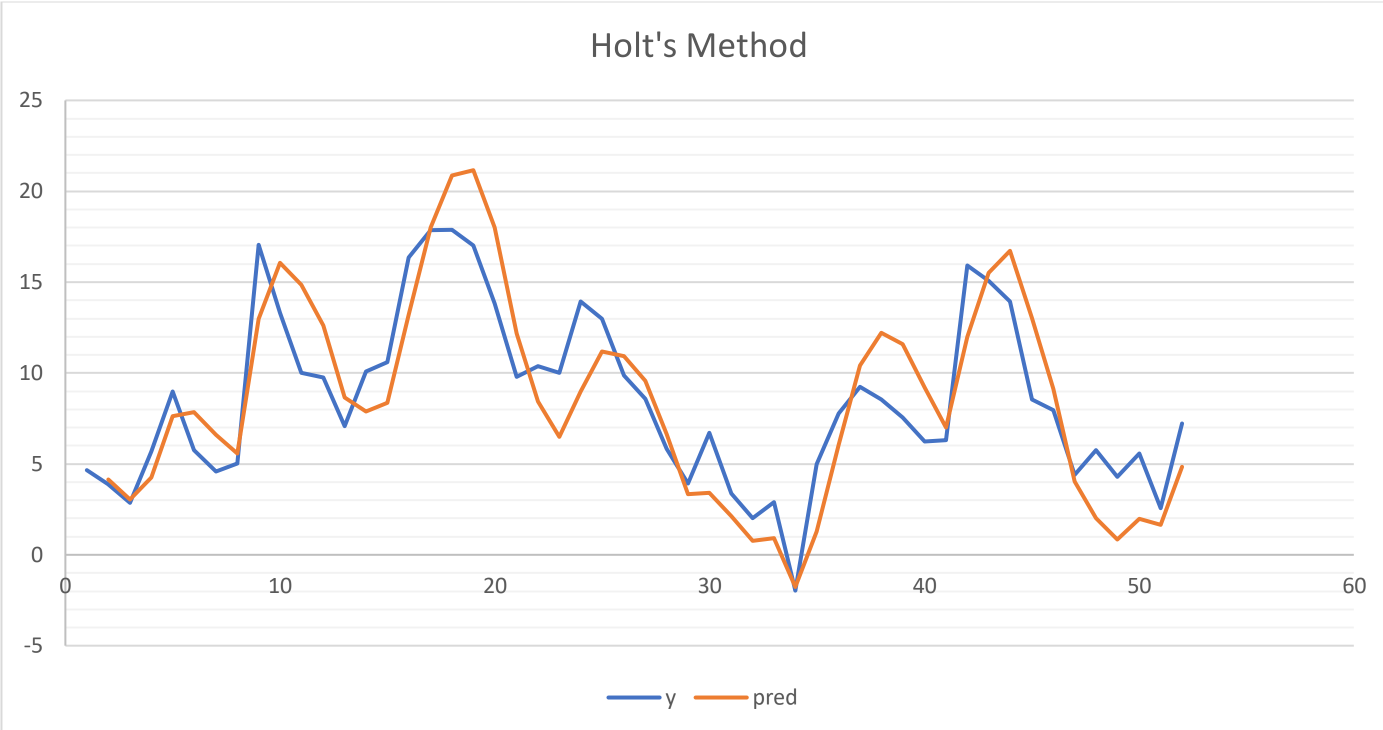

Fig. 6. $y, \hat{y}$ vs t

Overall, the Holt's linearization works as a fairly adequate forecast of the National Debt Growth Rate from 1967 to 2018. Especially in periods from 1968 to 1972, 1974 to 1976, 1982 to 1983, 1992 to 2003, and 2007 to 2012, the similarity is evident. There are discrepancies in multiple periods, however, in periods such as from 1983 to 1985, 1988 to 1991, 2004 to 2007, and 2013 to 2016. However, the maximum discrepancy is about 4 to 5 percentage points (about 20% in a relative scale), so it is not very big of a margin.

Holt’s Linear Trend for National Unemployment Rate

Fig. 7. National Unemployment Rate (%) vs Years After 1966

6) Label column A as ‘t’, column B as ‘Date’, and column C as ‘Y’ (National Unemployment Rate) and enter them in appropriate values in each cell, which is shown in Table XIX below.

TABLE XIX. T AND NATIONAL UNEMPLOYMENT RATE

T	Date	Y (rate)
1	01/12/1967	3.8
2	01/12/1968	3.4
3	01/12/1969	3.5
4	01/12/1970	6.1
5	01/12/1971	6
6	01/12/1972	5.2

7	01/12/1973	4.9
8	01/12/1974	7.2
9	01/12/1975	8.2
10	01/12/1976	7.8

7) Create a fourth column called u and a fifth column called v. Then, create a sixth column called 'pred', which measures $\hat{y}_i$. Using the equations for Holt's Linear Trend, fill out the column values. Assume $\alpha = 0.4$ and $\beta = 0.7$.

