

# **The Impact of Socioeconomic Factors on State Suicide Rates: Revisited**

**Mitch Kunce<sup>1</sup>**

## **Abstract**

This paper revisits the purported impact of socioeconomic and social environment factors on annual, U.S. state-level suicide rates. Special attention is paid to the right-hand-side linking covariates directly to Durkheim's (1897/1951) significant contributions to established ecological suicide research. Results from a Hausman-Taylor panel specification lend little support to Durkheim's social integration/regulation hypothesis that aggregate social forces matter in explaining variations in regional suicide rates. Data from 1990-2019 and the advanced empirical method support the mounting sentiment of an abiding ecological fallacy plaguing suicidology.

**JEL classification numbers:** C51, R11, I31.

**Keywords:** Hausman-Taylor, Suicide rates, Socioeconomic factors.

---

<sup>1</sup> Douglas Mitchell Econometric Consulting Laramie, WY USA.

## 1. Introduction

The sociological study of suicide has a rich and long history rooted in Durkheim's (1897/1951) *Suicide*. Durkheim posits that suicide rates of a region are determined by the degree of both social integration and social (moral) regulation in a society. Suicide will be common when either of these social factors is too weak or too strong. In a related vein, rational choice theory (Becker 1962) influencing lifetime economic utility models (Hamermesh and Soss 1974) tend to bolster the socioeconomic premise explaining spatial differences in suicide. However, a mounting literature (Breault 1995, Cutchin and Churchill 1999, Kposowa 2000, Cubbin et al 2000, Kunce and Anderson 2002, Kushner and Sterk 2005, Abrutyn and Mueller 2014, Mueller et al 2021) casts doubt on the ecological socioeconomic hypothesis.

"One of the oldest and most notable limitations of Durkheim is methodological.

Durkheim fails to adequately address the ecological fallacy of studying suicide rates to understand individual behavior. Durkheim forcefully argued that societal- or macro-level social forces (integration and regulation) caused individual-level behavior (suicide), and yet the link between societal-level social forces and individual behavior is challenging yet crucial to document."

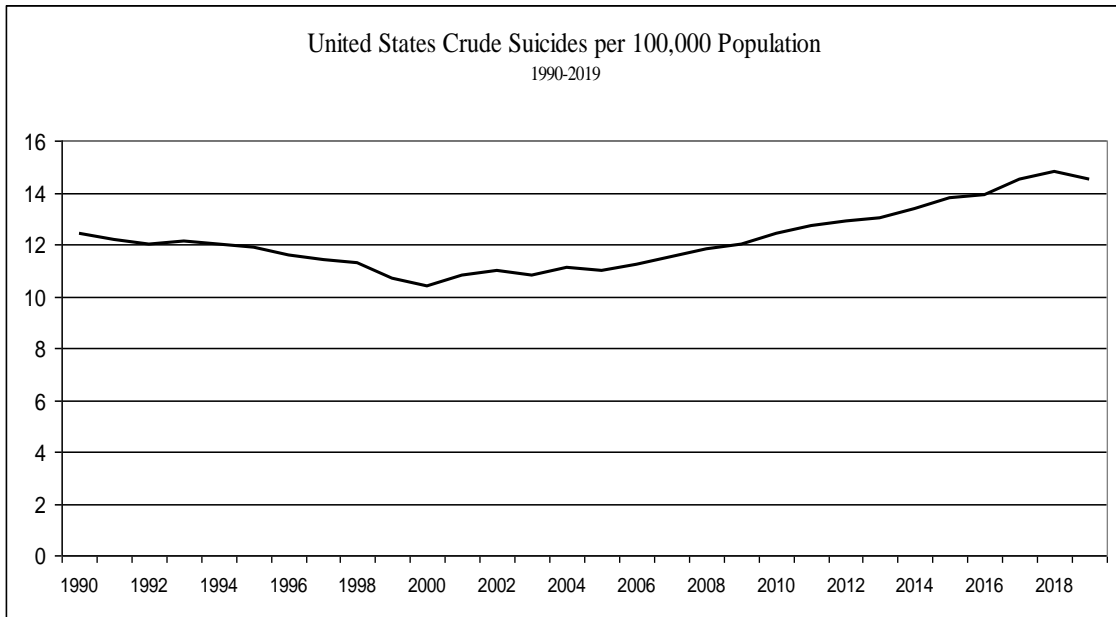
(Mueller et al 2021, p. 3)

Two noteworthy empirical exceptions to this rising sentiment include Neumayer (2003) and Yamamura (2010). Neumayer (2003) examines a select sample of 62 countries, where complete data are available, finding strong support for the social causes proposition. Yamamura (2010) analyzes gender stratified data from Japanese prefectures lending somewhat mixed evidence. Both are critical of the selection of right-hand-side variables in prior empirical constructs. They argue that econometric specifications should adhere strictly to Durkheim's seminal analysis and tie directly to well established ecological suicide research.

