reading group summer 24
Place-Based Drivers of Mortality: Evidence from Migration (2021)

vaidehi
jk: Amy Finkelstein, Matthew Gentzkow, and Heidi Williams
July 2024

Introduction

- mortality rates vary highly across the US
- health capital + current environment factors
- Chetty et al (2016) showed weak correlation between mortality and environmental conditions
- this paper: identifying causal effect of current location
- movers in Medicare data
- main outcome: life expectancy at the age of 65

Preview of results

- moving from an area at the 10th percentile of estimated place effects to an area at the 90th percentile would increase life expectancy at age 65 by 1.1 years
- findings suggest that health capital also plays an important role
- equalising place effects across areas would reduce the cross-sectional variation in life expectancy at age 65 by 15 percent
- but equalising health capital across areas would reduce the cross-sectional variation by about 70 percent.
- areas with positive place effects tend to have higher-quality hospitals, more primary care physicians and specialists per capita, and higher health care utilisation

Model: Part 1

- $\mathcal{J}$ locations indexed by j, individuals indexed by i
- mortality hazard rate modelled using Gompertz specifiation

$$\log\left(m_{ij}(a)\right) = \beta a + \gamma_j + \theta_i \tag{1}$$

- γ: place effect, θ: health capital, β: age effect

Model: Part 2

Two key assumptions

- Age, place effects, health capital are additively separable - strong assumption,
- γ_j and θ_i are time-invariant

Model: Part 3

Some math

- Non-movers average mortality rate: $\bar{m}_j(a) = \exp\left[\beta a + \gamma_j + \bar{\theta}_j\right]$
- $\gamma_j + \bar{\theta}_j$: avg. mortality index of place j
- avg. life expectancy of area j: $L_j = 65 + \int_{65}^{\infty} S_j(a)da$
- where $S(a)$ is the survival function, $S_j(a) = \exp\left[-\int_{65}^{a} \bar{m}_j(y)dy\right]$

Model: Part 4

Treatment effect

- counterfactual mortality rates: $\bar{m}_j^*(a) = \exp\left[\beta a + \gamma_j + \bar{\theta}\right]$
- counterfactual life expectancy: $L_j^* = 65 + \int_{65}^{\infty} S_j^*(a)da$,
  $S_j^*(a) = \exp\left[-\int_{65}^{a} \bar{m}_j{}^*(y)dy\right]$
- treatment effect: $\Delta = L_j^* - \bar{L}_j$, where $\bar{L}_j$ is the life expectancy computed using average values of both γ and θ over population of non-movers

Empirical Strategy: Part 1

- Health capital
$$\theta_i = X_i\psi + H_i\lambda + \eta^{nm}_{j(i)} + \eta^{\text{orig}}_{o(i)} + \eta^{\text{dest}}_{j(i)} + \tilde{\eta}_i \quad (2)$$
- ηs fixed effects from hypothetical regression of θ on X, H, fixed effects for non-movers' locations, movers' origins, and movers' destinations
- By construction:
- $E(\tilde{\eta}_i \mid X_i, H_i, o(i), j(i)) = 0$ for movers
- $E(\tilde{\eta}_i \mid X_i, H_i, j(i)) = 0$ for non-movers
- Not structural: picks up causal effects and effects of unobservables correlated with X and H

Empirical Strategy: Part 2

- Combine equations 1 and 2

$$\log\left(m_i(a)\right) = \beta a + X_i\psi + H_i\lambda + \tau^{\text{orig}}_{o(i)} + \tau^{\text{dest}}_{j(i)} + \tau^{nm}_{j(i)} + \tilde{\eta}_i \tag{3}$$

, with $\tau^{\text{orig}}_{o(i)} = \gamma^{\text{orig}}_{o(i)}$, $\tau^{\text{dest}}_{j(i)} = \gamma^{\text{dest}}_{j(i)} + \eta^{\text{dest}}_{j(i)}$ for movers, $\tau^{nm}_{j(i)} = \gamma^{nm}_{j(i)} + \eta^{nm}_{j(i)}$ for non-movers