TABLE XX. $u_i, v_i, \hat{y}_i$ OF NATIONAL UNEMPLOYMENT RATE VALUES

t	*Date*	*y*	*u*	*v*	*pred*
1	01/12/1967	3.8	3.8	0	
2	01/12/1968	3.4	3.64	-0.112	3.528
3	01/12/1969	3.5	3.5168	-0.11984	3.39696
4	01/12/1970	6.1	4.478176	0.6370112	5.1151872
5	01/12/1971	6	5.46911232	0.88475878	6.3538711
6	01/12/1972	5.2	5.89232266	0.56167487	6.45399754
7	01/12/1973	4.9	5.83239852	0.12655556	5.95895409
8	01/12/1974	7.2	6.45537245	0.47404842	6.92942087
9	01/12/1975	8.2	7.43765252	0.82981058	8.2674631

continued...

42	01/12/2008	7.3	5.62553032	0.40257387	6.02810418
43	01/12/2009	9.9	7.57686251	1.4867047	9.0635672
44	01/12/2010	9.3	9.15814032	1.55290588	10.7110462
45	01/12/2011	8.5	9.82662772	0.93381294	10.7604407
46	01/12/2012	7.9	9.6162644	0.13288956	9.74915395
47	01/12/2013	6.7	8.52949237	-0.7208736	7.80861882
48	01/12/2014	5.6	6.92517129	-1.3392868	5.58588447
49	01/12/2015	5	5.35153068	-1.5033345	3.84819621
50	01/12/2016	4.7	4.18891773	-1.2648294	2.92408831

51	01/12/2017	4.1	3.39445299	-0.9355741	2.45887885
52	01/12/2018	3.9	3.03532731	-0.5320602	2.50326709

8) Using the fact that $|e| = |y - \hat{y}|$, calculate $|e|$ and e^2, as seen in Table XXI below.

TABLE XXI. $y, \hat{y}_i, |e|, e^2$ OF NATIONAL UNEMPLOYMENT RATE VALUES

y	***pred***	**|e|**	***e^2***
3.8			
3.4	3.528	0.128	0.016384
3.5	3.39696	0.10304	0.01061724
6.1	5.1151872	0.9848128	0.96985625
6	6.3538711	0.3538711	0.12522476
5.2	6.45399754	1.25399754	1.57250982
4.9	5.95895409	1.05895409	1.12138376
7.2	6.92942087	0.27057913	0.07321306
8.2	8.2674631	0.0674631	0.00455127
7.8	8.77939877	0.97939877	0.95922195
6.4	7.86032851	1.46032851	2.13255937
6	6.62799438	0.62799438	0.39437694
6	5.71275547	0.28724453	0.08250942
7.2	6.06004059	1.13995941	1.29950745

continued...

4.9	5.80753389	0.90753389	0.82361776
4.4	4.81343752	0.41343752	0.17093058
5	4.50921719	0.49078281	0.24086776
7.3	6.02810418	1.27189582	1.61771897
9.9	9.0635672	0.8364328	0.69961982
9.3	10.7110462	1.4110462	1.99105138
8.5	10.7604407	2.26044066	5.10959199
7.9	9.74915395	1.84915395	3.41937034

6.7	7.80861882	1.10861882	1.22903569
5.6	5.58588447	0.01411553	0.00019925
5	3.84819621	1.15180379	1.32665197
4.7	2.92408831	1.77591169	3.15386232
4.1	2.45887885	1.64112115	2.69327863
3.9	2.50326709	1.39673291	1.95086282

9) MAE is the average of $|e|$ values, MSE is the average of e^2 values, and RMSE is the square root of MSE. Using these facts, calculate MAE, MSE, and RMSE values.

TABLE XXII. SPECIAL VALUES

MAE
0.90901958
MSE
1.12228894
RMSE
1.05938139

Graphing Holt's Linear Trend

10) Go to 'insert' tab and select charts. Plot 't' vs Rate Value. On the same graph, also plot 't' vs $\hat{y}_i$ (pred). Label the chart title and the axes on the obtained graph. Figure 8 (below) shows how the graph may look like.