Herein, we revisit the impact of socioeconomic factors on state suicide rates by exploiting a rich and expansive balanced panel of time-series, cross-section data from 1990 to 2019. The selection of covariates closely follows Durkheim's foundational insight as suggested by Neumayer (2003). Results from a two-way Hausman and Taylor (1981) specification finds little evidence supporting the social causes hypothesis that aggregate socioeconomic factors explain the variation in state suicide rates. The balance of this examination is divided into four sections. Section 2 describes the data, provides sources and presents descriptive statistics. Additional attention is paid to the dependent variable, state suicide rates. Section 3 presents the empirical model and discusses the econometric issues. Section 4 interprets the empirical inference with conclusions and implications drawn in section 5.

## 2. Data

Figure 1 shows the rise in crude suicide rates over the last 20 years in the U.S. This trend is considered a public health emergency in need of fresh solutions (Mueller et al 2021). Contemporary solutions appear to be moving away from suicidology and more toward prevention science. Prevention, at present, addresses more distal risk factors requiring complex interventions that take a public health and population-based approach rather than addressing the issues of high-risk populations.



**Figure 1: Crude suicides per 100,000 population, 1990-2019**

Means, over the 30 year panel, of crude suicide rates exhibit considerable variation across states. Highest suicide rates are found in Montana 21.6, Wyoming 21.4, Nevada 21.1 and Alaska 20.8, whereas the lowest rates appear in DC 6.2, New Jersey 7.4, New York 7.7 and Massachusetts 8.2. Table 1 describes, provides data sources and shows descriptive statistics for all variables examined. The choice of regressors closely follows Neumayer's (2003) discussion of the nuance and richness of Durkheim's analysis. Moreover, the right-hand-side is consistent with well established specifications in ecological suicide research (Yamamura 2010). Mindful of the multicollinearity issues inherent in state level aggregate data, attention is paid to the orthogonality of covariates entered in  $X_{it}$ . A correlation matrix for the right-hand-side is provided in Table 2. Moreover, variance inflation factors (VIF) are estimated for each regressor and shown in the far-right column of Table 2. Multicollinearity, as a matter of degree, does not appear problematic with this suite of regressors.

**Table 1: Data descriptions, sources and descriptive statistics**

|  |
|--|
| <p><b>Suicide Rate.</b> Age Standardized per 100,000 total state population. National Center for Health Statistics, National Vital Statistics System, 1990-2019.<br/>Mean 13.63, STD 4.03.<br/>Natural Log Suicide Rate, Mean 2.57, STD 0.30.</p>  |
| <p><b>Gross State Product Per Capita.</b> In millions of 2012 dollars per 1,000 total state population. U.S. Department of Commerce, Bureau of Economic Analysis, Regional Economic Analysis Division, 1990-2019.<br/>Mean 49.02, STD 19.33.</p>   |
| <p><b>Birth Rate.</b> Live births occurring in a state per 1,000 total state population. National Center for Health Statistics, National Vital Statistics System, 1990-2019.<br/>Mean 13.50, STD 1.96.</p>   |
| <p><b>Female Labor Force Participation.</b> In percent of total female working age population by state. U.S. Bureau of Labor Statistics, 1990-2019.<br/>Mean 60.11, STD 4.37.</p>  |
| <p><b>Distilled Spirits Consumption.</b> In gallons per capita, total state population. Beer Institute, Beer Almanac, 1990-2019.<br/>Mean 1.53, STD 0.56.</p>  |
| <p><b>Divorce Rate.</b> Based on counts of divorces occurring in a state per 1,000 total state population. National Center for Health Statistics (NCHS), National Vital Statistics System, 1990-2019. California, Georgia, Hawaii, Indiana, Louisiana and Minnesota have stopped consistently reporting divorce occurrence to the NCHS. This missing data can be recovered from the U.S. Bureau of the Census, American Community Survey (ACS), Public Use Microdata Sample (PUMS). See Mayol-Garcia et al (2021) for a similar application of ACS PUMS data.<br/>Mean 3.96, STD 1.13.</p> |
| <p><b>Marriage Rate.</b> Based on resident marriages occurring in a state per 1,000 total state population. ACS PUMS 1990-2019 resident data was used in order to correct for the large number (outliers) of out of state residents marrying in Las Vegas, Nevada (see U.S. Bureau of the Census 2021).<br/>Mean 8.01, STD 2.42.</p>   |
| <p><b>Household Size.</b> Average size of households by state. U.S. Bureau of the Census, 1990-2019.<br/>Mean 2.53, STD 0.16.</p>  |
| <p><b>Unemployment Rate.</b> Seasonally adjusted average annual rates by state, in percent. U.S. Bureau of Labor Statistics, 1990-2019.<br/>Mean 5.48, STD 1.87.</p>   |