- This is the main estimating equation.
- MLE $\rightarrow$ consistency: $\hat{m}_j(a) = \exp(\hat{\beta}a + \bar{X}_j\hat{\psi} + \bar{H}_j\hat{\lambda} + \hat{\tau}^{nm}_{j(i)})$
- with consistent estimates of γ_j, estimate $\bar{\theta}$ as mean of $\bar{X}_a\hat{\psi} + \bar{H}_a\hat{\lambda} + \hat{\tau}_j{}^{nm} - \hat{\gamma}_j$ across non-movers
- Key issue: unobservables correlated with X_i and H_i that distinguish movers from non-movers

Empirical Strategy: Part 3

Main Idea: Selection Correction

- allow for $\eta_j^{dest} \neq 0$
- standard approach: 2 assumptions.
- equal selection: relationship between observables and treatment is similar to that between unobservables and treatment
- R^2 assumption: the overall importance of the unobservables relative to the observables is small

Empirical Strategy: Part 4

Main Idea: Selection Correction in this setting

- $h_i = H_i\lambda$: observed health index
- Estimate $\hat{\lambda}$ from equation 3 and then estimate following equation using $\hat{h}_i$:

$$h_i = \beta^h a + X_i\psi^h + h_{o(i)}^{\text{orig}} + h_{j(i)}^{\text{dest}} + \tilde{h}_i \tag{4}$$

- Two assumptions in this setting:
  - Proportional selection
  - Relative importance

Empirical Strategy: Part 5

Assumption 1: Proportional Selection

- Selection on unobservables is proportional to selection on observables
- $\text{corr}(T_{ij}, h_{j(i)}^{\text{dest}}) = \varphi_1 \text{corr}(T_{ij}, \eta_{j(i)}^{\text{dest}})$
- $\varphi_1 = 1 \implies$ equal selection

Empirical Strategy: Part 6

Assumption 2: Relative Importance

- Variance of these unobservables relative to the variance of the destination observables is proportional to the corresponding ratio for movers' origins
- $\frac{\text{std}(\eta_{(i)}^{\text{orig}})}{\text{std}(h_{j(i)}^{\text{orig}})} = \varphi_2 \frac{\text{std}(\eta_{j(i)}^{\text{dest}})}{\text{std}(h_{j(i)}^{\text{dest}})}$
- Better than the R^2 assumption

Empirical Strategy: Part 7

Wrapping up

- Baseline assumption: $\varphi_1 = \varphi_2 = 1$
- simple but ugly math yields that A1 $\iff \eta_j^{dest} = \frac{1}{\varphi_1} \frac{\text{std}(\eta^{\text{dest}}_{j(i)})}{\text{std}(h^{\text{dest}}_{j(i)})} h_j^{dest}$
- A2 $\iff \hat{\eta_j}^{\text{dest}} = \frac{1}{\varphi_1 \varphi_2} \frac{\hat{\text{std}}(\tau^{\text{orig}}_{j(i)})}{\hat{\text{std}}(hj(i)^{\text{orig}})} \hat{h_j}^{dest}$
- and finally, $\hat{\gamma}_j = \hat{\tau}_j^{dest} - \hat{\eta}_j^{dest}$ is consistent for γ_j

Data

- 100 % panel of Medicare beneficiaries from 1999 to 2014
- time-variant observed indicators: zipcode of residence for each year, enrollment in Medicare Advantage, Part A, Part B, and Medicaid, age
- time-invariant unobserved indicators: race, gender
- entire universe of Medicare claims
- date of death if person died during the period

