Modelling and Forecasting Suicide Rates

Yunus Emre Ergemen and Malene Kallestrup-Lamb

CREATES and Aarhus University

September 21, 2018

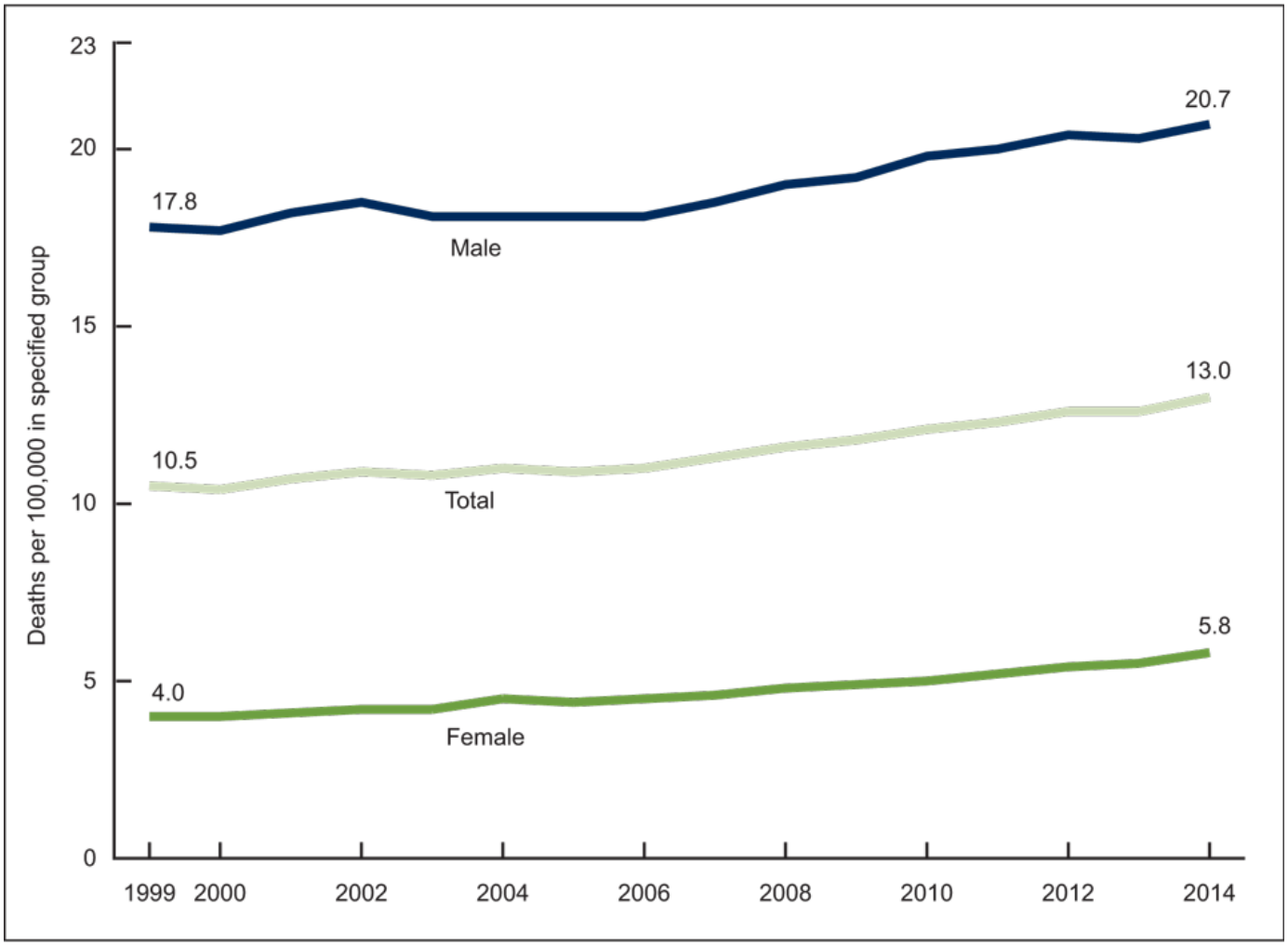
14th International Longevity Risk and Capital Markets Solutions Conference,
Amsterdam

Motivation – Why suicide rates?

