# Trajectories of Suicide Mortality in US Counties from 2000-2017 and the Influence of Selective Migration

Emma Gause

A thesis

submitted in partial fulfillment of the

requirements of the degree of

Master of Science

University of Washington

2020

Committee:

Ali Rowhani-Rahbar

Alice Ellyson

Jonathan Mayer

Program Authorized to Offer Degree:

Epidemiology

University of Washington

**Abstract**

Trajectories of Suicide Mortality in US Counties from 2000-2017 and the Influence of Selective Migration

Emma Gause

Chair of the Supervisory Committee:

Ali Rowhani-Rahbar

Department of Epidemiology

**Objectives:** To identify five distinct latent trajectory classes of suicide mortality among US Counties from 2000 to 2017 and to compare these trajectories to each county's cumulative net migration over the study period to assess the influence of population change due to migration as a potential measure of economic opportunity.

**Methods:** The annual county suicide mortality rates were prepped using a spatial Bayesian smoothing approach using Integrated Nested Laplace Approximations to reduce the influence of extreme high rates due to small numbers. Latent growth mixture modelling was employed to identify five linear mortality trajectories. The model allowed for random slopes and intercepts to allow individual county trajectories to vary around the class means. Each class was compared descriptively against three categories of cumulative net migration over the study period:

persistent population loss due to migration, net even migration, and persistent population gain due to migration.

**Results:** All five classes of latent trajectories displayed a positive slope indicating rising suicide mortality rates in all classes of counties from 2000 to 2017. While two classes exhibited greater slopes, the classes seemed to be more differentiated by their intercepts – the suicide mortality rate in 2000. The proportion of counties in each of the three migration categories differed between the five classes. However, the results of the Lo-Mendell-Rubin test failed to reject the null hypothesis that five classes were more appropriate than four classes, calling into question the validity of the individual county classifications.

**Conclusions:** As the five-class trajectory model was not found to be statistically significant, further research should be done to explore different numbers of classes as well as higher order polynomial slopes to identify the best fit for the data. However, given the similarity in the five class slopes it may be possible that county suicide mortality rates exist on a continuum rather than clustered within distinct classes. Lastly, while the latent trajectories seemed to be marginally associated with cumulative net migration, it is possible that aggregating migration counts across the study period masked local temporal variation that should be explored in a more nuanced manner.

## Dedication

Dedicated to my Grandmother, a math teacher, who taught me not to be intimidated by statistics.

Table of Contents

Background

Suicide is the 10th leading cause of death in the United States overall. Among younger individuals it is an even more prominent contributor to the proportional mortality as the second and fourth leading cause of death among those ten to thirty-four and thirty-five to fifty-four years of age respectively.[1] Additionally, national vial statistics surveillance indicates that suicide mortality rates have been rising steadily in the past two decades at one percent per year in the first half of the first decade of the twenty-first century and more rapidly at two percent per year thereafter.[2] However, the burden of suicide mortality is not shared equally across the United States. Areas of the West and Northwest have considerably higher suicide rates than many other parts of the country, particularly in rural counties which consistently have a higher proportion of deaths by suicide compared to urban counties.[3] While suicide rates have been rising overall in the United States, it is not clear if all counties in the United States are experiencing this trend similarly or if the aggregate reporting masks important local variability.

Contextual dynamics of suicide mortality require additional attention in order to recognize potentially protective or harmful environmental and community factors that contribute to suicide deaths upstream from the individual. Understanding how suicide rates have changed in the past two decades at the county level will help identify healthy communities and their organizational, economic, or social contexts. Since certain known suicide risk factors such as health care access, access to lethal means, and public health infrastructure are most relevant at the local level, identifying outcomes at this same scale is crucial in order to discover associations to potential risk or protective factors, particularly in an ecological study. County geographies may be the smallest unit of area that can reliably be used in ecological research with rare outcomes, and even then close attention should be paid to small numbers issues.

Suicide has been studied ecologically as a social phenomenon for decades, perhaps most famously by famed sociologist Émile Durkheim in his well-known book, "Le Suicide."[4] Though his specific findings are now contested, Durkheim claimed that the level of community integration in the aggregate was a predictor of the suicide rate for an area and although individuals who die by suicide leave the population, the area will continue to experience a similar rate of suicide because of the emergent properties of the social environment. In public health, these emergent properties might be called contextual or upstream factors. Identifying specific contextual factors that influence suicide is an important step in creating healthier environments.

The recent and well publicized deaths of despair literature has implicated suicide deaths in the narrative of self-injurious behaviors as a result of economic stagnation.[5] Ecological economic measures such as poverty, unemployment, and median income have been found to be associated with the rate of drug, alcohol and suicide deaths that comprise deaths of despair in US counties.[6,7] However, the deaths of despair hypothesis is at risk of falling prey to the ecological fallacy without a clear mechanism by which these contextual factors relate to a community's suicide rate. This may partly be an issue of data unavailability for the true measure of interest – economic opportunity. Economic opportunity in this framework can be conceptualized as the perceived availability and accessibility of employment or other benefit in an area that would allow an individual or family to provide for their basic needs and improve their quality of life. Opportunity is similar to hopefulness as an abstract concept that represents a real or perceived potential benefit in the future compared to the present. In this way, opportunity can be considered a contextual measure as it may be available to all regardless of who ultimately capitalizes on it. This paper suggests cumulative net migration as a new measure of both real and perceived opportunity and a possible protective factor for suicide mortality at a community level.

In the economic literature, the decision to move is an economic calculus in which the potential marginal monetary or social benefit of relocating must outweigh the marginal cost of moving, which can be quite large.[8–10] It has been estimated that an interstate move in the US may cost, on average, two-thirds of a household's annual income.[11] There are also other social and emotional considerations that individuals and households weigh in the migration calculus that may affect their quality of life in a new place. Ultimately, a threshold of perceived potential gains must be overcome in order for the decision to migrate to become a better option than remaining in place. In this way, the net migration rate within counties can be considered an observed measure of the sum of underlying individual choices based on the expected benefit of being in a place. This theory has been called, "voting with your feet" since the act of relocating reveals individual preferences or expectations of higher returns or greater opportunity in a new location.[12] However, it is important to note that contrary to the economics' supply and demand systems approach to migration, the return on the migration investment is not equal across demographic groups and may be most advantageous for individuals of higher education.[13–15] Consequently, migration may be most correlated with economic opportunity for those who already have the resources and skills to take advantage of it.

This analysis first attempts to leverage latent growth mixture modelling to identify latent classes of suicide mortality trajectories from 2000-2017 at the county level in the United States. These methods are useful because they can summarize a large pool of complex, heterogeneous data into a few distinguishable categories.[16] Five different latent classes of suicide rate trajectories are hypothesized to exist, characterized by their suicide rates at the beginning of the study period and their rate of change in the death rate over time. As the national rates of suicide have been steadily increasing during this period,[2] the most populous classes are expected to have exhibited increasing mortality rates, while a subgroup of counties with more protective factors may have maintained steady rates over time, and finally a small selection of counties might have experienced a decrease in suicide deaths. The hypothesized five classes include: low suicide rate areas that remain low for the duration of the study period, high suicide rate areas that remain high, low suicide rate areas whose suicide rate increases, high suicide rate areas whose death rates rise further, and lastly a minority of counties who experience a decrease in their suicide rates over the study period, perhaps among counties that start the study period with high suicide rates as a result of regression to the mean. Secondly, I hypothesize that trajectories with low or non-increasing suicide rates over time will have a higher proportion of counties characterized by persistent, cumulative population gain due to migration, reflecting the geography's ability to draw migrants as an area of perceived opportunity or wellbeing.