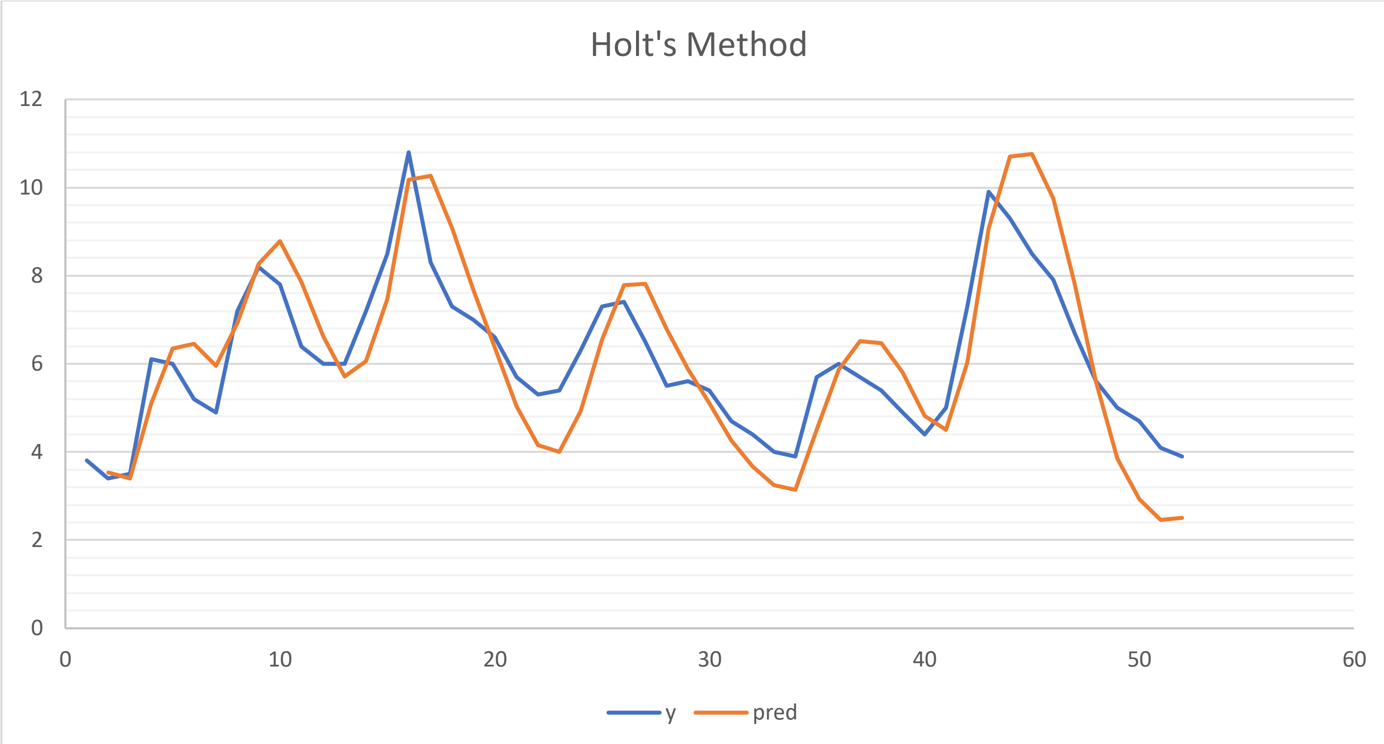

Fig. 8. $y, \hat{y}$ vs t

Overall, the Holt's linearization works as a decent forecast of the National Unemployment Rate from 1967 to 2018 – in fact, better than the previous case of National Debt Growth Rate. Especially in periods from 1968 to 1971, 1974 to 1976, 1979 to 1982, 2001 to 2002, and 2006 to 2010, the similarity is evident. There are discrepancies in multiple periods, however, most notable in periods such as from 1983 to 1992, 2002 to 2006, and 2010 to 2018. However, the maximum discrepancy is about 2 percentage points (about 16% in a relative scale), so it is not very big of a margin.

Discussion

In economics, multiple types of data, such as stock prices, interest rates, GDP, and other indices have been collected over time. Because these data sets display change as time progresses, time series analysis is necessary to be able to analyze such sequential data by using math, statistics, and computation.

In order to prevent any noise from disturbing the data, linear and nonlinear smoothers were applied to the data to make the time series analysis more efficient and accurate.

In MATLAB®, we can use the least sum of squares line to get the best-fit line for the data. This format of linear modeling is named so because of its method of minimizing the sum of the squares of the residuals, the deviations between the model and the actual data. Because certain data points may be lower than the ones predicted by the model, the deviations are squared to account for negative residuals. Also, curve fitting involves interpolation, where the model is an exact fit to the data, or smoothing, where the noise found in the data is reduced and a function is able to approximate the overall trend of the data. Curve fitting can not only visualize the data and make it computable, but also enable the extrapolation of other data points based on the trend suggested by the model.

For time series analysis and forecasting, we used the Gumbel distribution and the concept of return periods to map out how frequently (or how unfrequently) certain data values appear. We also used Holt's linearization method to get a model of how the two sets of data change over time. The data for national unemployment rate had a Holt's linearization model that had better accuracy but had a theoretical return period evaluation that was relatively less accurate. The data for national debt growth rate had a Holt's linearization model that had relatively less accuracy but had a theoretical return period evaluation that was more accurate.