Non-Stationarity in U.S. All-Cause Mortality Rates: Probing The Usefulness of the Idiom 'Excess Death'

Mitch Kunce

DouglasMitchell Econometric Consulting, Laramie WY USA

Citation: Mitch Kunce (2022) Non-Stationarity in U.S. All-Cause Mortality Rates: Probing The Usefulness of the Idiom 'Excess Death', *European Journal of Statistics and Probability*, Vol.10, No.2, pp., 1-13

ABSTRACT: Borrowing a bit from the author Yuval Harari – death is a chaos not particularly influenced by predictions made about it. This paper examines the non-stationarity of aggregate U.S. age standardized all cause mortality rates over the period 1968-2021. Both univariate and state-level panel unit root tests confirm that the underlying stochastic process generating U.S. mortality rates changes over time. Examining non-stationary death in the aggregate, controlling for age and population, establishes proper context to scrutinize the usefulness of the idiom 'excess death'.

KEYWORDS: age standardized death rates, unit root, panel data, excess death **JEL CLASSIFICATIONS:** C22, C23, C53

INTRODUCTION

In 1798, President John Adams signed legislation that commissioned the U.S. Marine Hospital Service to provide care and relief for sick and injured seamen of the merchant fleet. At the time, the merchant fleet was vital to the nation's national defense and was steadily becoming the heart of the country's economic lifeline. By mid 1870, the Marine Hospital Service was reorganized as a national hospital system led by a supervising surgeon who was eventually given the title 'Surgeon General'. This reorganization paved the way for the statutory beginnings of the Public Health Service, circa 1912. At this time the U.S. Marine Hospital Service became the U.S. Public Health Service (PHS) reflecting the expanded functions of the service including, but not limited to, infectious disease prevention, hygiene research and water pollution detection and study. The Public Health Service Act of 1944 helped integrate and extend the many authorities of the service and divided the PHS into four bureaus: the Office of the Surgeon General, the National Institutes of Health, the Bureau of Medical Services and the Bureau of State Services (source: Department of Health, Education and Welfare 1976).

Concurrently in the 1790s, city-level health and sanitary boards emerged in response to outbreaks of disease. From meager beginnings, these boards began collecting vital statistics identifying disease and death. The next 50 years saw a rather haphazard interest

in perfecting the country's vital registration and vital statistics. Beginning in 1850, heavily influenced by Lemuel Shattuck of Massachusetts,¹ the U.S. Census adopted the collection of and published the first national death statistics. These counts were based on information collected by enumerators during the decennial census. Because of the 10 year time intervals, these data suffered from inaccuracy and incompleteness. By 1890, the U.S. Census set up select regional registration areas for deaths. Mortality data coming from these registration areas appeared to be more accurate, but counts for the entire country were still compiled by census enumerators. The move to collect death data on an annual basis began in 1902 through legislation. The now Bureau of the Census, Division of Vital Statistics could obtain records from those states with adequate death registration systems. At the time, the registration area covered roughly 40 percent of the U.S. population. By 1933 the registration area encompassed the entire country. Beginning in 1940, statistical series were forming for registration states allowing general comparisons year to year – though data quality and completeness continued to be an issue. In the summer of 1946, the vital statistics responsibility of the Bureau of the Census was transferred to the U.S. Public Health Service, Office of the Surgeon General. There, the National Office of Vital Statistics was established. This office produced the annual Vital Statistics of the U.S. and in the early 1950s began publishing Morbidity and Mortality *Weekly Report* (MMWR).² (source: Department of Health, Education and Welfare 1967).