<u>Data</u>

All suicide deaths in the contiguous United States from 2000 through 2017 were included in the study. Alaska and Hawaii were excluded from the analysis because it is prohibitively costly to move to these states and thus their migration patterns are characterized by unique pull factors. Mortality data were obtained from the Centers for Disease Control and Prevention (CDC) National Vital Statistics System (NVSS) restricted Detailed Mortality files, 2000-2017.[17] The NVSS detailed mortality files record the primary causes of death reported on the death certificate as well as up to five contributing causes of death. Deaths were classified as a suicide if any of the six cause of death fields contained an ICD-10 code for suicide (Table 1). All classified suicides were then aggregated to the decedent's county of residence. As some counties changed their geography or FIPS code between 2000 and 2017, a set of unified county boundaries developed by the Institute for Health Metrics and Evaluation (IHME) at the University of Washington was used to ensure consistent geographical units for the full study period.[18] Annual suicide rates were calculated for each unified county area using the CDC intercensal Race-Bridged population files.[19]

Migration data used in this analysis are from the Internal Revenue Service (IRS) Statistics of Income Division (SOI).[20] Annual migration files for 2011-2016 were downloaded directly from the IRS website and supplemented with 1999-2010 data that had already been compiled by Dr. Matthew Hauer and colleagues.[21] Migration data were lagged one year prior to the suicide deaths data to avoid potential issues of temporality. Aggregate inflow and outflow migration counts for each unified county were then calculated using the IRS exemptions estimates. The IRS suppresses data from counties where fewer than twenty returns are filed. Hot deck imputation was used within unified counties to impute inflow and outflow values in years with missing data. In order to characterize the cumulative migration trends within each county over the study period, the annual inflow counts were summed over the eighteen years of data, and the same was done for the outflow counts. The county's net migration was calculated by subtracting the total outflow from the total inflow. For one particularly small county for which all years of data were surpassed in the IRS data, the cumulative net value was set to zero. Negative values represented counties with fewer incoming migrants than departing migrants across the study period, values close to zero represented net even migration, and positive values represented counties with a greater number of incoming migrants than outgoing migrants. Since the distribution of net migration values was roughly normally distributed (Figure 1), tertials were calculated and categorized in to three equal groups: persistent population loss due to migration, net even migration, and persistent net population gain due to migration.

Methods

Due to the high variability of rates calculated from small numbers, annual county level raw suicide rates were smoothed using a spatial Bayesian approach. Besag, York and Mollié (reparametrized as BYM2) Poisson models were fit to the annual county suicide count data using a population offset, and posterior distributions were calculated using Integrated Nested Laplace Approximations.[22] BYM2 default priors were applied. Spatial neighbors were defined as any counties that shared a common border. This approach allows counties to borrow information from their neighbors as the weighted mean of the rates in adjacent counties. The degree of smoothing in an area is dependent upon the number of outcomes and size of the population in that area, so small number counties are more affected than larger counties. Statistically this is done by including both a spatial random effect and an unstructured random error term into the model. For each county i, the smoothed death rate, $\log(\theta_i)$, is estimated as:

$$\log(\theta_i) = \alpha + S_i + e_i$$

where $\alpha$ is the fixed effect, $S$ is the spatial random effect, and $e$ is the unstructured random effects centered around the true mean. The output from this model provides an estimate of the suicide rate for a hypothetical infinite population living in each county for a single year.

The smoothed county rates from 2000 through 2017 were used as the inputs for the latent growth mixture models (LGMM). Latent growth mixture models are an extension of finite mixture modeling in which latent classes are identified based on the observed longitudinal time-series of suicide mortality.[23] This method assumes that there are distinct subgroups within the larger heterogeneous US population that can be statistically separated into clusters of counties with similar longitudinal trajectories. Each class is defined by two parameters, the intercept and slope, referred to as growth factors.[24] In this analysis the intercept represents the suicide morality rate in 2000 and the slope is the rate of change in the mortality rate with each year from 2000 to 2017. The resulting clusters can thus differ from each other both in the magnitude of the suicide rate at the beginning of the study period as well as the mean rate of change in suicide mortality over time. LGMM models differ from other growth models by their use of random effects for each unit of analysis, in this case counties. This allows individual county intercepts and slopes within each trajectory class to vary around the class mean. This is a key departure from other growth modelling approaches such as latent class growth analysis (LCGA) and the original semi-parametric group-based approach (SGBA), also known as the group-based trajectory modeling (GBTM), which constrain the within class variance of the slope and intercept to be zero.[25–27] For this reason, LGMM models are considered the most flexible of the commonly used growth models.

For this analysis a LGMM model was fit to the county-level Bayesian smoothed suicide rates from 2000 to 2017. The LGMM model specified the identification of five linear trajectories, meaning that each trajectory would be characterized by an intercept and a single linear slope as opposed to allowing higher order polynomials such as quadratic. The model included random intercepts and slopes to allow individual county trajectories to vary around the class slope and intercept. The first model using 1000 random starts did not converge and thus 2000 random starts were used to fit the final model, which did achieve convergence. Each of the five trajectory classes were then characterized by their growth parameters - the mean intercept and mean slope, and their associated ninety-five percent confidence intervals. Counties were then grouped into their latent trajectory classes and plotted over time as well as geographically on a map to visually inspect the longitudinal and spatial clustering. The Lo-Mendell-Rubin test (LMR) test implemented in the Mplus TECH11 output, which tests the goodness of fit between k and k-1 classes,[28] was consulted to test the hypothesis that a five class model was a good fit for the data by comparing it to a simpler model with four classes.

Lastly, the five county-level latent suicide mortality trajectories were compared against the three categories of cumulative migration. Descriptive statistics were used to characterize the relationships between trajectory classes and cumulative migration trends. This work is considered "non-human subjects" research and thus did not require IRB review. Data prepping and spatial analyses were conducted in R version 3.6.2[29] using the following packages: inla,[22,30] rgdal,[31] sp,[32,33] spdep,[34,33] and dplyr.[35] The latent growth mixture models were fit in Mplus version 8.4[36] (Appendix A).

## Results

Eighteen years of data for 3,084 unified counties were compiled for the analysis. Annual county-level suicide rate data were roughly normally distributed, though slightly skewed to the right representing very few counties with particularly high suicide mortality rates. The median raw suicide rate within counties in 2000 was 1.07 deaths per 10,000 people and rose to 1.60 deaths per 10,000 by 2017. The raw rates ranged from 0.00 to 22.83 deaths per 10,000 people in 2000 and 0.00 to 16.42 deaths per 10,000 people in 2017. After smoothing, the median suicide rates changed slightly (1.16 deaths per 10,000 people in 2000 and 1.68 deaths per 10,000 people in 2017), but the range of rates was compressed as the extreme high variable rates were shrunk toward the mean of their neighbors. The range of smoothed suicide rates within counties in 2000 was 0.35 to 2.83 deaths per 10,000 people and 0.54 to 4.05 deaths per 10,000 people in 2017 (Figure 2).

County migration also exhibited clear geographic patterns, though overall it seemed to vary less through time. The median per capita migration inflow in counties remained very stable over the period from 1999 to 2016 with a median rate of 480.3 and 584.7 people per 10,000 for the beginning and end of the period respectively. Unsurprisingly, the median per capita outflow migration rate among counties also held steady for the duration of the study period with a median outflow rate of 481.6 people per 10,000 in 1999 and 563.8 people per 10,000 in 2016. There were, however, marked differences in which counties experienced higher levels of inflow versus outflow. Figure 3 shows the ratio of inflow to outflow counts in each study county for the year 2016. Blue areas display counties with a larger number of people moving to the county than are departing the county, whereas red areas represent areas that are losing more population than they are gaining in migration. Over the eighteen-year study period, there were concentrated areas of consistent net population loss due to migration, including areas of Appalachia, Upstate New York, along the Mississippi River, and four corners area of the United States. Conversely, the Pacific Northwest coast, the Texas Triangle, and Atlanta metro area are dependably areas of high net inflow migration.