- WHO Report 2014 on suicide:
  - Every year, more than 800,000 deaths by suicide
  - Deaths every 40 seconds in the world
  - Second leading cause of death globally for 15-29 years of age
  - Top 10 cause of death for several countries including the US
- Spillover effects
  - Impact on families, friends and communities
  - Loss of knowledge accumulation
- Recent increasing trends in the US and several other countries

Increasing Trends in US Suicide Rates

Figure 1. Age-adjusted suicide rates, by sex: United States, 1999–2014



NOTES: Suicide deaths are identified with codes U03, X60–X84, and Y87.0 from the *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision*. Access data for Figure 1 at: http://www.cdc.gov/nchs/data/databriefs/db241_table.pdf#1.
SOURCE: NCHS, National Vital Statistics System, Mortality.

Characteristics of Suicide Rates Data

- Reflecting a complex phenomenon influenced by the interplay of several factors
  - Social, economic, psychological, cultural, biological, environmental
- Measurement errors
  - Taboo, stigma, shame and guilt leading to under-reporting
  - Miscoding as a mental illness
  - Difficult to analyze in a cross-state (or cross-country) study
- Complex, possibly nonlinear, dynamics

Overview of the Talk

Analyzing Suicide Rates

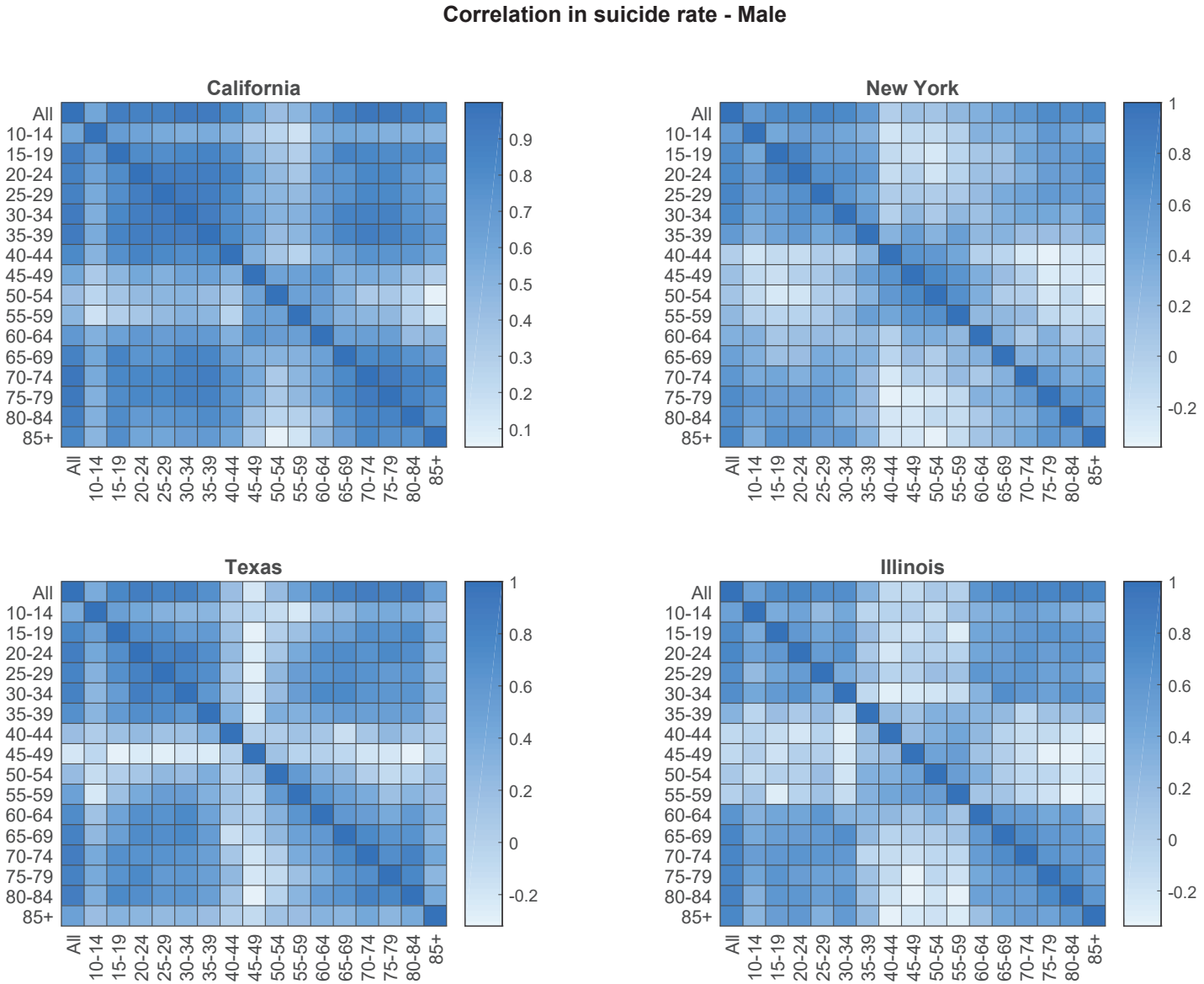
In the literature,

- either psychological or economic factors considered
- oversimplified specifications disregarding interplays of suicide risk factors
- measurement errors neglected

