

Abstract

Question

We know that stress, discrimination, etc can accumulate biologically in individuals and across generations leading to adverse health effects (Henry and Stephens 2013; Hobfoll, Stevens, and Zalta 2015).

Can adverse health accumulate in places? Can we develop a model for place-based health as a function of migration patterns?

- Goals**
- Better characterize and model the relationship between population migration and population health.
 - Expand upon existing spatial-temporal autoregressive models of county-level health

Hypotheses

Hypothesis A: County-to-county migration patterns improve the explainability of autoregressive models of county-level premature mortality.

Hypothesis B: The role that intercounty migration plays in county-level premature mortality differs significantly between rural and urban counties.

Hypothesis C: Considering bias in migration patterns improves our ability to explain county-level premature mortality as well as the differential role that migration plays in urban versus rural counties

Data

Outcome: county-level age-adjusted premature mortality from CDC WONDER for 2011 through 2019

Primary explanatory factor: county-level IRS migration flow from 2011 through 2019

Secondary explanatory factor: urbanicity as defined by the NCHS, grouped into urban (n=1155) and rural (n=1612)

All analyses are from the perspective of *destination* counties. Counties are included in our analyses if they:

- Are a migration destination
- Have at least 4 years of premature mortality data available

There are a total of **2767 US counties** included in our analyses

Methods

Baseline autoregressive model:

$$y_{it} = \beta_0 + \beta_1 t_{13i} + \beta_2 t_{14i} + \beta_3 t_{15i} + \beta_4 t_{16i} + \beta_5 t_{17i} + \beta_6 t_{18i} + \beta_7 t_{19i} + \beta_8 y_{t-1,i} + v_i t + \mu_i + \epsilon_{it}$$

Where:

- y_{it} represents the predicted premature age-adjusted mortality rate of each destination county i at year t .
- t_{ni} are binary factor variables for each year of available data.
- t is time, in years, treated continuously.
- β_k are unknown regression coefficients.
- $y_{t-1,i}$ is the lagged premature mortality rate for each destination county i at year $t - 1$.
- μ_i is a random intercept for each destination county i .
- v_i is a random slope for the effect of time t on each county i .
- ϵ_{it} represents an error term for each destination county i at year t .

To test Hypothesis A:

We add the following to our baseline autoregressive model:

$$mig_{it} = \frac{\sum_{j \neq i} o ut_{jit} y_{j,t-1} + y_{i,t-1} (pop_{i,t-1} - \sum_{j \neq i} o ut_{ijt})}{\sum_{j \neq i} o ut_{ijt} + (pop_{i,t-1} - \sum_{j \neq i} o ut_{ijt})}$$

Where:

- y_{it} is defined above
- out_{ijt} represents the number of migrants from a unique origin county j who migrated to a destination county i between year $t - 1$ and year t .
- $pop_{i,t-1}$ is the population of a unique destination county i at initial year $t - 1$

Then we compare models with this term to the baseline model

To test Hypothesis B:

We create separate datasets for rural and urban counties and then built rural and urban-specific models using the same baseline autoregressive model and migration weighted average.

To test Hypothesis C:

We add parameter k to our weighted average migration term to simulate health-related selection of individuals for migration:

$$smig_{it} = \frac{\sum_{j \neq i} o ut_{jit} (y_{j,t-1} + k_j) + (y_{i,t-1} + ki) (pop_{i,t-1} - \sum_{j \neq i} o ut_{ijt})}{\sum_{j \neq i} o ut_{ijt} + (pop_{i,t-1} - \sum_{j \neq i} o ut_{ijt})}$$

When $k < 0$ (i.e., decreased premature mortality rate), we simulate migration by individuals who are healthier than the average of their origin county.

When $k > 0$ (i.e., increased premature mortality rate), we simulate migration by individuals who are unhealthier than the average of their origin county.