Correspondingly, in the summer of 1946, the malaria control program under the PHS, Bureau of State Services was transitioned into the Communicable Disease Center (CDC 2018). By 1960, the CDC had become a major operating component of the newly formed, cabinet-level, Department of Health, Education and Welfare. In January of 1961, the CDC assumed responsibility for the publication, Morbidity and Mortality Weekly Report, from the National Office of Vital Statistics. Just beginning its tenth year, the MMWR was comprised of a few short analytical reports, notifiable disease case tables and the weekly morbidity and mortality tables and graphs for select cities of regions in the U.S.³ Specifically, the January 7, 1961 MMWR documented death certificate counts for 117 major cities grouped into 9 geographic divisions. Along with the death counts for a given week, the report provided a simple 5-week moving average and an adjusted average for comparison. The adjusted average was computed as follows: "From the total deaths reported each week for the years 1956-1960, three central figures were selected by eliminating the highest and lowest figure reported for that week. A 5-week moving average of the arithmetic mean of the three central figures was then computed with adjustment to allow for population growth" (CDC 1961, p. 6). For the first week ended

@ECRTD-UK: https://www.eajournals.org/

¹ Key author of the *Report of the Sanitary Commission of Massachusetts, 1850.* Arguably the most influential figure in U.S. vital registration and public health.

² Effective in 1963, the National Office of Vital Statistics became part of the newly formed National Center for Health Statistics (NCHS).

³ As part of the Preventive Health Amendments of 1992, the Communicable Disease Center formally became the Centers for Disease Control. The MMWR is still published today by the Centers for Disease Control (CDC).

January 7, 1961, the CDC reported 12,650 deaths in the 117 cities monitored. Compared to the adjusted average first week (1956-1960) of 12,311, actual deaths were 2.8% above what was shown in past death counts, during the same week, adjusted for population (CDC 1961, p. 6, Table 3).

What was referred to as a comparison, 2.8% above the computed adjusted average in January of 1961, has evolved into the phrasing 'excess death' of late. According to the CDC, excess death is now a 'standard' metric defined as observed mortality exceeding predictions from historical experience (CDC 2022). The idiom appeared intermittently over the last 60 years in work mostly associated with the CDC. The bulk of these 'excess death' examinations focused on specific causes of death, demographic groups and/or particular situations like the lack of access to adequate health care (e.g. Ayala 2000, Flegal et al 2018, CDC 2021). Comparable predictions, with few exceptions⁴, were derived from raw death counts not standardized rates of death. For example, Ahmad et al (2022) referenced forecasted weekly deaths, from all causes, for the year 2021 derived from weekly death counts over the period 2014-2020. In any case, from March 2020 forward, the phrase 'excess death' has become common vernacular (e.g. Achenbach 2022).

The methods to predict and forecast mortality have likewise evolved. Classic statistical forecasting methods, such as seasonal autoregressive integrated moving average (sARIMA), appear as relics when compared to new machine learning platforms like Prophet or Greykite (Wang et al 2022). Accurately forecasting non-stationary time series involves multiple complexities (Grillenzoni 1998). Time varying stochastic processes have no problem drifting away – while previous data points (lags) may not provide relevant information for predictions. Crucial issues include: (i) consistency of unit root assumptions, (ii) feasibility of unit root modeling, (iii) relationships between trend and seasonal components and (iv) effects of data transformations (differencing, etc.). Modern day death forecasting may indeed be much improved over what was convention in 1961, but death is a form of chaos that does not react to predictions made about it (Harari 2014).

Interestingly, the literature is deficient in analyses of non-stationarity (unit root) in death and/or death rate time-series. Herein, we will focus on annual, U.S. all cause age standardized rates of mortality per 1,000 population. Examining death in the aggregate, controlling for age and population, establishes proper context to scrutinize the usefulness of the idiom 'excess death'. The paper proceeds as follows; the next section examines non-stationarity in the univariate death rate time-series for the entire U.S. Section three bolsters the univariate analysis with a panel treatment including all 50 states and the District of Columbia. Section four concludes with remarks probing the usefulness of the phrase 'excess death'.

⁴ The annual United Nations: World Population Prospects does forecast mortality rates for the U.S. and other nations.