**Table 2: Correlation matrix and Variance Inflation Factors**

|                     | <b>SGP/Capita</b> | <b>Birth</b> | <b>Female</b> | <b>Spirits</b> | <b>Divorce</b> | <b>Marriage</b> | <b>Household</b> | <b>VIF</b> |
|---------------------|-------------------|--------------|---------------|----------------|----------------|-----------------|------------------|------------|
| <b>GSP/Capita</b>   | 1.00              |              |               |                |                |                 |                  | 1.78       |
| <b>Birth</b>        | -0.04             | 1.00         |               |                |                |                 |                  | 2.04       |
| <b>Female</b>       | 0.25              | 0.09         | 1.00          |                |                |                 |                  | 1.47       |
| <b>Spirits</b>      | 0.55              | -0.28        | 0.33          | 1.00           |                |                 |                  | 1.85       |
| <b>Divorce</b>      | -0.35             | 0.35         | -0.18         | -0.17          | 1.00           |                 |                  | 2.00       |
| <b>Marriage</b>     | -0.27             | 0.33         | -0.02         | -0.22          | 0.59           | 1.00            |                  | 1.75       |
| <b>Household</b>    | -0.22             | 0.58         | -0.10         | -0.31          | 0.21           | 0.35            | 1.00             | 1.74       |
| <b>Unemployment</b> | 0.06              | 0.01         | -0.35         | 0.01           | 0.07           | -0.11           | 0.03             | 1.20       |

The dependent variable examined is a state's annual age-standardized suicide rate (per 100,000 population). Age-standardized (adjusted) rates ( $R'$ ) are commonly used in the ecological suicide literature to compare relative indexes across groups and over time. The National Center for Health Statistics (NCHS) computes the standardized rates by weighting age-specific suicide rates ( $R_j$ ) as follows,

$$R' = \sum_j \frac{P_{sj}}{P_s} R_j \quad (1)$$

where  $P_{sj}$  is the standard state population for age group  $j$  and  $P_s$  is the total state standard population (all ages combined).

All age-standardized rates examined herein are based on the 2000 U.S. standard population statistics (National Vital Statistics Reports 2017). A referee points out that pinning a time sensitive behavioral standard on 20 year old population weights seems nebulous. In response see Appendix A for a comparison of 2019 crude and age-standardized rates by state. Additionally, Appendix A details in tables the computation of the 2019 age-standardized rate for the U.S. and shows a comparison of the 2000 NCHS weights versus 2019 population weights. The referee's concern appears appropriate, the population age distribution has changed markedly over the last two decades. Regardless, the NCHS has adopted the year 2000 population standard.

Although, there is nothing in socioeconomic suicide theory that dictates functional form, previous examinations have used linear and semilog linear specifications (see Lester 2020 for a review). On a priori grounds, the semilog functional form has considerable appeal relative to the linear form. By taking the natural logarithm of suicide rates, estimates of right-hand-side coefficients vary proportionately with suicide rates rather than effect the overall level as in the linear construct. Moreover, semilog form can lessen the effects of vertical outliers on coefficient estimates (see the application in Verardi and Wagner 2011). This monotonic transformation of the left-hand-side also corrects for a common form of error variance heteroscedasticity in which the standard error is correlated with the conditional expectation of the

dependent variable (Glejser 1969). Consequently, the semilog form was selected for this examination.

### 3. Econometrics

The model estimated becomes,

$$\ln Y_{it} = \alpha + X_{it}\beta + \mu_i + \lambda_t + \eta_{it} \quad (2)$$

where  $\alpha$  is a scalar intercept,  $X_{it}$  are observable socioeconomic variables that vary across states  $i$  and over time  $t$ ,  $\beta$  is a vector of estimated coefficients,  $\mu_i$  and  $\lambda_t$  are latent state and time specific effects, and  $\eta_{it}$  denotes the remainder disturbance. Generally, two specifications of equation (2) are considered. Fixed effects treats  $\mu_i$  and  $\lambda_t$  as fixed yet unknown constants differing across states and over time. Alternatively, random effects assumes that  $\mu_i$  and  $\lambda_t$  are random, distributed independently across states and over time. The potential correlation of  $\mu_i$  and  $\lambda_t$  with the variables in  $X_{it}$  is a primary consideration. If these correlations are present, random effects estimation yields biased and inconsistent estimates of  $\beta$  and the variances of  $\mu_i$ ,  $\lambda_t$  and  $\eta_{it}$ . By transforming the data, into deviations from the group means, the fixed effects estimator is not impacted by this lack of orthogonality but is not fully efficient since it ignores variation across states and over time. The choice of estimator generally rests on statistical considerations and hypothesis testing. Hausman (1978) outlines a specification test of the null hypothesis of orthogonality between the latent effects and  $X_{it}$ .