Five classes were fit to the smoothed suicide rate trajectories from 2000 to 2017. The 3,084 study counties were unequally distributed between the class groups with the smallest class containing just 36 counties and the largest class containing 1,359 counties (Table 2). The five classes varied in both their intercept values, representing the mean smoothed suicide rate in 2000, and their slopes, the average change in the rate of suicide in each year of the study period. The latent class intercepts resembled the overall median smoothed suicide rate in 2000 though class two had the lowest rate of 0.98 deaths per 10,000 people whereas classes three and four began the study period with high suicide rates at 1.44 and 1.67 deaths per 10,000 people respectively (Table 2). The rate of change in suicide rates annually was similar among classes one, two, and five, but again classes three and four exhibited a markedly higher rate of change with an increasing rate of more than 0.04 deaths per year (Figure 4). A pattern emerged in which classes with higher intercept values also had higher estimated slopes suggesting that counties that started the study period with high suicide rates also experienced larger increases in suicide over time. Geographically, classes one and two tended to occur together while class four counties were always near to class three counties (Figure 5). Class five counties, the largest group of counties, did not exhibit a marked geographical pattern, seeming to exist around the margins of the other class clusters. Although the five classes showed some distinct patterning, the results of the LMR test found that the five-class model fit to the data was not superior to a four-class model ($p=0.21$). This suggests that five may not be the most statistically appropriate number of classes to fit to the county level suicide rate trajectories.

When the latent trajectory groups were compared against the cumulative migration trends, they appear to be fairly evenly distributed across the tertials, with a few exceptions (Table 3). Since the cumulative migration categories were created into three even groups, we would expect that for each trajectory class approximately 33% of the member counties would appear in each category if no association between cumulative migration and the latent trajectories existed. This seemed to be the case for classes one and five. Class two had similar proportions of counties with persistent population loss and net even migration with a smaller proportion of counties with persistent population gain. However, classes three and four exhibit decidedly imbalanced allocations of counties by cumulative migration type with much higher proportions in the persistent population gain category.

## Discussion

This study is unique in two key ways. LGMM models have typically been applied to data looking at an individual's developmental trajectory, or their change over time using a life course perspective, and have been used frequently in psychology and criminology research.[16,37] This work is among the first studies to apply LGMM methods to ecological units,[38] and the only example able to be found by the author in the field of public health. Secondly, while economists and geographers have long been interested in the spatial, economic, and social effects of migration,[9,11,13,39–43] this measure is not common in health research despite its favorable qualities.

Many previous studies have found that there is an association between the decision to migrate and economic characteristics such as a labor market conditions, per capita incomes, and housing affordability.[41] Furthermore, migration may be a better benchmark of economic opportunity compared to other commonly used variables such as unemployment because the ability to obtain and retain work is a more significantly associated with relocation than the average unemployment rate in an area.[42] Recent work has also identified an imbalance in who migrates and who does not. Differing from the neoclassical supply and demand approach to migration in which relocation is a natural systems response to disequilibrium in labor market, research has found that highly skilled or educated workers are more likely or more able to move to new opportunities, and have greater returns on their skills, thus potentially deepening existing spatial disparities in the places left behind.[14,15] As such, migration data may be a significant tool in the repertoire of social determinants of health research.

The perceived benefit of a migration has been considered economically, but recent work has also shown that states with a higher level of subjective well-being attract more inflow migrants.[44] Warmer temperatures, more sunshine, state parks, outdoor recreation, amenities, and lower crime have also been found to be positively associated with greater in migration.[43,45,46] Migration is thus more than simply an economic measure as it might also be related to the belief that things in the future may be better than they have been in the past, and these concepts have measurable spatial structure.

The LGMM failed to reject the null hypothesis that five distinct classes of suicide mortality were a more appropriate fit for the data than four classes. The hypothesized trajectories of persistent low rates, persistent high rates, low rates that increased over time, high rates that continued to increase, and a minority of counties with decreasing rates were not identified. Notably, all five latent classes exhibited a positive slope, indicating that suicide rates were rising in all classes of counties during the study period from 2000 to 2017. In fact, the longitudinal patters of counties were very similar across classes, differing most notably in the value of the suicide mortality rate in 2000 (the intercept). This may draw into question the utility of using an LGMM analysis if the true patterning of suicide mortality trajectories lies along a continuum rather than in distinct subgroups.[23] The persistent positive slopes in all suicide mortality classes found is in agreement with national and state-level analyses that have reported rising suicide rates overall,[2,3] but it is significant that no class of counties was found with a decreasing or non-increasing average slope. The overall trend of rising suicide mortality nationwide seems to be replicating itself in all local contexts and sparing none. This discovery further highlights the growing problem of suicide in all of America.

Descriptively, classes one and two both started with low rates of suicide and had a smaller increasing rate than the other classes. Given that the result of the LMR test did not find the five-class model to be a better fit than the simpler four-class model, the distinction between classes one and two may not be meaningful. Moreover, these classes tended to occur in similar areas of the United States – much of the Midwest, South Texas, Southern California, and much of Georgia – further suggesting they should not be considered distinct. However, although net migration was not included in the growth model, class two contained a much smaller proportion of counties that experienced a persistent net gain in population. It is possible that including a measure of migration as a predictor in the LGMM model might further elucidate whether a meaningful distinction between these two groups exists.

The differences between the other latent trajectories were more apparent. Class three and four both displayed a pattern of high suicide rates in 2000 which continued to rise more steeply in the following eighteen years. Class four, though only containing a total of 36 counties, had both the highest rate of suicide to begin the study period as well as the highest increase in suicide annually, marking these counties as top priorities for suicide prevention strategies (Table 2). While there were too few counties in class four to truly distinguish a geographical pattern, the counties seemed to exist in rural areas predominantly in the West, but also clustered in the Upper Peninsula (Figure 5). In terms of their demographics, class four counties were predominantly smaller population counties, though they were approximately evenly split between counties with persistent population gain, net even, and persistent population loss due to migration (Table 4). Class 5 seemed to include the rest of the counties that fell somewhere in the middle both in terms of their suicide rates in 2000 and the magnitude of their increasing suicide rates thereafter.

Overall, the hypothesis that net in migration would be associated with lower suicide rates or with persistent or decreasing rates over time was not supported by this analysis. In fact, surprisingly and in contradiction to what was expected, the two latent trajectories with the highest starting rates of suicide as well as the largest increases in suicide mortality over the study period also contained the largest proportion of counties that experienced persistent population gain from migration and the smallest proportions of persistent population loss. While there does appear to be an association between suicide mortality trajectories and cumulative net migration, the directionality of the association appears here to be the inverse of what was expected.

Limitations & Next Steps

The LGMM model fit in this analysis allowed for linear slopes only, and thus was unable to identify any other functional forms of the suicide mortality trajectories such as quadratic, exponential, or logarithmic. Given the shape of the raw longitudinal data from Figure 4, it is possible that the addition of quadratic terms might result in a better model fit and new or different classifications identified. It is important to note that the estimation of a higher order polynomial model would substantially increase the complexity and computational burden of the analysis which could lead to convergence issues or other analysis barriers.