SINGLE EQUATION UNIT ROOT TEST

The data for this first set of unit root tests represents the age standardized rates of death per 1,000 population for the U.S. aggregate from 1968 to 2021. The National Center for Health Statistics (NCHS) provides this data publicly through the CDC/NCHS website. As mentioned in the introduction, death certificate counts prior to 1968 were faulty rendering the older faction of the series non-conforming to the more recent. Age standardized (adjusted) rates are commonly used in the mortality literature to compare relative indexes across groups and over time. Year 2000 population weights apply. At the time of final composition of this paper (June 2022), the 2021 death counts were provisional, though comparative to what the final tallies will eventually show ⁵. Figure 1 depicts the crude and age standardized rates over the sample time period. In order to reduce time dependence in the variance of the data, on the level, the rates are transformed into natural logarithms.



Interestingly, the age standardized log rate in 2021 is substantially (17%) lower than rates from the late 1960s.⁶ The U.S. death rate graph depicts a possible structural break towards the end of the time-series. This potential structural deviation calls for further scrutiny. Structural change and unit roots are closely related and unit root tests are biased toward a false unit root null when the data may be trend stationary exhibiting structural breaks (Perron 1989). Unit root with structural break tests were performed on the aggregate U.S. time-series. Test results found in Table 1 indicate a single structural break at or about the year 2019. Accordingly, we separated a 1968-2019 sub-sample for

⁵ Ahmad et al (2022) reported that the 2021 death data as of April 12, 2022 was greater than 99% complete. They derive an age-standardized rate of 8.42 for 2021.

⁶ Non-log age standardized rates declined from 13.04 in 1968 to 8.43 in 2021.

structural comparison to the full sample. The unit root with structural break test fails to reject the null hypothesis of non-stationarity.

Table 1. Unit Koot with Break Test							
Minimize Dickey-Fuller t-statistic							
ADF, Constant	Lag: 1	-2.85					
ADF, Constant, Trend	Lag: 3	-2.96					
Break Date 2019							
-Break type, Innovational Outlier. Akaike information criterion lag selection (AIC).							

Table 1 Unit Root with Break Test

Tables 2 and 3 depict descriptive statistics and unit root test results for both the sub- and full-sample time-series. Ng and Perron (2001) provide four (NP M-GLS) tests denoted MZ α , MZt, MSB and MPt for investigating the presence of unit roots. The MZ α and MZt test statistics are obtained by modifying the Phillips (1987) and Phillips and Perron (1988) Za and Zt tests. MSB is derived from the Bhargava (1986) R-test and lastly the MPt test stat is adopted from the point optimal work of Elliot et al (1996). All four tests fail to reject the null of unit root for each sample and model specification. Nonstationarity is also confirmed by more conventional testing. Both the augmented Dickey and Fuller (1979, 1981) (ADF) and Kwiatkowski et al (1992) (KPSS) tests affirm unit root in both U.S. death rate time-series. The structural break had little effect on the test results, however, could indicate the start of a new drift for death rates into the future.

Table 2.	Unit R	oot Te	sts, 196	8-2019		
	19	68-2019				
	Α	ge				
	Stand	ardize	LN			
Mean	9.	27	2.21			
STD	1.	.60	0.17			
Observations	5	52		-		
				-		
	Lag	MZα	MZt	MSB	MPT	
NP, Constant	3	0.89	0.74	0.83	48.62	
NP, Constant, Trend	1 -4.18		-1.32	0.32	20.57	
-Null hypothesis, Unit ro	ot. AIC l	ag select	tion.			
ADF, Constant	-2	-2.23		AIC		
ADF, Constant, Trend	-2	-2.47		AIC		
-Null hypothesis, Unit ro	ot.			<u>.</u>		
KPSS, Constant 0.955***						
KPSS, Constant, Trend	0.1	23*	1			
-Null hypothesis, Station	ary series	. *** (<	1% sign	ificance), * (<	< 10%	
significance),	-		•			

T-LL 2 H-.: (D--4 T-.4. 10(0 2010

European Journal of Statistics and Probability

Vol.10, No.2, pp., 1-13, 2022

Print ISSN: 2055-0154(Print),

Online ISSN 2055-0162(Online)