Hausman and Taylor (1981) (HT) propose a third specification for estimating equation (2) where  $X_{it}$  is split into two sets of variables, those assumed uncorrelated with  $\mu_i$ ,  $\lambda_t$  and  $\eta_{it}$ , and those correlated with  $\mu_i$  and  $\lambda_t$ , but not  $\eta_{it}$ . The uncorrelated set serves two functions, (i) using mean deviations, unbiased estimates of the respective  $\beta$  are produced, and (ii) the set provides valid instruments for the variables in  $X_{it}$  correlated with the latent effects. The two sets of variables need not rely solely on a priori assumptions, correlation hypothesis can be tested.

### 4. Inference

The question of poolability naturally arises with panel data. The restriction is the pooled model representing a behavioral question with the same slope parameters across states and over time. These structural tests share roots with Chow (1960) and test the null hypothesis of equal slope coefficients. Table 3 shows F-tests for pooling across states and over time. Though not statistically imperious, both fail to reject the null at the < 5% level.

**Table 3: Hypotheses tests**

|                                       |                                |
|---------------------------------------|--------------------------------|
| Pooling slopes across states          | F(400,1071) = 1.14 (p 0.054)   |
| Pooling slopes over time              | F(232,1260) = 1.16 (p 0.064)   |
| LM Random Effects                     | $\chi^2_2 = 8083.81$ (p 0.000) |
| One-way State Effects vs. Pooled OLS  | F(50,1471) = 170.91 (p 0.000)  |
| Adding Year Effects vs. One-way model | F(29,1442) = 24.04 (p 0.000)   |
| Pooled Durbin-Watson                  | 2.01                           |
| Variance Inflation Factor Threshold   | 1.35                           |

Careful testing denoted in Table 3 confirms state and year heterogeneity and verifies the importance of controlling for unobservable state and year effects. The Lagrange multiplier test statistic of 8,083.81 distributed  $\chi^2$  with 2 degrees of freedom soundly rejects the null hypothesis,

$$H_0 : \sigma_\mu^2 = \sigma_\lambda^2 = 0.$$

Second, the test statistic  $F(50, 1471) = 170.91$  firmly rejects the null hypothesis of state homogeneity at the  $< 1\%$  level. The  $F(29, 1442) = 24.04$  statistic favors the two-way model over the one-way specification, again at the  $< 1\%$  level. Lastly, the Durbin-Watson statistic (2.01) indicates that serial correlation poses no direct problem in the pooled sample.

Results from the two-way error components estimators are presented in Table 4. The random effects estimates in the first column are easily challenged due to the sizable Hausman test statistic,  $\chi^2_8 = 91.53$ . The null,  $H_0 : E(\mu_t, \lambda_t | X_{it}) = 0$ , of latent effects orthogonality is soundly rejected at the  $< 1\%$  level. The random effects estimator tests biased and inconsistent. The unbiased yet less efficient two-way fixed effects estimates offer little support to the social forces hypothesis in explaining the variation in state suicide rates. Only distilled spirits consumption and average household size test significant at any conventional p level. The coefficient of average household size (-8.48) implies that, for the average state, the suicide rate per 100,000 population will decrease proportionately by around 8.5% for a marginal increase in average household size.

**Table 4: Two-way error components estimates<sup>a</sup>**

| Variable  | Random Effects     | Fixed Effects   | Hausman-Taylor <sup>b</sup> |
|---|--------------------|-----------------|-----------------------------|
| Constant (t)  | 2.52 (15.21)***    | 2.64 (15.73)*** | 2.49 (15.63)***             |
| GSP/Capita (t)  | -0.09 (-1.68)*     | -0.06 (-1.09)   | -0.04 (-0.61)               |
| Birth Rate (t)  | 0.75 (2.24)**      | 0.43 (1.23)     | 0.35 (1.02)                 |
| Female Labor (t)  | -0.17 (-1.01)      | -0.14 (-0.83)   | 0.16 (0.95)                 |
| Spirits (t)   | 8.19 (4.33)***     | 9.85 (4.81)***  | 10.04 (5.07)***             |
| Divorce Rate (t)  | 1.05 (1.79)*       | 0.02 (0.04)     | 0.12 (0.20)                 |
| Marriage Rate (t)   | -0.12 (-0.47)      | -0.32 (-1.20)   | -0.40 (-1.48)               |
| Household Size (t)  | -3.92 (-0.85)      | -8.48 (-1.75)*  | -8.72 (-1.80)*              |
| Unemployment (t)  | 0.08 (0.30)        | 0.25 (0.89)     | 0.24 (0.86)                 |
| R <sup>2</sup>  | c                  | 0.93            | c                           |
| Hausman Test  | $\chi_8^2 = 91.53$ |                 | $\chi_{34}^2 = 8.05$        |
| <sup>a</sup> Non-intercept variables in percentage terms.<br><sup>b</sup> Year dummies included in set X <sub>1</sub> representative of a two-way model.<br><sup>c</sup> No precise counterpart to R <sup>2</sup> in these specifications.<br>***, **, * significance at the 1%, 5%, 10% level.<br>Observations 1,530 |                    |                 |                             |