Similarly, this analysis assessed the association between the five latent trajectories of suicide mortality with overall migration trends descriptively, but a more sophisticated and sensitive method of assessment would be to include the annual, net, per capita migration values to the LGMM model as time varying covariates. The inclusion of time-varying covariates to the model, including both net migration and other known suicide risk factors, would allow for the direct interpretation of the effect of net migration on the rate of suicide in the following year and may help explain the changes in suicide rates from year to year. This may have a bearing on how latent trajectories are estimated and could change the group membership of individual counties or even the best number of latent classes since a portion of the model variance would be explained by the included covariates.

Even though the Bayesian spatial smoothing was done only over space and not over space and time, it is possible that performing the smoothing to reduce the random Poisson variance of rates based on small numbers also decreased the separation between county trajectories, thus making it more difficult for the LGMM model to identify statistically significant distinct classes. Bayesian spatial smoothing methods are not used in spatial cluster analyses for this reason, although it is less clear how this may affect longitudinal clustering. Future research should assess the LGMM model's performance using raw suicide rate data, even though county level rates can be unstable, in order to inform small numbers research best practice.

Finally, LGMM is an exploratory method of classification and should be interpreted cautiously. Since US counties form the entire universe of this data, there is no way to test the classification scheme identified from his type of analysis on an independent validation dataset to assess if the same trajectories are found. As such, interpretation of individual county contexts using these classifications should be done with care. Future research into longitudinal trajectories of suicide mortality at the county should assess different numbers of classes and model parameterizations, such as quadratic slopes, to determine the best fit for the data.

Conclusions

As an exploratory analysis, the results presented here of the latent growth mixture model should be considered a first step towards understanding how suicide rates have changed locally in the US in the past two decades, and how they may be association with the theory of selective migration. However, the fact that the hypothesized five latent classes were not found to be statistically superior using an LMR test to a four-class model suggests that further research must be done to determine whether distinct classes of longitudinal suicide mortality are present. Given that the classes all had relatively similar slopes and were differentiated more by their intercept values, it is possible that county suicide mortality rates exist on a continuum rather than clustering within distinct classes.[23] If this is the case, then the utility of identifying classes of latent trajectories may be more in the recognition of risk patterns overall than a truly meaningful classification scheme for individual counties. Notably, all classes had a positive slope indicating that suicide rates have been increasing in the United States from 2000 to 2017 in all five trajectories, and classes with a higher suicide rate in 2000 also experienced larger increases in morality over the study period. Lastly, latent trajectories of suicide mortality overall appeared to be marginally associated with cumulative net migration. However, it is possible that aggregating migration counts across the study period concealed local temporal variation that should be explored in a more nuanced manner. Further research is needed to disentangle these two important concepts.

References


1. Hedegaard H, Curtin S, Warner M. *Increase in Suicide Mortality in the United States, 1999–2018*. CDC National Center for Health Statistics; 2020. Accessed June 19, 2020. https://www.cdc.gov/nchs/data/databriefs/db362-h.pdf

2. Hedegaard H, Curtin S, Warner M. *Suicide Rates in the United States Continue to Increase*. CDC National Center for Health Statistics; 2018. Accessed June 19, 2020. https://www.cdc.gov/nchs/products/databriefs/db309.htm

3. Rossen LM, Hedegaard H, Khan D, Warner M. County-Level Trends in Suicide Rates in the U.S., 2005–2015. *American Journal of Preventive Medicine*. 2018;55(1):72-79. doi:10.1016/j.amepre.2018.03.020

4. Durkheim E. *Suicide, a Study in Sociology*. 1951 Edition. (Spaulding J, ed.). Routledge; 1897.

5. Case A, Deaton A. Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *Proc Natl Acad Sci USA*. 2015;112(49):15078-15083. doi:10.1073/pnas.1518393112

6. Monnat SM. "Deaths of Despair" From the Cities to the Hollers: Explaining Spatial Differences in U.S. Drug, Alcohol, and Suicide Mortality. In: *2018 Annual Meeting*. ; 2017.

7. Knapp EA, Bilal U, Dean LT, Lazo M, Celentano DD. Economic Insecurity and Deaths of Despair in US Counties. *American Journal of Epidemiology*. 2019;188(12):2131-2139. doi:10.1093/aje/kwz103

8. Sjaastad LA. The Costs and Returns of Human Migration. *Journal of Political Economy*. 1962;70(5):80-93.

9. Gardner J, Hendrickson JR. If I Leave Here Tomorrow: An Option View of Migration When Labor Market Quality Declines: Option View of Migration. *Southern Economic Journal*. 2018;84(3):786-814. doi:10.1002/soej.12249

10. Davies PS, Greenwood MJ, Li H. A Conditional Logit Approach to U.S. State-to- State Migration. *Journal of Regional Science*. 2001;41(2):337-360. doi:10.1111/0022-4146.00220

11. Bayer C, Juessen F. On the dynamics of interstate migration: Migration costs and self-selection. *Review of Economic Dynamics*. 2012;15(3):377-401. doi:10.1016/j.red.2012.02.002

12. Faggian A, Olfert MR, Partridge MD. Inferring regional well-being from individual revealed preferences: the "voting with your feet" approach. *Cambridge Journal of Regions, Economy and Society*. 2012;5(1):163-180. doi:10.1093/cjres/rsr016

13. Leibbrand C. Unequal Opportunity? Racial, Ethnic, and Gender Disparities in the Returns to Internal U.S. Migration. *Social Currents*. 2020;7(1):46-70. doi:10.1177/2329496519869339

14. Reisinger ME. Sectoral Shifts and Occupational Migration in the United States. *The Professional Geographer*. 2003;55(3):383-395. doi:10.1111/0033-0124.5503014

15. Borjas GJ, Bronars S, Trejo S. Self-Selection and internal Migration in the United States. *Journal of Urban Economics*. 1992;32:159-185.

16. Nagin DS, Tremblay RE. Developmental Trajectory Groups: Fact or a Useful Statistical Fiction? *Criminology*. 2005;43(4):873-904. doi:10.1111/j.1745-9125.2005.00026.x

17. National Center for Health Statistics. *Detailed Mortality Files (2000-2017)*, as Compiled from Data Provided by the 57 Vital Statistics Jurisdictions through the Vital Statistics Cooperative Program..

18. Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, et al. US County-Level Trends in Mortality Rates for Major Causes of Death, 1980-2014. *JAMA*. 2016;316(22):2385. doi:10.1001/jama.2016.13645

19. United States Department of Health and Human Services (US DHHS), Centers for Disease Control and Prevention (CDC), National Center for Health Statistics (NCHS), Bridged-Race Population Estimates, United States July 1st Resident Population by State, County, Age, Sex, Bridged-Race, and Hispanic Origin. Compiled from Revised Bridged-Race 2000-2009 Intercensal Population Estimates (Released by NCHS on 10/26/2012); and Bridged-Race Vintage 2017 (2010-2017) Postcensal Population Estimates (Released by NCHS on 6/27/2018).

20. Internal Revenue Service (1999-2016). *IRS County-to-County Migration and Income Data*. Washington DC: IRS Statistics of Income Division.

21. Hauer M, Byars J. IRS county-to-county migration data, 1990–2010. *DemRes*. 2019;40:1153-1166. doi:10.4054/DemRes.2019.40.40

22. Rue H, Riebler A, Sorbye SH, Illian JB, Simpson DP, Lindgren FK. Bayesian Computing with INLA: A Review. *Annual Reviews of Statistics and Its Applications*, 4(March):395-421, 2017. URL Http://Arxiv.Org/Abs/1604.00860.