Data

- unit of area: commuting zones (CZs); 709 in the US, aggregated counties that approximate local labour markets
- CZs within states aggregated to larger areas if not enough movers have chosen these destinations
- total health-care utilisation measured as total inpatient and outpatient spending
- observable H_i: vector of indicators for 27 Chronic Conditions + log(utilisation + 1)
- observable X_i: race, gender, interaction
- H_i, X_i are measured at year $t^* - 1$, where t^* is year of move for movers, first year for which there is no missing data in $t^* - 1$ for non-movers

Sample Restrictions

- movers: only move once during this period, survive through the end of the move year
- non-movers: random 10% sample of non-movers with complete data
- final sample: 2.3 million movers, 4.3 million non-movers
- movers tend to be older, female, white, and less health

Preliminary evidence of selection

- mean of residuals from a regression of observed health index on age and demographics for movers to destination j

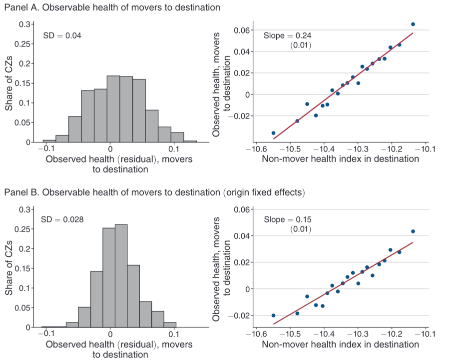

Figure 2. Observable Health and Non-Mover Mortality

Figure 1:

Results: Estimated Treatment Effects

- most favourable: major cities like Chicago
- least favourable: rural areas in the deep South and Southwest

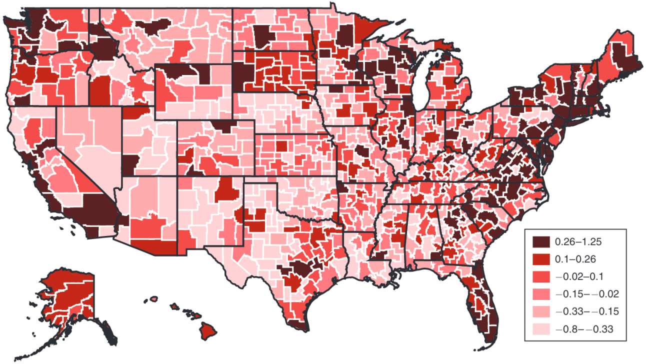

FIGURE 3. LIFE EXPECTANCY TREATMENT EFFECTS

Notes: The map shows the EB-adjusted estimates of life-expectancy treatment effects $(L_j^* - \bar{L})$. Note that small CZs have been aggregated within state (see online Appendix Figure A.1) and a single life expectancy estimate is reported for each aggregate CZ.

Results: Estimated Treatment Effects vs Average Life Expectancy

- positively correlated

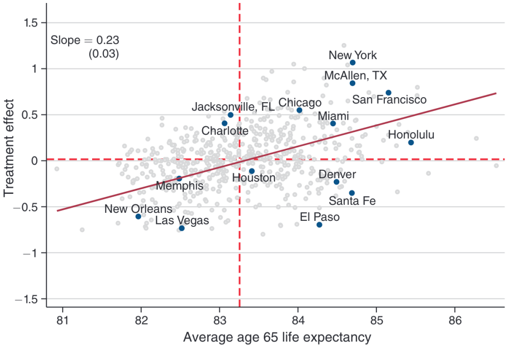

Figure 4. Life Expectancy Treatment Effects versus Life Expectancy

Figure 3:

Results: Interpreting these numbers

TABLE 4—LIFE EXPECTANCY DECOMPOSITIONS

Cross-CZ standard deviation of:	
(1) Age 65 life expectancy (L_j)	0.79 [0.76, 0.83]
(2) Treatment effects ($L_j^* - \bar{L}$)	0.44 [0.32, 0.55]
(3) Health capital effects	0.73 [0.60, 0.83]
(4) Correlation of treatment and health capital effects	−0.04 [−0.15, 0.09]
Share variance would be reduced if:	
(5) Place effects were made equal	0.15 [−0.10, 0.46]
(6) Health capital was made equal	0.69 [0.53, 0.83]