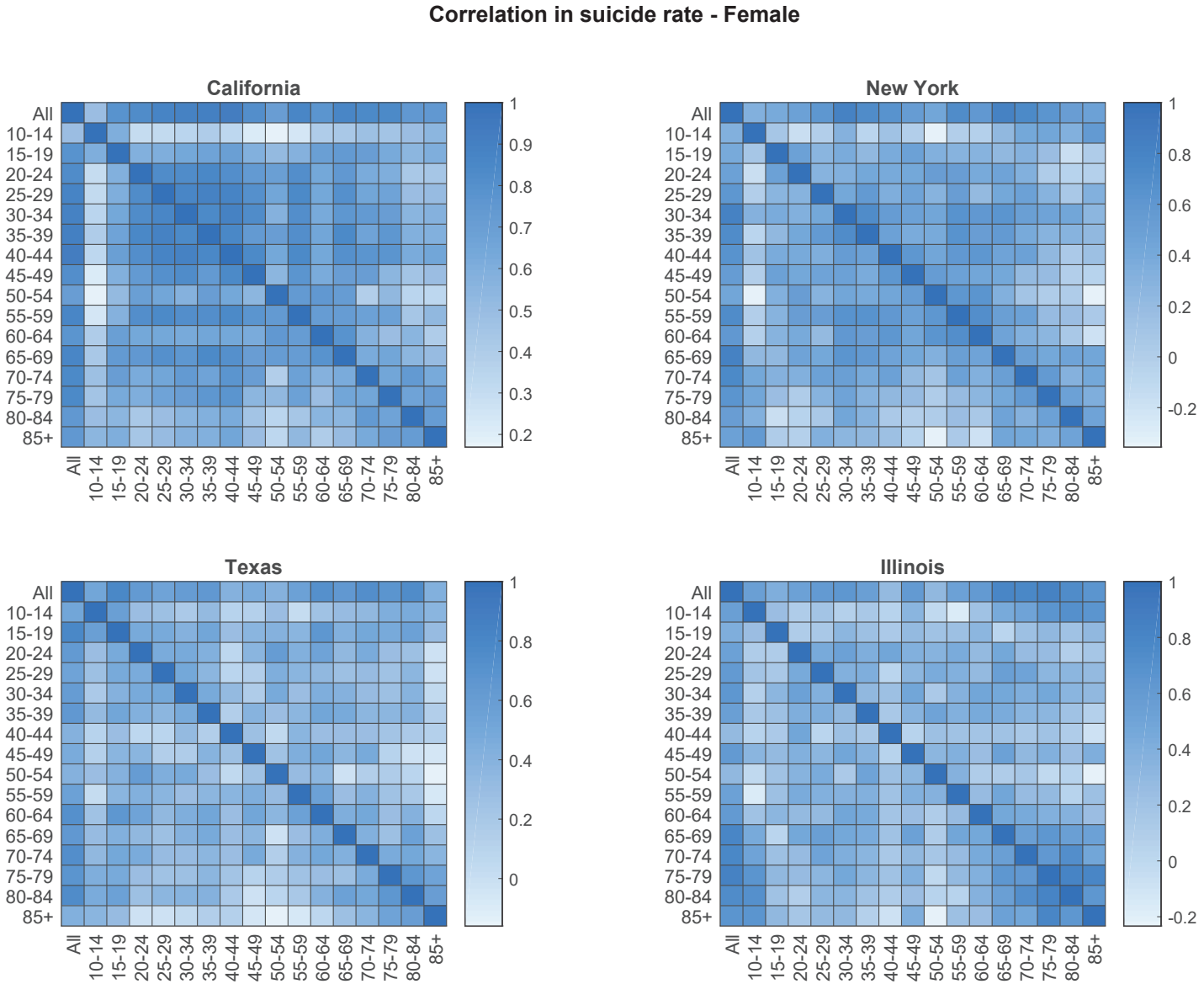
In this paper, we can

- include a combination of all relevant factors affecting suicide rates.
- study possibly nonlinear characteristics.
- account for measurement errors.
- obtain post-available-data forecasts.