A marginal increase in distilled spirits consumption per capita increases the suicide rate for the average state by roughly 10%. Interestingly, Durkheim (1897/1951) did not support the presumption that alcohol consumption could explain regional differences in suicide. Durkheim viewed alcoholism as a psychopathic state rather than a symptom of the level of integration. Conversely, strong evidence of an alcohol consumption - suicide link is supported herein. The inconspicuous explanatory power of the fixed effects specification ( $R^2 = 0.93$ ) rests on the latent effects. A note of interpretive caution, fixed effects estimation places great demands on the data. For example,  $\mu_i$  capture any between state variation leaving only within state variation to be picked up by regressors.

If we are inclined to assume that certain variables in  $X_{it}$  are uncorrelated with the latent effects, Hausman and Taylor (1981) outlines an estimator that may produce consistent and efficient estimates of  $\beta$ . In order to identify variable sets, iterative two-way random effects regressions were performed - varying the sets of variables. Table 5 shows the resulting Hausman specification tests. For example, the first row depicts the resulting test statistic when the Gross State Product variable is dropped from the right-hand-side. Note that the Hausman test statistic reduces to 60.22 from 91.53. The Gross State Product variable appears to be a significant 'correlation contributor'. Following this logic, set  $X_1$  (uncorrelated) includes Birth Rate, Spirits Consumption, Marriage Rate, Household Size and Unemployment Rate with set  $X_2$  (correlated) containing Gross State Product, Female Labor Force Participation and Divorce Rates.



**Table 5: Latent effect correlation tests\***

|                                       | $\chi^2_7$ |
|---------------------------------------|------------|
| <b>GSP/Capita</b>                     | 60.22      |
| <b>Birth</b>                          | 89.63      |
| <b>Female</b>                         | 77.43      |
| <b>Spirits</b>                        | 88.04      |
| <b>Divorce</b>                        | 30.23      |
| <b>Marriage</b>                       | 88.40      |
| <b>Household</b>                      | 88.81      |
| <b>Unemployment</b>                   | 89.62      |
| *All RHS variables $\chi^2_8 = 91.53$ |            |

With the variable sets identified, *LIMDEP*<sup>®</sup> Version 11 provides a Hausman-Taylor estimator for the one-way random effects model. In order to estimate a comparable two-way specification,  $T - 1$  time dummies are included in variable set  $X_I$ . Results are presented in the last column of Table 4. Again, only distilled spirits consumption and average household size test significant at any conventional p level. Proportional marginal effect coefficients are also comparable to the fixed effects specification. Interestingly, the significant effect of average household size found herein is fairly analogous to the effect of single occupant households found in Kuncce and Anderson (2002).

A Hausman test based on the difference between the HT and FE estimator yields,  $\chi^2_{34} = 8.05$ , which fails to reject the null hypothesis of orthogonality. The HT use of within model instruments appears appropriate and estimates test as unbiased, consistent *and* asymptotically efficient.

## 5. Conclusion

Despite a wide reaching literature, divergent findings regarding the impact of socioeconomic factors on suicide rates persist. Twenty years ago, Kuncce and Anderson (2002) argued that the reasons for the contrariety rests with data aggregation and estimation methodology. In a systematic review of the 'at the time' relevant empirical literature, Rehkopf and Buka (2006) examined 86 publications which accounted for 221 separate analyses. All 221 analyses used aggregated data varying from cities (48 in fact looked at sub-city data) to select countries. Justification of this now conventional approach is that the data are mainly available in the aggregate form. There is a vast literature critical of the use of aggregate data to explain heterogeneous individual behavior (see Stroker 1993 and Holderness 2016 for reviews). Statistical properties and the biases introduced by using aggregated or averaged data have yet to be adequately explained. Additionally, Rehkopf and Buka (2006) found that 87% (192) of the analyses used solely cross-sectional data. Cross-sectional estimation methods are deficient in controlling for grouped or individual-specific effects -

possibly unobservable - which may be correlated with other covariates in the specification of a socioeconomic relationship (Kunce and Anderson 2001-2002).