Table 3. Unit Root Tests, Full Sample							
Full Sample 1968-2021							
Age							
	Stand	ardize	LN				
Mean	9.2	24	2.21				
STD	1.:	58	0.17				
Observations	5	4					
	Lag	MZα	MZt	MSB	MPT		
NP, Constant	2	0.13	0.14	1.10	68.39		
NP, Constant, Trend	1 -2.12		-0.50	0.24	21.03		
-Null hypothesis, Unit roc	ot. AIC la	ig select	ion.				
ADF, Constant	-2.	-2.59 Lag 1		AIC			
ADF, Constant, Trend	-0.35		Lag 3	AIC			
-Null hypothesis, Unit roc	ot.						
KPSS, Constant 0.833***							
KPSS, Constant, Trend 0.148**							
-Null hypothesis, Stationa	ry series	. *** (<	1% sign	ificance), ** (< 5%		
significance).				·			

PANEL UNIT ROOT TEST

Single equation tests, as derived above, can suffer from a low power defect when examining shorter time spans. The number of time observations, T = 54, in the univariate series above, could be considered relatively small. A popular remedy for this problem is to use panel unit root tests that augment power by exploiting cross-sectional information. However, conventional panel unit root tests have been criticized, of late, for assuming that cross-section cointegrating relationships are not present (Westerlund and Breitung 2013). Assuming cross-section independence, when perhaps dependence is in play, tends to distort the size of the estimated test statistics that reject the null of non-stationarity too often.

Pesaran (2007) proposes a test statistic for cross-section dependence,

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \right),$$
(1)

where $\hat{\rho}_{ij}$ denotes the pair-wise correlation coefficient from the residuals of crosssectioned (*N*) Augmented Dickey-Fuller regressions. The *CD* statistic, testing the null of independence, is distributed asymptotically normal and possesses good small sample properties. Faced with the likelihood of cross-section dependence among mortality rates – we opt for Bai and Ng's (2004, 2010), Panel Analysis of Non-stationarity in Idiosyncratic and Common components (PANIC), method for panel unit root testing.⁷ The PANIC unit root test is based on a factor model in which non-stationarity can arise from common factors, idiosyncratic components, or both. Consider the following stochastic process for death rates,

$$D_{it} = c_i + \lambda_i' F_t + \eta_{it}, \qquad (2)$$

where the series D_{it} is the sum of a deterministic component c_i , a common component $\lambda'_i F_t$, and an error η_{it} that is idiosyncratic.⁸ Herein, factor selection follows the information criteria proposed by Bai and Ng (2002). Relative to the number of cross-sections (*N*) and time periods (*T*), the number of common factors are usually small. Multivariate common factors from equation (2) are tested using the modified version of the, more general, Qc test developed by Stock and Watson (1988). For each idiosyncratic component $\hat{\eta}_{it}$, the Augmented Dickey-Fuller test is applied to each cross-section. Accordingly, a pooled panel unit root statistic (distributed N(0,1)) for the idiosyncratic terms can be constructed,

$$P_{\hat{\eta}} = \frac{-2\sum_{i=1}^{N} \ln(p_{\hat{\eta}_i}) - 2N}{2\sqrt{N}},$$
(3)

where $p_{\hat{\eta}_i}$ denotes the probability values from the cross-sectioned Augmented Dickey-Fuller tests. It is important to note that tests on the common factors are asymptotically independent of tests on the idiosyncratic components. Lastly, a series with a factor structure is non-stationary (unit root) if one or more of the common factors are nonstationary, or the idiosyncratic error is non-stationary, or both. Data for this part of examination include state-level log age standardized all-cause mortality rates for all 50 states and the District of Columbia over the time period 1968-2021. Because the potential structural break had little effect in the univariate specification, we focus on the full sample for the panel analysis. Table 4 shows the data source, descriptive statistics and cross-section dependence test statistic.

@ECRTD-UK: <u>https://www.eajournals.org/</u>

⁷ This approach is arguably the workhorse in panel unit root testing, however, can suffer from small sample distortion particularly when the number of cross-sections is 'small'.