23. Warren JR, Luo L, Halpern-Manners A, Raymo JM, Palloni A. Do Different Methods for Modeling Age-Graded Trajectories Yield Consistent and Valid Results? *American Journal of Sociology*. 2015;120(6):1809-1856. doi:10.1086/681962

24. Frankfurt S, Frazier P, Syed M, Jung KR. Using Group-Based Trajectory and Growth Mixture Modeling to Identify Classes of Change Trajectories. *The Counseling Psychologist*. 2016;44(5):622-660. doi:10.1177/0011000016658097

25. Nagin DS. Analyzing Developmental Trajectories: A Semiparametric, Group-Based Approach. *Psycological Methods*. 1999;4(2):139-157. doi:10.1037/1082-989X.4.2.139

26. Nagin DS. Group-Based Trajectory Modeling: An Overview. *Ann Nutr Metab*. 2014;65(2-3):205-210. doi:10.1159/000360229

27. Jung T, Wickrama KAS. An Introduction to Latent Class Growth Analysis and Growth Mixture Modeling. *Social Pers Psych Compass*. 2008;2(1):302-317. doi:10.1111/j.1751-9004.2007.00054.x

28. Asparouhov T, Muthen B. *Using Mplus TECH11 and TECH14 to Test the Number of Latent Classes*.; 2012:17. https://www.statmodel.com/examples/webnotes/webnote14.pdf

29. R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing; 2019. https://www.R-project.org/

30. Rue H, Martino S, Chopin N. Approximate Bayesian Inference for Latent Gaussian Models Using Integrated Nested Laplace Approximations (with discussion). *Journal of the Royal Statistical Society*. 2009;71:319-392. doi:10.1111/j.1467-9868.2008.00700.x

31. Bivand R, Keitt T, Rowlingson B. *Rgdal: Bindings for the "Geospatial" Data Abstraction Library*.; 2019. https://CRAN.R-project.org/package=rgdal

32. Pebesma E, Bivand R. Classes and methods for spatial data in R. *The R Journal*. 2005;5(2). https://cran.r-project.org/doc/Rnews/

33. Bivand R, Brown E, Gomez-Rubio V. *Applied Spatial Data Analysis with R*. 2nd Edition. Springer; 2013. http://www.asdar-book.org/

34. Bivand R, Wong D. Comparing implementations of global and local indicators of spatial association. *TEST*. 2018;27(3):716-748. doi:https://doi.org/10.1007/s11749-018-0599-x

35. Wickham H, François R, Henry L, Müller K. *Dplyr: A Grammar of Data Manipulation*.; 2020. https://CRAN.R-project.org/package=dplyr

36. Muthén, L.K. and Muthén, B.O. (1998-2017). *Mplus User's Guide*. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

37. Berlin KS, Parra GR, Williams NA. An Introduction to Latent Variable Mixture Modeling (Part 2): Longitudinal Latent Class Growth Analysis and Growth Mixture Models. *Journal of Pediatric Psychology*. 2014;39(2):188-203. doi:10.1093/jpepsy/jst085

38. J. Bowers A, R. White B. Do principal preparation and teacher qualifications influence different types of school growth trajectories in Illinois?: A growth mixture model analysis. *Journal of Educational Admin*. 2014;52(5):705-736. doi:10.1108/JEA-12-2012-0134

39. Johnson KM, Curtis KJ, Egan-Robertson D. Frozen in Place: Net Migration in sub-National Areas of the United States in the Era of the Great Recession: Net migration in the US in the Great Recession. *Population and Development Review*. 2017;43(4):599-623. doi:10.1111/padr.12095

40. Hauer ME. Migration induced by sea-level rise could reshape the US population landscape. *Nature Clim Change*. 2017;7(5):321-325. doi:10.1038/nclimate3271

41. Sasser AC. Voting with their feet: Relative economic conditions and state migration patterns. *Regional Science and Urban Economics*. 2010;40(2-3):122-135. doi:10.1016/j.regsciurbeco.2010.02.001

42. Fields GS. Labor Force Migration, Unemployment and Job Turnover. *The Review of Economics and Statistics*. 1976;58(4):407. doi:10.2307/1935872

43. Cebula RJ. Internal Migration Determinants: Recent Evidence. *Int Adv Econ Res*. 2005;11(3):267-274. doi:10.1007/s11294-005-6656-8

44. Hummel D. Inter-State Internal Migration: State-level Wellbeing as a Cause. *J Happiness Stud*. 2016;17(5):2149-2165. doi:10.1007/s10902-015-9689-6

45. Johnson K, Lichter D. *Rural Depopulation in a Rapidly Urbanizing America*. University of New Hampshire Carsey School of Public Policy; 2019. doi:10.34051/p/2020.347

46. Ulrich-Schad JD, Henly M, Safford TG. The Role of Community Assessments, Place, and the Great Recession in the Migration Intentions of Rural Americans: Migration Intentions. *Rural Sociol*. 2013;78(3):371-398. doi:10.1111/ruso.12016

Figures & Tables

Table 1: ICD-10 codes used to classify suicide deaths

| ICD-10 Code | Definition |
|---|---|
| X60 | Intentional self-poisoning by and exposure to nonopioid analgesics, antipyretics and antirheumatics |
| X61 | Intentional self-poisoning by and exposure to antiepileptic, sedative-hypnotic, antiparkinsonism and psychotropic drugs, not elsewhere classified |
| X62 | Intentional self-poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified |
| X63 | Intentional self-poisoning by and exposure to other drugs acting on the autonomic nervous system |
| X64 | Intentional self-poisoning by and exposure to other and unspecified drugs, medicaments and biological substances |
| X65 | Intentional self-poisoning by and exposure to alcohol |
| X66 | Intentional self-poisoning by and exposure to organic solvents and halogenated hydrocarbons and their vapors |
| X67 | Intentional self-poisoning by and exposure to other gases and vapors |
| X68 | Intentional self-poisoning by and exposure to pesticides |
| X69 | Intentional self-poisoning by and exposure to other and unspecified chemicals and noxious substances |
| X70 | Intentional self-harm by hanging, strangulation and suffocation |
| X71 | Intentional self-harm by drowning and submersion |
| X72 | Intentional self-harm by handgun discharge |
| X73 | Intentional self-harm by rifle, shotgun and larger firearm discharge |
| X74 | Intentional self-harm by other and unspecified firearm and gun discharge |
| X75 | Intentional self-harm by explosive material |
| X76 | Intentional self-harm by smoke, fire and flames |
| X77 | Intentional self-harm by steam, hot vapors and hot objects |
| X78 | Intentional self-harm by sharp object |
| X79 | Intentional self-harm by blunt object |
| X80 | Intentional self-harm by jumping from a high place |
| X81 | Intentional self-harm by jumping or lying in front of moving object |
| X82 | Intentional self-harm by crashing of motor vehicle |
| X83 | Intentional self-harm by other specified means |
| X84 | Intentional self-harm by unspecified means |

Table 2: Latent trajectory class estimated slopes and intercepts

| Latent Trajectory Class | Intercept (95% CI)* | Slope (95% CI)* | n |
|---|---|---|---|
| Class 1 | 1.042 (1.011, 1.073) | 0.022 (0.020, 0.024) | 575 |
| Class 2 | 0.978 (0.943, 1.013) | 0.025 (0.023, 0.027) | 705 |
| Class 3 | 1.442 (1.385, 1.499) | 0.043 (0.039, 0.047) | 419 |
| Class 4 | 1.666 (1.546, 1.786) | 0.047 (0.035, 0.059) | 36 |
| Class 5 | 1.147 (1.118, 1.176) | 0.030 (0.028, 0.032) | 1,349 |

*Values represent suicide rates per 10,000 people

Table 3: US counties by their cumulative migration trends and latent suicide trajectories