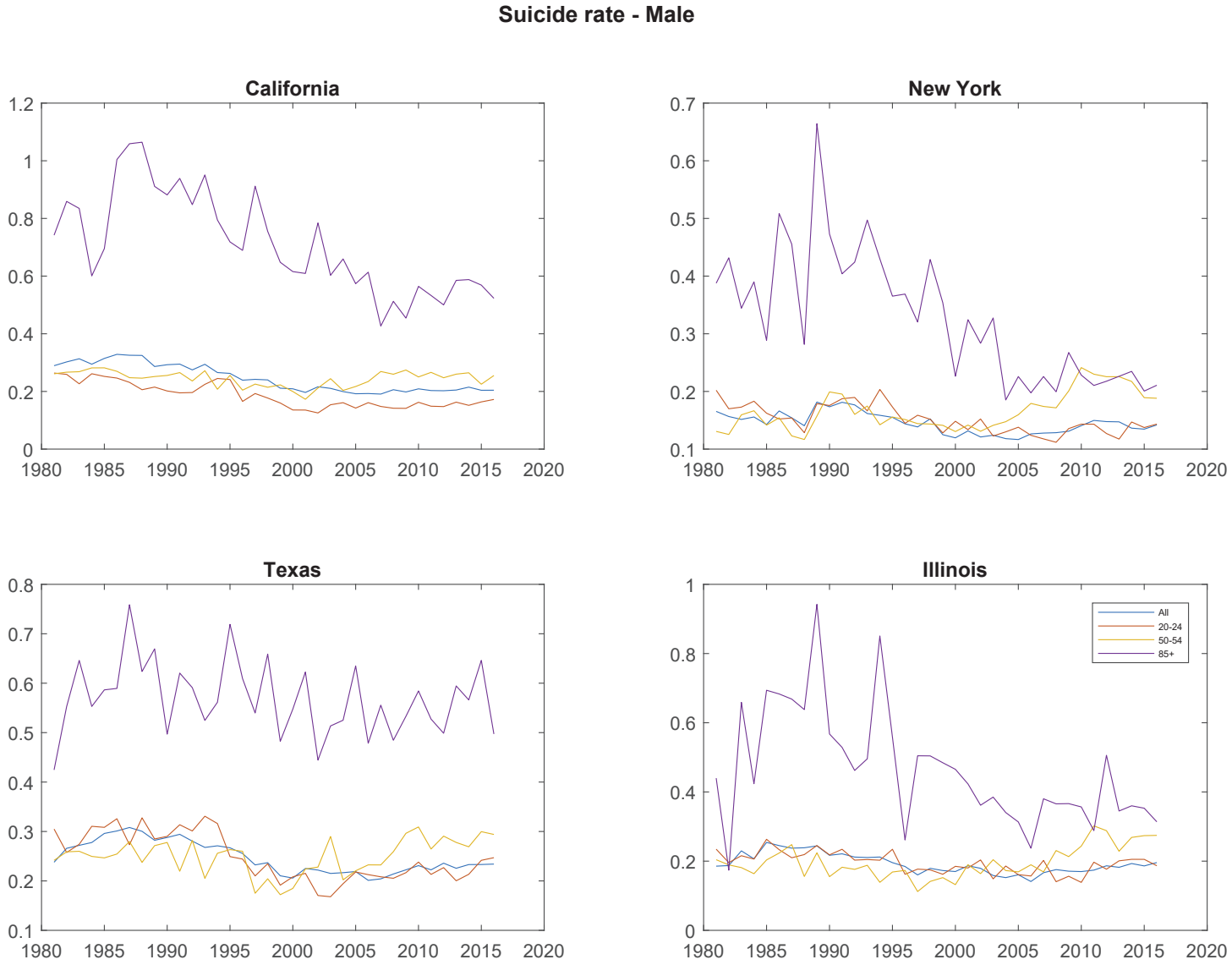
Age Grouping in Suicide Rates



Age Grouping in Suicide Rates