⁸ This factor model focuses on the intercept only specification where no deterministic trend is apparent in each cross-section (i).

Table 4. Source, Descriptive Statistics and Pesaran CD

All Cause Mortality Rate. Age Standardized rate per 1,000 total state population by jurisdiction of residence.					
Year 2000 population standard. National Center for Health Statistics, 1968-2021.					
	Full Sample				
	Rate	Rate LN			
Mean	9.28	2.21			
STD	1.69	1.69 0.18			
Observations (NT)		2,754			
Pesaran CD	2	254.04***			

***Significant at the < 1% level.

The last row of Table 4 shows the Pesaran (2007) test statistic for the full sample. The test rejects the null hypothesis of cross-section independence at any conventional significance level. Now we invoke PANIC – which does not require cross-section independence nor the stationarity of common components. Results are reported in Table 5.

	Lags	Stat		Lags	Stat
Alabama	8	-1.77*	Missouri	1	-1.86*
Alaska	1	-0.34	Montana	2	-1.19
Arizona	3	2.85	Nebraska	4	1.41
Arkansas	7	0.56	Nevada	4	-2.34**
California	0	-0.75	New Hampshire	1	-0.11
Colorado	2	1.48	New Jersey	1	-0.77
Connecticut	1	-0.99	New Mexico	2	0.37
DC	0	-0.63	New York	0	-0.51
Delaware	2	-0.02	North Carolina	3	1.76
Florida	4	0.23	North Dakota	9	2.18
Georgia	1	-1.33	Ohio	0	-1.23
Hawaii	7	-0.34	Oklahoma	10	-0.65
Idaho	4	0.21	Oregon	2	-0.48
Illinois	2	3.04	Pennsylvania	2	0.01
Indiana	1	-0.61	Rhode Island	10	-0.30
Iowa	1	0.91	South Carolina	4	-0.79
Kansas	1	-1.56	South Dakota	1	-2.41**
Kentucky	2	-0.87	Tennessee	10	1.76
Louisiana	1	-0.62	Texas	9	-0.77
Maine	3	2.15	Utah	0	-0.14
Maryland	2	-1.18	Vermont	6	-0.32
Massachusetts	0	0.26	Virginia	2	0.64

Table 5. PANIC Results

@ECRTD-UK: <u>https://www.eajournals.org/</u>

European Journal of Statistics and Probability

Vol.10, No.2, pp., 1-13, 2022

Print ISSN: 2055-0154(Print),

Online ISSN 2055-0162(Online)

Michigan	2	-0.13	Washington	6	2.15
Minnesota	1	-0.26	West Virginia	0	-1.02
Mississippi	1	1.05	Wisconsin	2	-0.55
			Wyoming	6	-2.67***
Null Rejections		5			
	7	11.24			
Common Factors	/	11.34			
Idiosyncratic		-0.58			

Significance at 1% (***), 5% (**), 10% (*) levels.

Only 5 state cross-sections reject the null hypothesis of unit root at conventional levels. Multiple (7) common factors are determined and then tested with an iterative procedure. Herein, we apply the more general MQc test which corrects for serial correlation, of arbitrary form, through non-parametric estimation. MQc parallels the multivariate procedure suggested by Phillips (1987). The null hypothesis states that r common factors have at most r common stochastic trends. As in our case, failure to reject the null of retaining the common factors indicates that all are non-stationary. The last row of Table 5 shows the pooled idiosyncratic component test (see equation (3)). The null hypothesis of this test is all cross-sections have a unit root (non-stationary). Note that the null holds only if no stationary combination of the D_i exists. As such, the pooled test mirrors a panel test for no cointegration. Herein, we fail to reject the null of no cointegration among states. Overall results are consistent with non-stationarity in U.S. age standardized mortality rates, pervasive in both the common factors and in the idiosyncratic components.