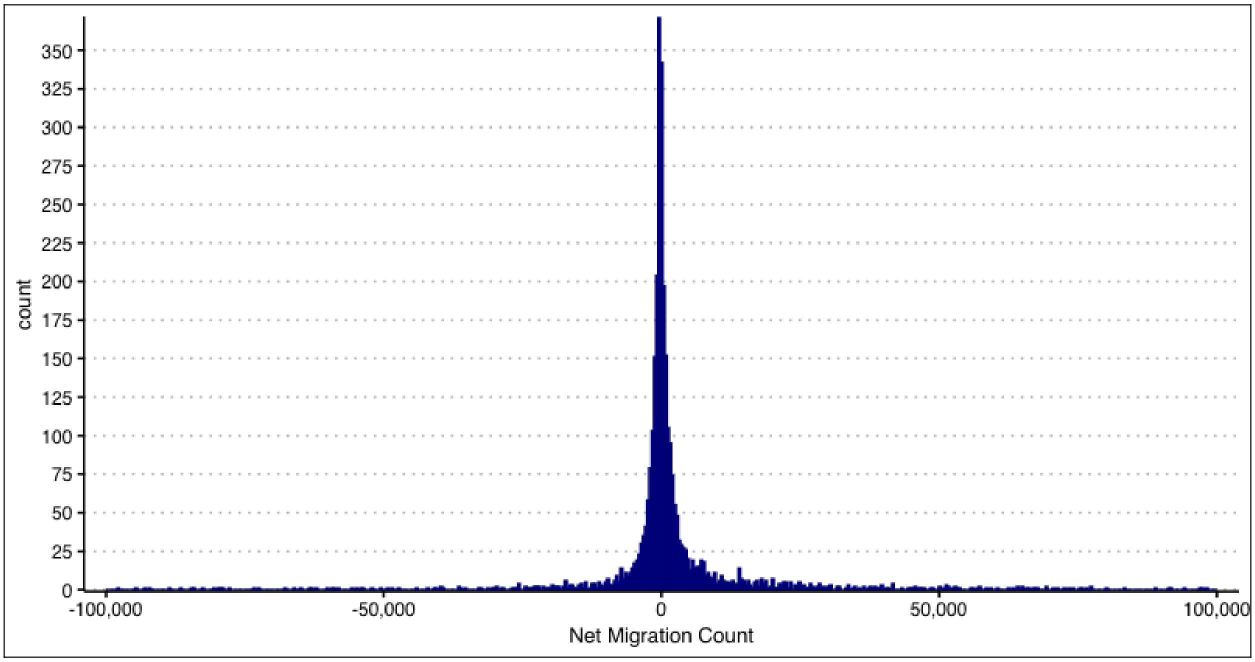
| | Persistent Net Loss<br>n = 1,028<br>n (%)* | Net Even<br>n = 1,028<br>n (%)* | Persistent Net Gain<br>n = 1,028<br>n (%)* |
|---|---|---|---|
| Class 1 | 218 (37.9) | 176 (30.6) | 181 (31.5) |
| Class 2 | 285 (40.4) | 286 (40.6) | 134 (19.0) |
| Class 3 | 89 (21.2) | 142 (33.9) | 188 (44.9) |
| Class 4 | 7 (19.4) | 8 (22.2) | 21 (58.3) |
| Class 5 | 429 (31.8) | 416 (30.8) | 504 (37.4) |

*row percentages calculated

Table 4: Trajectory class four - high suicide mortality counties

| County Name | State | 2017 Population | Cumulative Net Migration Count |
|---|---|---|---|
| Alcona County | Michigan | 10,351 | 1,418 |
| Alpena County | Michigan | 28,462 | -846 |
| Chippewa County | Michigan | 37,711 | -784 |
| Clatsop County | Oregon | 39,182 | 2,508 |
| Cook County | Minnesota | 5,398 | 373 |
| Coos County | Oregon | 63,888 | 4,782 |
| Crawford County | Michigan | 13,907 | 1,323 |
| Curry County | Oregon | 22,669 | 4,203 |
| Douglas County | Nevada | 48,309 | 7,753 |
| Fremont County | Colorado | 47,559 | 5,771 |
| Fulton County | Georgia | 1,041,423 | 26,282 |
| Garland County | Arkansas | 98,658 | 14,029 |
| Iosco County | Michigan | 25,162 | 1,202 |
| Josephine County | Oregon | 86,352 | 14,033 |
| Klamath County | Oregon | 66,935 | 1,381 |
| Kootenai County | Idaho | 157,637 | 37,127 |
| Lake County | Montana | 30,273 | 2,423 |
| Lincoln County | Oregon | 48,920 | 5,053 |
| Mackinac County | Michigan | 10,712 | -531 |
| Martinsville City | Virginia | 13,142 | -1,290 |
| Mohave County | Arizona | 207,200 | 46,118 |
| Monroe County | Florida | 77,013 | -7,421 |
| Montour County | Pennsylvania | 18,272 | -256 |
| Navajo County | Arizona | 108,956 | 9,571 |
| Northampton County | Virginia | 11,846 | 667 |
| Norton City | Virginia | 3,936 | -477 |
| Nye County | Nevada | 44,202 | 13,833 |
| Park County | Colorado | 17,905 | 4,312 |
| Roanoke City | Virginia | 99,837 | -11,162 |
| Schoolcraft County | Michigan | 8,049 | 572 |
| Silver Bow County | Montana | 34,602 | -1,328 |
| Taos County | New Mexico | 32,795 | 850 |
| Teller County | Colorado | 24,646 | 3,964 |
| Trinity County | California | 12,709 | 565 |
| Winchester City | Virginia | 27,932 | -1,879 |