Assumptions

- Individuals migrate uniformly (i.e., there is no selection of individuals due to health status, but there is selection of destination county)
- Age-adjusted county-level premature mortality is a proxy for county-level overall health
- Our baseline model is correct
- IRS migration flow data accurately represents US county-to-county migration patterns

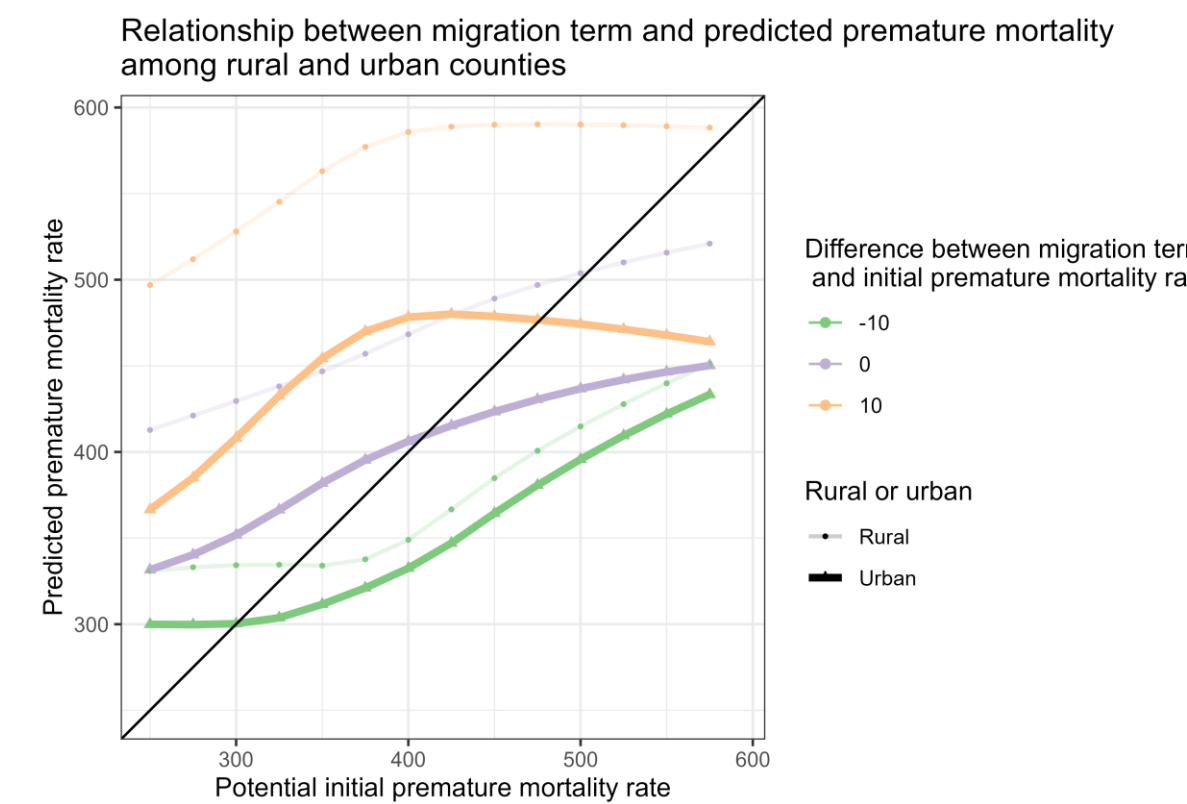
Results and Conclusion

Hypothesis A:

Migrational autoregressive models have more explanatory power than temporal autoregressive models.

Hypothesis B:

Migration from unhealthy counties appears to affect rural counties more than urban counties. Both rural and urban counties experience improvements in county-level health when experiencing in-migration from migrants from healthy counties.



Discussion

- Lagged premature mortality rate $y_{t-1,i}$ and mig_{it} are multicollinear
- We have not yet accounted for spatial autocorrelation, only migration and temporal autocorrelation.
- Premature age-adjusted county-level mortality is an imperfect proxy for county-level overall health
- Does net migration matter?
- IRS migration flow data only includes individuals and households that have filed taxes
- Some rural counties are missing from our analyses
- All analyses are from the perspective of “destination” counties