CONCLUDING REMARKS

In the aggregate, we find strong evidence of non-stationarity in age standardized all cause mortality rates for the United States. Moreover, after controlling for age and population, mortality in the U.S. has declined markedly since the late 1960s. While key causes of death may shuffle through the decades, sum-total rate of death has declined. What seems puzzling is the proliferation of the use of the phrase 'excess death'. Excess, in this modifying context, means over and above what is usual or ordinary.⁹ This presumes that there is a usual or ordinary rate of death. Herein, we show that U.S. death rate timeseries are non-stationarity encompasses cyclical fluctuations and shocks into permanent effects on the time-series. This is akin to the scientific concept of *hysteresis* which favors path-dependence and the inability of events to return to an initial level after being changed by an external force – even after the force is removed. Using the idiom 'excess death' to define observed death above forecasts smacks as superfluous, intending to alarm. Why not simply call it what it is, 'death above projection'?

⁹ Merriam - Webster, excess used as an adjective.

REFERENCES

- Achenbach, J. (2022) U.S. excess deaths surpass 1 million. The Washington Post. February 15.
- Ahmad, F., Cisewski, J. and Anderson, R. (2022) Provisional mortality data U.S. 2021. Morbidity and Mortality Weekly Report 71(17) pp. 597-600.
- Ayala, C. (2000) Age-specific excess deaths associated with stroke among racial/ethnic groups U.S. 1997. Morbidity and Mortality Weekly Report. CDC. February 11.
- Bai, J. and Ng, S. (2002) Determining the number of factors in approximate factor models. Econometrica 70(1), pp. 191-221.
- Bai, J. and Ng, S. (2004) A PANIC attack on unit roots and cointegration. Econometrica 72(4), pp. 1127-1177.
- Bai, J. and Ng, S. (2010) Panel unit root tests with cross-section dependence: A further investigation. Econometric Theory 26, pp. 1088–1114.
- Bhargava, A. (1986) On the theory of testing for unit roots in observed time series. Review of Economic Studies 53(3), pp. 369-384.
- CDC (1961) Morbidity and Mortality Weekly Report 10(1). Public Health Service. Department of Health, Education and Welfare.
- CDC (2018) https://www.cdc.gov/about/history.
- CDC (2021) Impact of hospital strain on excess deaths during the COVID-19 pandemic - U.S. July 2020 - July 2021. Morbidity and Mortality Weekly Report 70(46), pp. 1613-1616.
- CDC (2022) https://www.cdc.gov/nchs/vsrr/excess deaths.htm
- Department of Health, Education and Welfare (1967) 1965 Vital statistics of the U.S. Volume 2, Mortality. Public Health Service. U.S. Government Printing Office.
- Department of Health, Education and Welfare (1976) History, Mission and Organization of the Public Health Service. U.S. Government Printing Office.
- Dickey, D. and Fuller, W. (1979) Distribution of estimators for autoregressive time series with unit root. Journal of American Statistical Association 74, pp. 427-431.
- Dickey, D. and Fuller, W. (1981) Likelihood ratio tests for autoregressive time series with a unit root. Econometrica 49, pp. 1057-1072.
- Elliot, G., Rothenberg, T. and Stock, J. (1996) Efficient test for an autoregressive unit root. Econometrica 64, pp. 813-36.
- Flegal, K., Graubard, B., Williamson, D. and Gail, M. (2018) Excess deaths associated with underweight, overweight and obesity: An evaluation of a potential bias. Vital and Health Statistics 3(42). NCHS.
- Grillenzoni, C. (1998) Forecasting unstable and nonstationary time series. International Journal of Forecasting 14, pp. 469-482.
- Harari, Y. (2014) Sapiens: A brief history of humankind. Harvill Secker Publishing.
- Kwiatkowski, D., Phillips, P., Schmidt, P. and Shin, Y. (1992) Testing the null hypothesis of stationarity against the alternative of unit root. Journal of Econometrics 54, pp. 159-178.