Figure 1: Distribution of cumulative net migration, 1999-2016




*Excluding 62 outlier counties with more extreme net migration values

Figure 2: Smoothed suicide mortality rates in US counties in 2017, by quantile

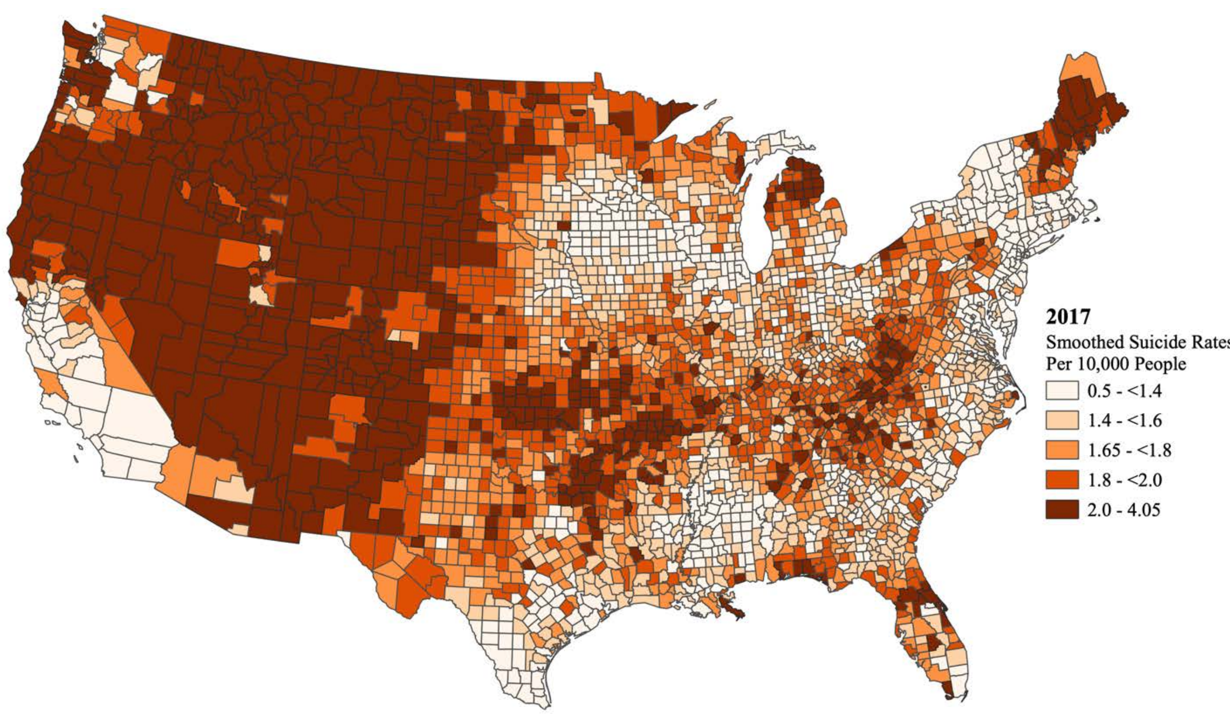


Figure 3: Ratio of inflow to outflow migration in US counties in 2016, by quantile

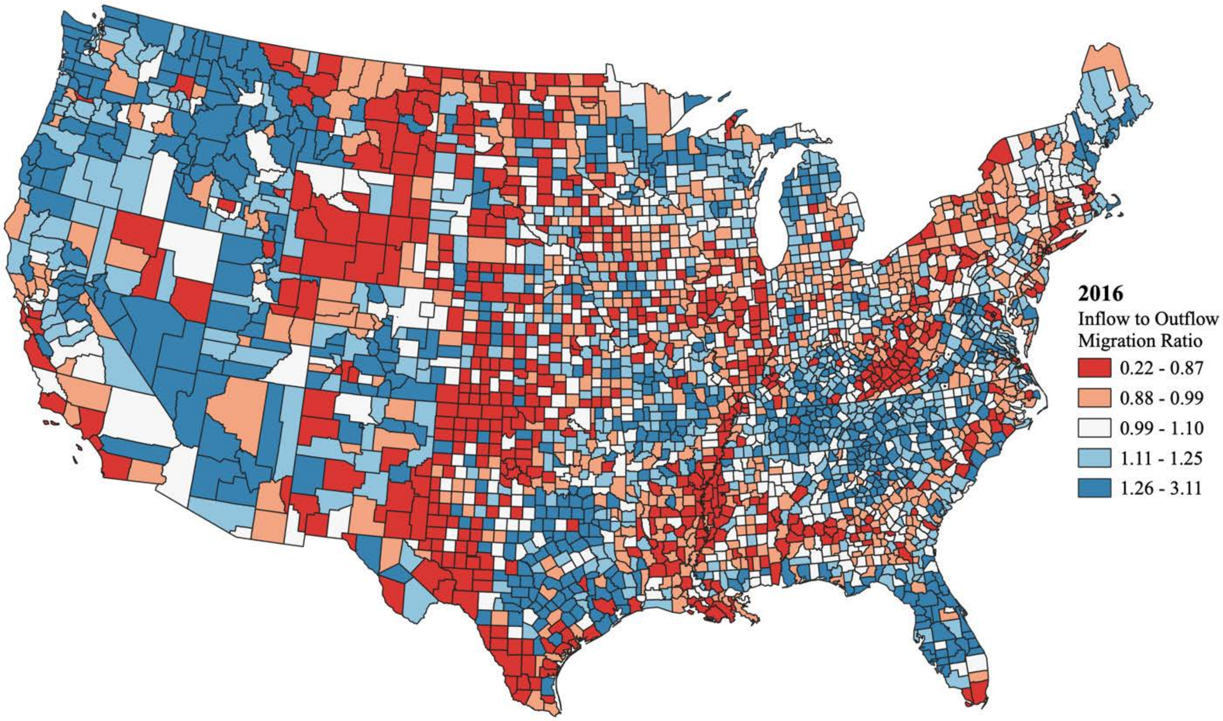


Figure 4: Individual county suicide mortality trajectories in each latent class

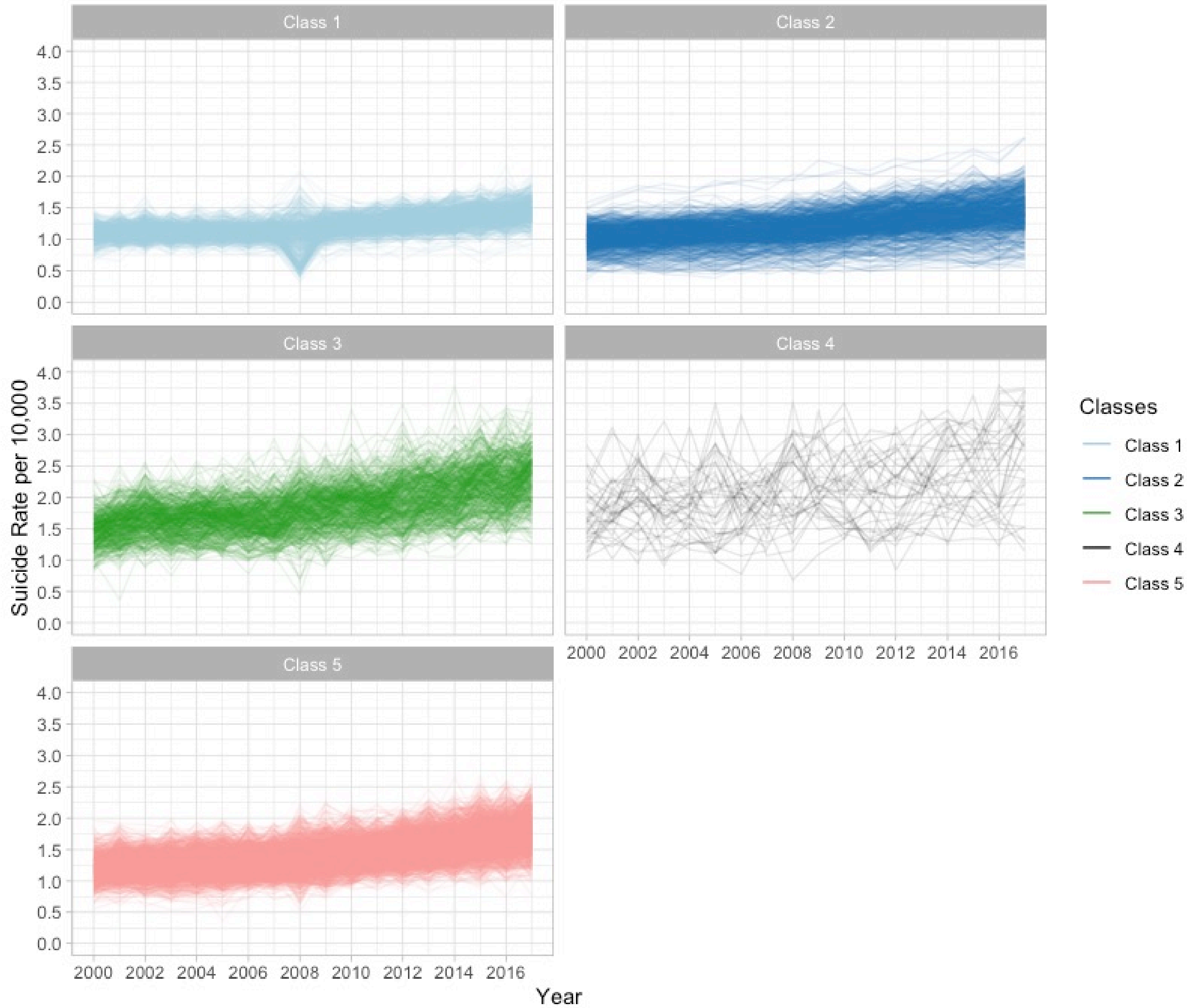
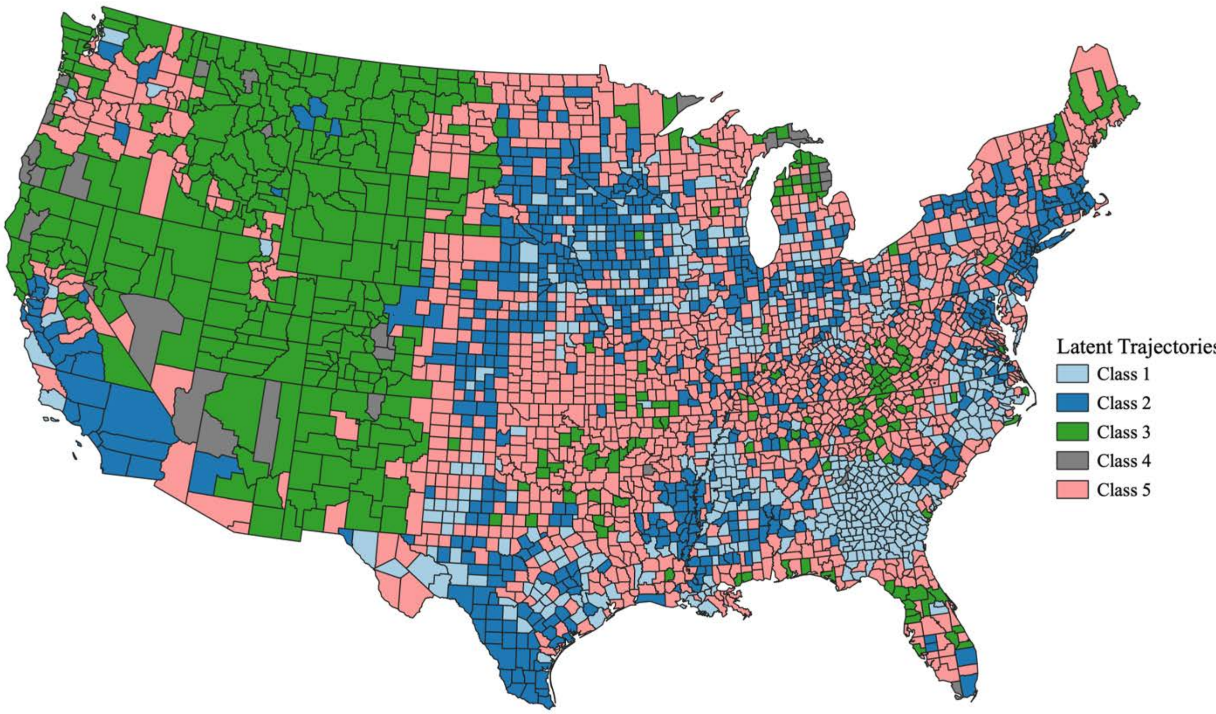