As noted above, the spirit of Durkheim's theory centers on two simple tenets:

- i. the structure of social relationships shapes the structure of suicide and
  - ii. integration and regulation are key structural dimensions of social relationships.
- At the core, societies that foster social relationships and cultivate a strong collective conscience are healthier, and low suicide occurrence is but one piece of the empirical evidence (Mueller et al. 2021). However, a consensus in the literature regarding this one piece of empirical evidence does not exist. Past ecological examinations of these tenets tend to substitute biased aggregate social profiles for individual behavior. When empirical studies are rigorously framed, constructs that involve more than just 'jumbling' (Hood-Williams 1996) the right-hand-side, results remain mixed at best. Figure 1 above shows that suicide rates in the United States are at their highest levels in over three decades. Perhaps David Lester is right, suicidology may indeed be dead (Connolly and Lester 2021).

## References

- [1] Durkheim, E. (1897/1951). *Suicide: A study in sociology*. Glencoe, IL: Free Press.
- [2] Becker, G. (1962). Irrational behavior and economic theory. *Journal of Political Economy* 70, pp. 1-13.
- [3] Hamermesh, D. and Soss, N. (1974). An economic theory of suicide. *Journal of Political Economy* 82, pp. 83-98.
- [4] Breault, K. (1995). Was Durkheim right? A critical survey of the empirical literature on *Le Suicide*. In: D. Lester (Ed.) *Celebrating the Centennial of Durkheim's Le Suicide*. pp. 35-62. Philadelphia, PA: Charles Press.
- [5] Cutchin, M. and Churchill, R. (1999). Scale, context, and causes of suicide in the United States. *Social Sciences Quarterly* 80, pp. 97-114.
- [6] Kposowa, A. (2000). Marital status and suicide in the National Longitudinal Mortality Study. *Journal of Epidemiology and Community Health* 54, pp. 254-261.
- [7] Cubbin, C., LeClere, F. and Smith, G. (2000). Socio-economic status and the occurrence of fatal and nonfatal injury in United States. *American Journal of Public Health* 90, pp. 70-77.
- [8] Kunce, M. and Anderson, A. (2002). The impact of socioeconomic factors on state suicide rates: A methodological note. *Urban Studies* 39(1), pp. 155-162.
- [9] Kushner, H. and Sterk, C. (2005). The limits of social capital: Durkheim, suicide, and social cohesion. *American Journal of Public Health* 95, pp. 1139-1143.
- [10] Abrutyn, S. and Mueller, A. (2014). The socioemotional foundations of suicide: A microsociological view of Durkheim's *Suicide*. *Sociological Theory* 32(4), pp. 327-351.

- [11] Mueller, A., Abrutyn, S., Pescosolido, B. and Diefendorf, S. (2021). The social roots of suicide: Theorizing how the external social world matters to suicide and suicide prevention. *Frontiers in Psychology* 12: Paper # 621569.
- [12] Neumayer, E. (2003). Socioeconomic factors and suicide rates at large unit aggregate levels: A comment. *Urban Studies* 40(13), pp. 2769-2776.
- [13] Yamamura, E. (2010). The different impacts of socio-economic factors on suicide between males and females. *Applied Economics Letters* 17(10), pp. 1009-1012.
- [14] Hausman, J. and Taylor, W. (1981). Panel data and unobservable individual effects. *Econometrica* 49(6), pp. 1377-1398.
- [15] Mayol-Garcia, Y., Gurrentz, B. and Kreider, R. (2021). Number, timing, and duration of marriages and divorces. *Current Population Report*. U.S. Census Bureau.
- [16] U.S. Census Bureau. (2021). Understanding and using American Community Survey (ACS) Public Use Microdata Sample (PUMS) Files: What Data Users Need to Know. U.S. Government Printing Office, Washington, DC.
- [17] National Vital Statistics Reports. (2017). Volume 66, Number 6.
- [18] Lester, D. (2020). *Suicide Studies* 1(1), pp. 19-36. <http://drdavidlester.net>.
- [19] Verardi, V. and Wagner, J. (2011) Robust estimation of liner fixed effects panel data models with an application to the exporter productivity premium. *Journal of Economics and Statistics* 231(4), pp. 546-557.
- [20] Glejser, H. (1969). A new test for heteroscedasticity. *Journal of the American Statistical Association* (644), pp. 316-323.
- [21] Hausman, J. (1978). Specification tests in econometrics. *Econometrica* 46(6), pp. 1251-1271.
- [22] Chow, G. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28, pp. 591-605.
- [23] Rehkopf, D. and Buka, S. (2006). The association between suicide and the socio-economic characteristics of geographical areas: A systematic review. *Psychological Medicine* 36, pp. 145-157.
- [24] Stoker, T. (1993). Empirical approaches to the problem of aggregation over individuals. *Journal of Economic Literature* 31, pp. 1827-1874.
- [25] Holderness, C. (2016). Problems using aggregate data to infer individual behavior: Evidence from law, finance, and ownership concentration. *Critical Finance Review* 5(1), pp. 1-40.
- [26] Kunce, M. and Anderson, A. (2001-2002). A natural rate of suicide for the U.S., revisited. *Omega, Journal of Death and Dying* 44(3), pp. 215-222.
- [27] Hood-Williams, J. (1996). Studying suicide. *Health and Place* 2, pp. 167-177.
- [28] Connolly, J. and Lester, D. (2021). Interviews with suicidologists. 4, p. 17. <http://drdavidlester.net>.