@ECRTD-UK: <u>https://www.eajournals.org/</u>

European Journal of Statistics and Probability Vol.10, No.2, pp., 1-13, 2022 Print ISSN: 2055-0154(Print),

Online ISSN 2055-0162(Online)

- Ng S. and Perron, P. (2001) Lag length selection and the construction of unit root tests with good size and power. Econometrica 69, pp. 1519-1554.
- Perron P. (1989) The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57(6), pp. 1361-1401.
- Pesaran, M. (2007) A simple panel unit root test in the presence of cross-section dependence. Journal of Applied Econometrics 22, pp. 265-312.
- Phillips, P. (1987) Times series regression with a unit root. Econometrica 55, pp. 277-301.
- Phillips, P. and Perron, P. (1988) Testing for a unit root in time series regression. Biometrika 75, pp. 335-346.
- Stock, J. and Watson, M. (1988) Testing for common trends. Journal of the American Statistical Association 83, pp. 1097-1107.
- Wang, Yanding, Yan, Z., Wang, D., Yang, M., Li, Z., Gong, X., Wu, D., Zhai, L., Zhang, W. and Wang, Yong. (2022) Prediction and analysis of COVID-19 daily new cases and cumulative cases: time series forecasting and machine learning tools. BioMed Central BMC Infectious Diseases 22, Article 495.
- Westerlund, J. and Breitung, J. (2013) Lessons from a decade of IPS and LLC. Econometric Review 32, pp. 547-591.

APPENDIX

I able A1. Descriptive Statistics by State						
	Age					
	Standardized Mean	In Mean	STD in Mean			
Alahama			STD In Mean			
Alabama	10.484	2.344	0.105			
Alaska	9.169	2.202	0.169			
Arizona	8.602	2.138	0.166			
Arkansas	9.994	2.298	0.093			
California	8.513	2.122	0.202			
Colorado	8.467	2.122	0.170			
Connecticut	8.443	2.116	0.187			
DC	11.054	2.379	0.225			
Delaware	9.601	2.246	0.181			
Florida	8.552	2.132	0.169			
Georgia	10.083	2.299	0.152			
Hawaii	7.313	1.975	0.172			
Idaho	8.700	2.153	0.145			
Illinois	9.459	2.230	0.187			
Indiana	9.752	2.268	0.134			
Iowa	8.576	2.138	0.145			
Kansas	8.862	2.174	0.122			
Kentucky	10.411	2.337	0.112			
Louisiana	10.590	2.352	0.127			
Maine	9.247	2.210	0.166			
Maryland	9.457	2.229	0.188			
Massachusetts	8.736	2.149	0.191			
Michigan	9.599	2.250	0.155			
Minnesota	8.155	2.084	0.169			
Mississippi	10.824	2.376	0.110			
Missouri	9.694	2.262	0.134			
Montana	9.070	2.193	0.153			
Nebraska	8.610	2.143	0.140			
Nevada	9.795	2.268	0.167			
New Hampshire	8.885	2.167	0.186			
New Jersey	9.117	2.189	0.205			
New Mexico	8.963	2.182	0.146			
New York	9.059	2.178	0.229			
North Carolina	9.728	2.263	0.155			
North Dakota	8.371	2.113	0.152			

Table A1. Descriptive Statistics by State

European Journal of Statistics and Probability

Vol.10, No.2, pp., 1-13, 2022

Print ISSN: 2055-0154(Print),

Online ISSN 2055-0162(Online)

			
Ohio	9.816	2.274	0.141
Oklahoma	9.953	2.294	0.083
Oregon	8.771	2.161	0.148
Pennsylvania	9.615	2.248	0.173
Rhode Island	8.846	2.165	0.171
South Carolina	10.208	2.311	0.152
South Dakota	8.628	2.144	0.148
Tennessee	10.191	2.315	0.110
Texas	9.265	2.215	0.148
Utah	8.398	2.117	0.146
Vermont	8.926	2.172	0.185
Virginia	9.390	2.223	0.180
Washington	8.608	2.138	0.172
West Virginia	10.615	2.356	0.113
Wisconsin	8.729	2.155	0.154
Wyoming	9.181	2.204	0.161