Suicide Rates across States



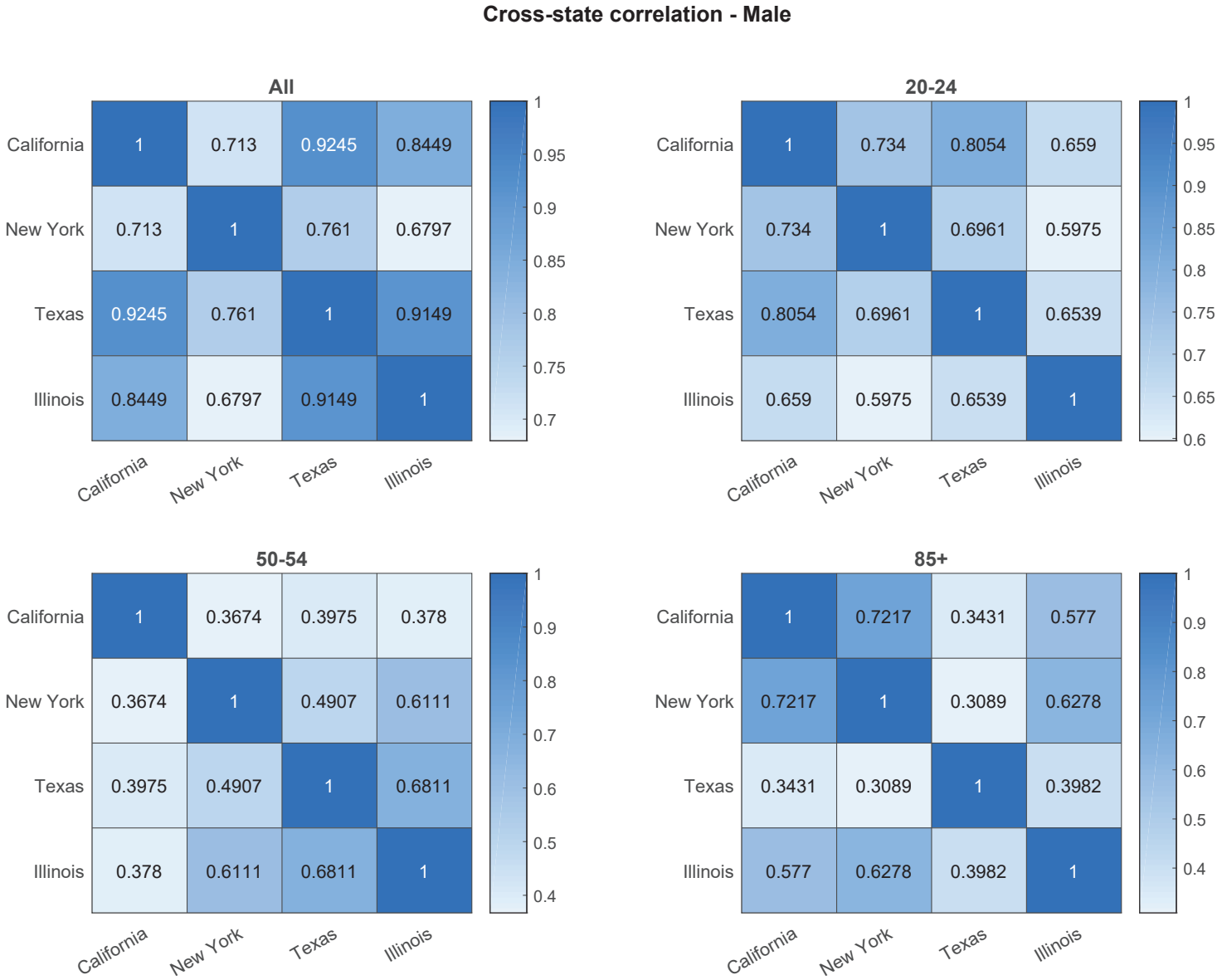
Suicide Rates across States

Data

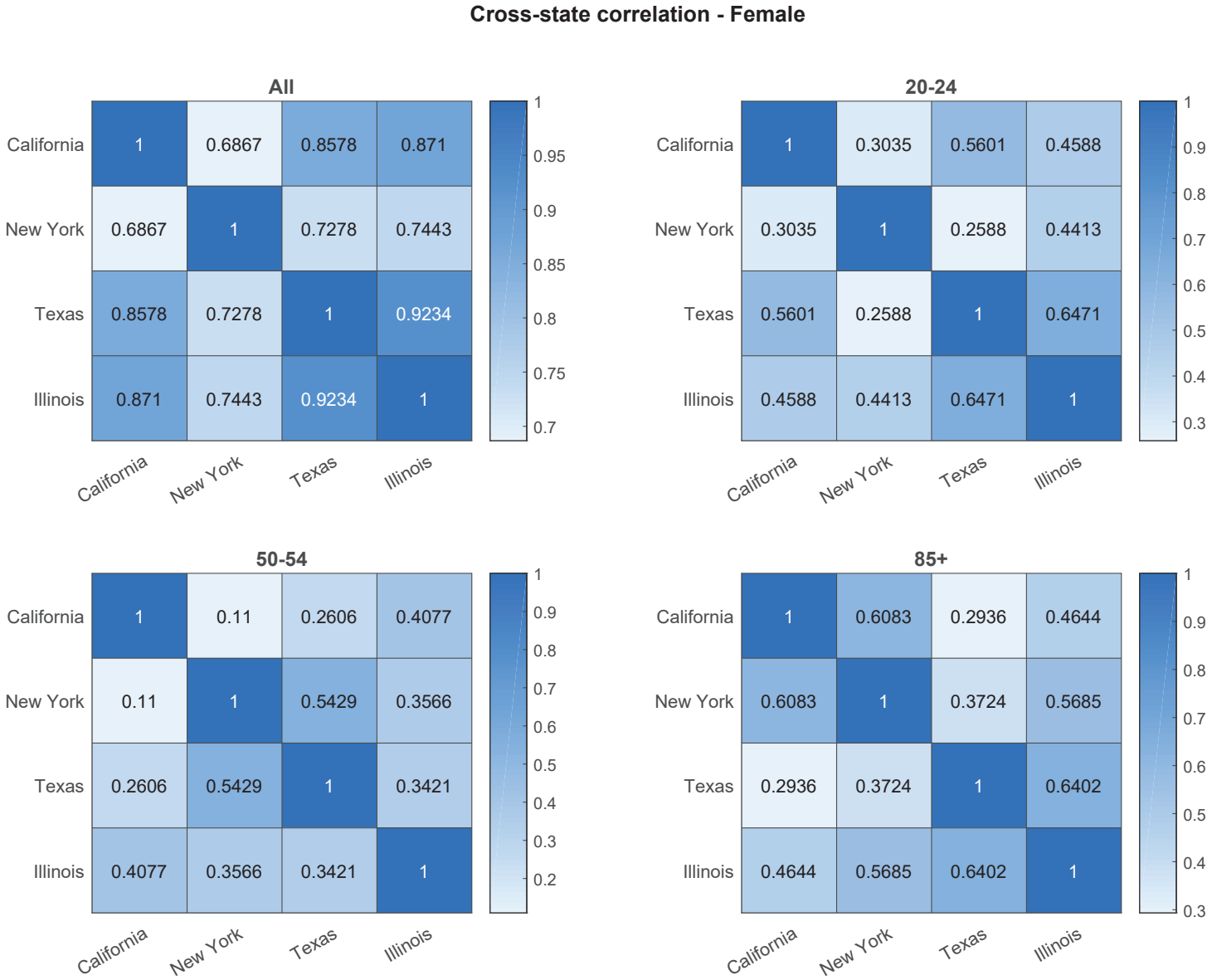
For the period 1981-2016 in California, Texas, New York and Illinois, and for age groups 20-24, 50-54 and 85+,