## Appendix A

**Table A1: 2019 crude and age-standardized suicide rates per 100,000 population**

|                | <b>Suicides</b> | <b>Population</b> | <b>Crude Rate</b> | <b>Age-Standardized Rate</b> |
|----------------|-----------------|-------------------|-------------------|------------------------------|
| Alabama        | 804             | 4,903,185         | 16.4              | 16.3                         |
| Alaska         | 210             | 731,545           | 28.7              | 28.5                         |
| Arizona        | 1,419           | 7,278,717         | 19.5              | 18.7                         |
| Arkansas       | 548             | 3,017,804         | 18.2              | 18.0                         |
| California     | 4,436           | 39,512,223        | 11.2              | 10.7                         |
| Colorado       | 1,312           | 5,758,736         | 22.8              | 22.1                         |
| Connecticut    | 435             | 3,565,287         | 12.2              | 11.4                         |
| Delaware       | 111             | 973,764           | 11.4              | 11.3                         |
| DC             | 44              | 705,749           | 6.2               | 6.2                          |
| Florida        | 3,465           | 21,477,737        | 16.1              | 14.5                         |
| Georgia        | 1,585           | 10,617,423        | 14.9              | 14.6                         |
| Hawaii         | 224             | 1,415,872         | 15.8              | 15.5                         |
| Idaho          | 365             | 1,787,065         | 20.4              | 20.4                         |
| Illinois       | 1,439           | 12,671,821        | 11.4              | 10.9                         |
| Indiana        | 972             | 6,732,219         | 14.4              | 14.2                         |
| Iowa           | 528             | 3,155,070         | 16.7              | 16.7                         |
| Kansas         | 523             | 2,913,314         | 18.0              | 18.2                         |
| Kentucky       | 756             | 4,467,673         | 16.9              | 16.5                         |
| Louisiana      | 704             | 4,648,794         | 15.1              | 15.0                         |
| Maine          | 276             | 1,344,212         | 20.5              | 19.4                         |
| Maryland       | 657             | 6,045,680         | 10.9              | 10.3                         |
| Massachusetts  | 647             | 6,892,503         | 9.4               | 8.7                          |
| Michigan       | 1,472           | 9,986,857         | 14.7              | 14.3                         |
| Minnesota      | 830             | 5,639,632         | 14.7              | 14.4                         |
| Mississippi    | 436             | 2,976,149         | 14.6              | 14.4                         |
| Missouri       | 1,141           | 6,137,428         | 18.6              | 18.2                         |
| Montana        | 289             | 1,068,778         | 27.0              | 26.2                         |
| Nebraska       | 309             | 1,934,408         | 16.0              | 16.1                         |
| Nevada         | 642             | 3,080,156         | 20.8              | 19.8                         |
| New Hampshire  | 255             | 1,359,711         | 18.8              | 17.5                         |
| New Jersey     | 762             | 8,882,190         | 8.6               | 8.0                          |
| New Mexico     | 513             | 2,096,829         | 24.5              | 24.0                         |
| New York       | 1,705           | 19,453,561        | 8.8               | 8.3                          |
| North Carolina | 1,358           | 10,488,084        | 12.9              | 12.5                         |
| North Dakota   | 136             | 762,062           | 17.8              | 18.1                         |
| Ohio           | 1,806           | 11,689,100        | 15.5              | 15.1                         |
| Oklahoma       | 816             | 3,956,971         | 20.6              | 20.5                         |

|                      |               |                    |             |             |
|----------------------|---------------|--------------------|-------------|-------------|
| Oregon               | 906           | 4,217,737          | 21.5        | 20.4        |
| Pennsylvania         | 1,896         | 12,801,989         | 14.8        | 14.1        |
| Rhode Island         | 123           | 1,059,361          | 11.6        | 10.7        |
| South Carolina       | 852           | 5,148,714          | 16.5        | 16.2        |
| South Dakota         | 182           | 884,659            | 20.6        | 20.9        |
| Tennessee            | 1,219         | 6,829,174          | 17.8        | 17.2        |
| Texas                | 3,891         | 28,995,881         | 13.4        | 13.4        |
| Utah                 | 654           | 3,205,958          | 20.4        | 21.2        |
| Vermont              | 110           | 623,989            | 17.6        | 16.0        |
| Virginia             | 1,140         | 8,535,519          | 13.4        | 12.8        |
| Washington           | 1,263         | 7,614,893          | 16.6        | 15.9        |
| West Virginia        | 330           | 1,792,147          | 18.4        | 18.5        |
| Wisconsin            | 845           | 5,822,434          | 14.5        | 14.0        |
| Wyoming              | 170           | 578,759            | 29.4        | 29.3        |
| <b>United States</b> | <b>47,511</b> | <b>328,239,523</b> | <b>14.5</b> | <b>13.9</b> |