- moving from a twenty-fifth percentile area to a seventy-fifth percentile area would increase life expectancy by 0.60 years; moving from a tenth to a ninetieth percentile area would increase life expectancy by 1.1 years
- about 15 percent of the cross-CZ variance in life expectancy would be eliminated if place effects were made equal across areas
- about 70 percent of the variation would be eliminated if health capital were equalised

Results: Heterogeneity by income and race

- using Medicaid enrollment as proxy for low income
- the standard deviation of life expectancy is larger for individuals on Medicaid compared to those not on Medicaid, and larger for non-White individuals compared to White individuals
- decomposition: the difference is mostly driven by differences in health capital and less so by place effects
- suggesting that variation in area life expectancy for low-income individuals is strongly correlated with health behaviors such as smoking and exercise

Results: Correlates of treament effects

- positively correlated with quality and quantity of healthcare
- negatively correlated with pollution, extreme temperatures, crime, and accidents
- positively correlated with good health behaviours: demand or sorting stories
- positively correlated with education and income: also demand or sorting stories

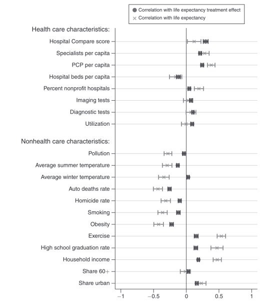

FIGURE 6. CORRELATIONS WITH PLACE CHARACTERISTICS

Figure 4:

Defending the assumptions 1

Assumption: Additive Separability

- health capital and current location affect the level of mortality multiplicatively
- the level of mortality of individuals with poor health capital will vary more across areas than that of individuals who have better health capital
- test: check whether place effects differ across subsets of enrollees partitioned by gender, age at move, health at move
- results: mixed. some deviations of additivity that are statistically significant.

Defending the assumptions 2

Assumption: Time-invariant Health Capital

- reasonable for this population and for the time frame
- threats: immediate changes to behaviour due to the move that have significant impacts on health capital
- defence: literature suggests that threats will be modest
- inelastic health behaviours, smaller impacts, gradual impact
- in order to show that place effects have an immediate impact on mortality, then estimate a logit model for mortality 1, 2, 3, & 4 years after the move

Defending the assumptions 3

Assumption: Selection Assumptions

- assumption: proportional importance, $\varphi = 1$
- test: subset H_i into K groups. Imagine that H_i^{-k} is observable and H_i^k is unobservable.
- want to check that $\frac{\text{std}(\eta^{\text{dest}}_{j(i),k})}{\text{std}(h^{\text{dest}}_{j(i),l})} = \frac{\text{std}(\eta^{\text{orig}}_{j(i),k})}{\text{std}(h^{\text{orig}}_{j(i),k})}, \forall k$
- also relax assumption that constant $= 1$ and re-estimate: results are not sensitive to this baseline assumption

Robustness Checks

- model with coefficients that are different for movers and non-movers
- add some controls, interaction terms
- restrict sample to significant moves, exclude moves to neighbouring CZs
- exclude CA, AZ, FL: snowbirds
- and more...

Conclusion

- significant impacts of current location on mortality
- equalising place effects would reduce the cross-sectional variation in life expectancy at 65 by 15 percent (but less than what equalising health capital would achieve)
- areas with positive place effects tend to have higher-quality hospitals, more primary care physicians and specialists per capita, and higher health care utilisation
- caveat: short-run, partial equilibrium impacts of place on life expectancy for an elderly population (future work for you?!!)
- other caveat: mortality is an absorbing state: assumptions stronger than those in panel analysis
- also important: what aspects of health capital and current environments are important causal determinants of life expectancy

See ya

Parts I didn't understand/try to understand

- sampling error corrections with empiral Bayes procedure and Bayesian bootstrap
- stuff in the appendix
- exact relationship between the R^2 assumption and the relative importance assumption