- mean female suicide rates highest in California
- mean male suicide rates highest in Texas
- male suicide rates around 4x female suicide rates
- nonlinear dynamics
- recent increasing trends in male and female suicide rates
- similarly evolving dynamics for the states pointing at cross-state dependence

Cross-State Dependence in Suicide Rates



Cross-State Dependence in Suicide Rates



Further Characteristics for the All-Age Group – Stochastics

- Unit roots in log male and female suicide rates in California, New York, Texas, Illinois
- Lack of mean reversion due to the existence of unit roots and the necessity for policy interventions
- The first principal component, measuring cross-state dependence, accounting for 86.15% and 85.76% of the total variation in log male and female suicide rates, respectively
- Possibility of developing a suicide index

Methodology

Denoting X_{it} the logarithm of suicide rate in state i at time t, we write

$$X_{it} = \alpha_i + \lambda_i' F_t + \varepsilon_{it}$$

- F_t account for common causes of suicide death, λ_i adjust the magnitude.
- Measurement errors contained either by $\lambda_i' F_t$ or idiosyncratically by ε_{it}
- Persistence allowed in model components

Forecast Model Considerations

- AR

$$X_{it+h} = \mu_{ih} + \beta_{ih}(L)X_{it} + \varepsilon_{it+h},$$

- AR + common factor(s)

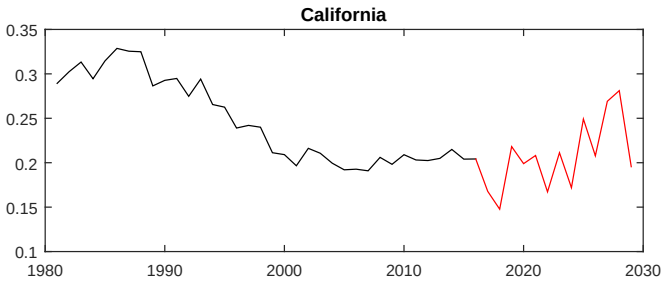
$$X_{it+h} = \mu_{ih} + \beta_{ih}(L)X_{it} + \gamma_{ih}(L)'\hat{F}_t + \varepsilon_{it+h}$$

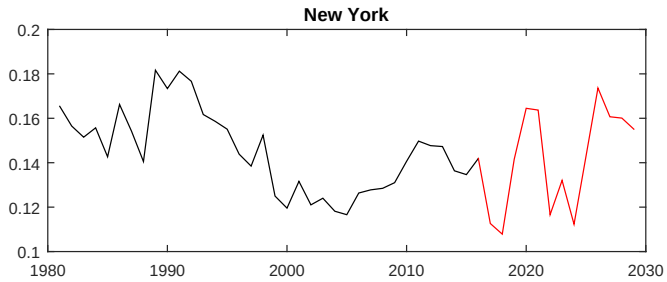
- Nonlinear AR Artificial Neural Network

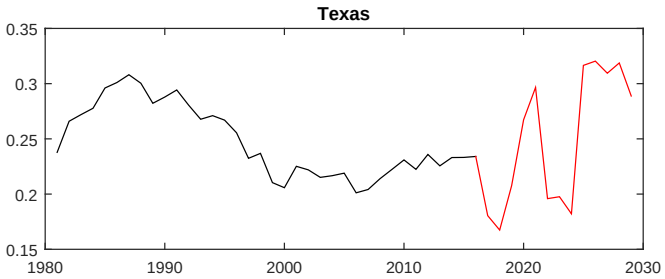
$$X_{it+h} = \mu_{i0h} + \sum_{j=1}^{q} \mu_{j0h}\, g\left(\beta_{i0j} + \sum_{k=1}^{p} \beta_{ikj} X_{it-k}\right) + \varepsilon_{it+h}.$$