Figure 5: US counties classified by their latent suicide mortality trajectories



Appendix A: Mplus Analysis Code

```
TITLE: LGMM, Linear slopes, 5 classes

DATA:
   FILE IS H:/csv_for_mplus.csv;
   FORMAT IS FREE;

VARIABLE:
   NAMES ARE mcnty smooth00 smooth01 smooth02 smooth03 smooth04 smooth05 smooth06
   smooth07 smooth08 smooth09 smooth10 smooth11 smooth12 smooth13 smooth14
   smooth15 smooth16 smooth17
   sm10_00 sm10_01 sm10_02 sm10_03 sm10_04 sm10_05 sm10_06
   sm10_07 sm10_08 sm10_09 sm10_10 sm10_11 sm10_12 sm10_13 sm10_14
   sm10_15 sm10_16 sm10_17;

   IDVARIABLE = mcnty;

   USEVARIABLES ARE
   mcnty sm10_00 sm10_01 sm10_02 sm10_03 sm10_04 sm10_05 sm10_06
   sm10_07 sm10_08 sm10_09 sm10_10 sm10_11 sm10_12 sm10_13 sm10_14
   sm10_15 sm10_16 sm10_17;

CLASSES = c(5);

ANALYSIS:
   type = MIXTURE;
   MITERATION = 5000;
   STARTS = 2000 500;
   STITERATIONS = 50;
   ALGORITHM = INTEGRATION;

MODEL: %overall%

i s | sm10_00@0 sm10_01@1 sm10_02@2 sm10_03@3 sm10_04@4 sm10_05@5
sm10_06@6 sm10_07@7 sm10_08@8 sm10_09@9 sm10_10@10 sm10_11@11
sm10_12@12 sm10_13@13 sm10_14@14 sm10_15@15 sm10_16@16 sm10_17@17;

%c#1%
i with s;
i s;
sm10_00 sm10_01 sm10_02 sm10_03 sm10_04 sm10_05 sm10_06
sm10_07 sm10_08 sm10_09 sm10_10 sm10_11 sm10_12 sm10_13
sm10_14 sm10_15 sm10_16 sm10_17 (1);

%c#2%
```

```
i with s;
i s;
sm10_00 sm10_01 sm10_02 sm10_03 sm10_04 sm10_05 sm10_06
sm10_07 sm10_08 sm10_09 sm10_10 sm10_11 sm10_12 sm10_13
sm10_14 sm10_15 sm10_16 sm10_17 (2);

%c#3%
i with s;
i s;
sm10_00 sm10_01 sm10_02 sm10_03 sm10_04 sm10_05 sm10_06
sm10_07 sm10_08 sm10_09 sm10_10 sm10_11 sm10_12 sm10_13
sm10_14 sm10_15 sm10_16 sm10_17 (3);

%c#4%
i with s;
i s;
sm10_00 sm10_01 sm10_02 sm10_03 sm10_04 sm10_05 sm10_06
sm10_07 sm10_08 sm10_09 sm10_10 sm10_11 sm10_12 sm10_13
sm10_14 sm10_15 sm10_16 sm10_17 (4);

%c#5%
i with s;
i s;
sm10_00 sm10_01 sm10_02 sm10_03 sm10_04 sm10_05 sm10_06
sm10_07 sm10_08 sm10_09 sm10_10 sm10_11 sm10_12 sm10_13
sm10_14 sm10_15 sm10_16 sm10_17 (5);


Output:
   sampstat TECH1 TECH4 TECH11 TECH14;

SAVEDATA:
   SAVE=CPROBABILITIES;
   FILE IS CPROBS_Linear_5class.dat;
   FORMAT IS FREE;
   ESTIMATES=MIXESTIMATES_Linear_5class.dat;
```

Acknowledgements

I would like to thank my committee members for all of their guidance and support as I finished this thesis. I must give particular thanks to Prof. Ali Rowhani-Rahbar and Prof. Jonathan Mayer for serving on both of Master's Thesis committees this year, and to Dr. Alice Ellyson for her generous willingness to discuss any and all new methods, theories, findings, and setbacks I encountered. I could not have had a better team. Lastly, I would like to thank Dilan Kiley for doing more than his fair share of errands and chores to allow me to devote all of my mental energy on this thesis.

Partial support for this research came from a Eunice Kennedy Shriver National Institute of Child Health and Human Development research infrastructure grant, P2C HD042828, to the Center for Studies in Demography & Ecology at the University of Washington.