**Table A2: Age-standardized 2000 weights**

|                 | <b>Population</b>  | <b>2000 P<sub>sj</sub>/P<sub>s</sub></b> |
|-----------------|--------------------|--|
| Under 1 year    | 3,794,901          | 1.382%                                   |
| 1-4 years       | 15,191,619         | 5.532%                                   |
| 5-14 years      | 39,976,619         | 14.556%                                  |
| 15-24 years     | 38,076,743         | 13.865%                                  |
| 25-34 years     | 37,233,437         | 13.557%                                  |
| 35-44 years     | 44,659,185         | 16.261%                                  |
| 45-54 years     | 37,030,152         | 13.483%                                  |
| 55-64 years     | 23,961,506         | 8.725%                                   |
| 65-74 years     | 18,135,514         | 6.604%                                   |
| 75-84 years     | 12,314,793         | 4.484%                                   |
| 85 + years      | 4,259,173          | 1.551%                                   |
| <b>All ages</b> | <b>274,633,642</b> |  |

**Table A3: 2019 Age-standardized computation**

|              | <b>Suicides</b> | <b>2019 Population</b> | <b>R<sub>j</sub></b> | <b>2000 P<sub>sj</sub>/P<sub>s</sub></b> | <b>P<sub>sj</sub>/P<sub>s</sub> * R<sub>j</sub></b> |
|--------------|-----------------|------------------------|----------------------|--|---|
| < 1 year     | 0               | 3,783,052              | 0                    | 1.382%                                   | 0   |
| 1-4 years    | 0               | 15,793,631             | 0                    | 5.532%                                   | 0   |
| 5-14 years   | 546             | 40,994,163             | 1.3                  | 14.556%                                  | 0.189232  |
| 15-24 years  | 5,954           | 42,687,510             | 13.9                 | 13.865%                                  | 1.927174  |
| 25-34 years  | 8,055           | 45,940,321             | 17.5                 | 13.557%                                  | 2.372561  |
| 35-44 years  | 7,516           | 41,659,144             | 18                   | 16.261%                                  | 2.927046  |
| 45-54 years  | 8,007           | 40,874,902             | 19.6                 | 13.483%                                  | 2.642761  |
| 55-64 years  | 8,228           | 42,448,537             | 19.4                 | 8.725%                                   | 1.69263   |
| 65-74 years  | 4,864           | 31,483,433             | 15.4                 | 6.604%                                   | 1.016944  |
| 75-84 years  | 3,012           | 15,969,872             | 18.6                 | 4.484%                                   | 0.834039  |
| 85 + years   | 1,329           | 6,604,958              | 20.1                 | 1.551%                                   | 0.311722  |
| <b>Total</b> | <b>47,511</b>   | <b>328,239,523</b>     | <b>14.5</b>          |  | <b>13.9</b>   |

$$\Sigma_j P_{sj}/P_s * R_j$$
**Table A4: Differences in age population weights 2000 to 2019**

|             | <b>2000 P<sub>sj</sub>/P<sub>s</sub></b> | <b>2019 P<sub>sj</sub>/P<sub>s</sub></b> | <b>% Change</b> |
|-------------|--|--|-----------------|
| < 1 year    | 1.382%                                   | 1.153%                                   | -16.593%        |
| 1-4 years   | 5.532%                                   | 4.812%                                   | -13.016%        |
| 5-14 years  | 14.556%                                  | 12.489%                                  | -14.202%        |
| 15-24 years | 13.865%                                  | 13.005%                                  | -6.200%         |
| 25-34 years | 13.557%                                  | 13.996%                                  | 3.234%          |
| 35-44 years | 16.261%                                  | 12.692%                                  | -21.952%        |
| 45-54 years | 13.483%                                  | 12.453%                                  | -7.644%         |
| 55-64 years | 8.725%                                   | 12.932%                                  | 48.222%         |
| 65-74 years | 6.604%                                   | 9.592%                                   | 45.250%         |
| 75-84 years | 4.484%                                   | 4.865%                                   | 8.502%          |
| 85 + years  | 1.551%                                   | 2.012%                                   | 29.750%         |