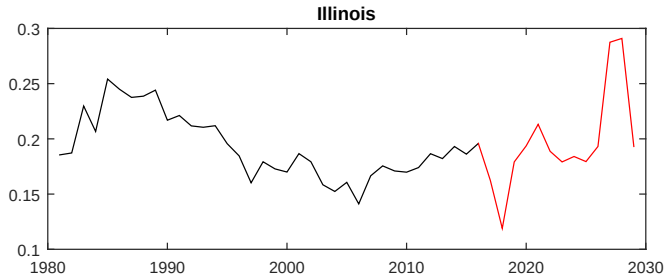
- Nonlinear AR Artificial Neural Network + common factor(s)

Data and Post-Data Forecasts for Male Suicide Rates (All Ages)

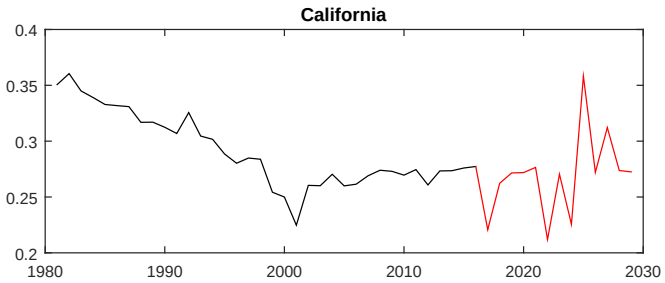


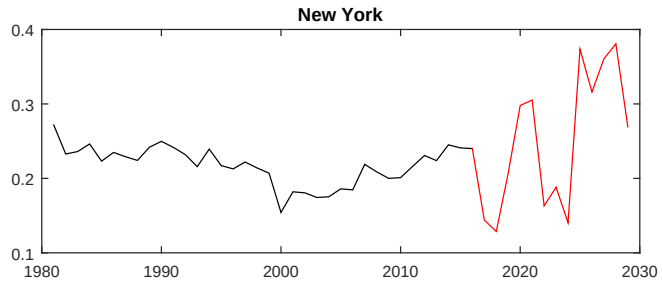


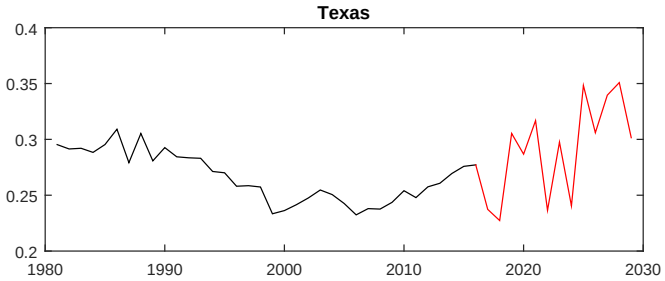


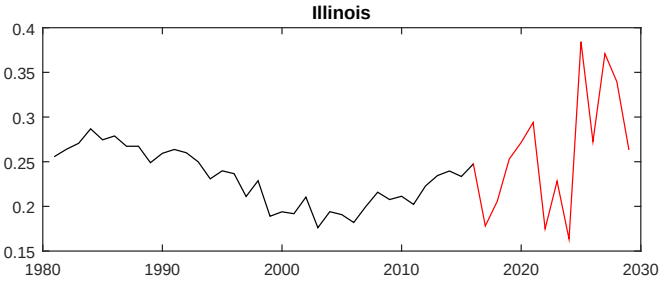


Data and Post-Data Forecasts for Female Suicide Rates (All Ages)









Conclusion

- Analyzing suicide rates is important.
- Our data-driven approach shows
  - there are heterogeneities across the states in terms of their characteristics.
  - there are common characteristics between the states.
  - suicide rates are forecast to increase in some states and/or fail to mean revert.
- Possible implications for pension and insurance planning