Motivation

• We face important questions regarding the future of health
  – Health disparities
  – Future burden of diseases
  – Access to care and health care costs

• Tackling these questions is of utmost policy importance

• Answers are difficult because of the complexity of health processes and powerful trends in demography, health behavior, and medical technology.
• Two central models: FEM and FAM
  – Future Elderly Model — Ages 51+, centered around Health and Retirement Study (HRS)
  – Future Americans Model – Ages 25+, centered around Panel Study of Income Dynamics (PSID)

• Both models can be used at the national and the Los Angeles county level
USC’s Roybal Center for Health Policy Simulation develops policy models to tackle these questions

**Funding Sources (since 1999)**

- NIH
- National Institute on Aging
- MacArthur Foundation
- Department of Labor
- Social Security Administration
- National Academy of Sciences
- MacArthur Foundation
- Congressional Budget Office
- Department of Labor
- Social Security Administration
- World Economic Forum
- *Economic Report of the President*

**Accomplishments**

- 62 papers/chapters/briefs, including 2 special issues of *Health Affairs*
- Multiple conferences on aging policy
- Contributions to:
  - National Academy of Sciences
  - MacArthur Foundation
  - Congressional Budget Office
  - Department of Labor
  - Social Security Administration
  - World Economic Forum
  - *Economic Report of the President*
- Diverse set of topics:
  - Obesity, smoking, cardiovascular risk factors
  - Value of delayed aging, Costs of dementia
  - Pharmaceutical price controls, Medicare reform
  - Progressivity of government programs
Who We Are

• Researchers at the University of Southern California (n=20)
  – 9 faculty members and fellows, 2 postdoctoral fellows, 5 mathematicians/statisticians/programmers, 4 PhD students

• External collaborators
  – **International**: OECD, University of Tokyo, University of Rome, National University of Singapore, University of Quebec in Montreal, University of Colima, Korea Institute for Health and Social Affairs
  – **United States**: University of Chicago, Stanford University, University of Pittsburg, University of Texas, University of South Carolina, RAND
  – **Los Angeles**: Los Angeles County Department of Public Health
Today’s Outline

Background and Motivation

Simulation Methods

Data Requirements
FEM tracks the complex interaction between health, mortality, and economic outcomes

- Estimated on
  - Health and Retirement Study Data (longitudinal) for the 51+ population (FEM)
- It tracks economic outcomes such as work, earnings, wealth, medical expenditures (Medicare parts A/B/D, Medicaid and Private), and federal program participation/benefits
- It simulates actual survey respondents and synthetic replenishing cohorts
Modeling Approach

• Demographic and health risk factors -> morbidity/disability/mortality -> economic
• First-order Markov transitions
• Reduced-form models
• Data-driven
<table>
<thead>
<tr>
<th>Category</th>
<th>Health</th>
<th>Economic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binary outcomes</td>
<td>Mortality, cancer, diabetes, heart disease, hypertension, chronic lung disease, stroke, depressive symptoms, Alzheimer’s disease, dementia, congestive heart failure, heart attack</td>
<td>Working for pay, OASI claiming, DI claiming, SSI claiming, live in nursing home, health insurance type</td>
</tr>
<tr>
<td>Ordered outcomes</td>
<td>Activities of daily living (0,1,2,3+), instrumental activities of daily living (0,1,2+), smoking status, subjective well-being</td>
<td></td>
</tr>
<tr>
<td>Continuous outcomes</td>
<td>BMI</td>
<td>Earnings, wealth, property taxes, transfers, helper hours received, volunteer hours, grandchild care hours</td>
</tr>
</tbody>
</table>
Inputs and Outcomes of the Transitions Model

**Inputs**
- Age, sex, race, education
- Lagged risk factors
- Lagged disease status
- Lagged functional status
- Fixed factors from childhood

**Economic outcomes**
- Employment
- Earnings
- Wealth
- Health insurance
- Social security claiming
- Disability insurance claim
- SSI claiming

**Health outcomes**
- Mortality
- Heart disease
- Stroke
- Cancer
- Hypertension
- Diabetes
- Lung disease
- Nursing home status
- BMI
- Smoking (start/stop)
- ADL
- IADL
Diabetes Transition

• Incident diabetes is a function of:
  – Time-invariant: sex, race, education, BMI at 50
  – Time-varying (via 2-year lagged variables): age splines, smoking status, any exercise, log BMI splines
# Data Sources

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health and Retirement Study (HRS)</td>
<td>Host data and estimation of the transition models.</td>
</tr>
<tr>
<td>Social Security Covered Earnings files</td>
<td>Estimation of individual earnings histories. (Subsample of HRS)</td>
</tr>
<tr>
<td>Aging, Dementia and Memory Study (ADAMS)</td>
<td>Estimation of incidence for Alzheimer's disease. (Subsample of HRS)</td>
</tr>
<tr>
<td>National Health Interview Survey (NHIS), National Health and Nutritional Examination Survey (NHANES)</td>
<td>Projection of health trends for replenishing cohorts.</td>
</tr>
<tr>
<td>Medical Expenditure Panel Survey (MEPS)</td>
<td>Estimation of medical costs for non-Medicare individuals.</td>
</tr>
<tr>
<td>Medicare Current Beneficiary Survey (MCBS)</td>
<td>Estimation of medical costs for Medicare recipients</td>
</tr>
<tr>
<td>Census forecasts</td>
<td>Demographics of replenishing cohorts.</td>
</tr>
</tbody>
</table>
Structure of an FEM cohort simulation

Population ages 51-52, 2018

Policy outcomes, 2018

Population ages 53-54, 2020

Policy outcomes, 2020

Population ages 55-56, 2022

Policy outcomes, 2022

Population age 57-58, 2024

Policy outcomes, 2024

Transitions Module

Policy Outcomes Module
Structure of FEM population simulation

1. Population age 51+, 2018
   - Policy outcomes, 2018

2. Replenishing cohort ages 51-52, 2020
   - Population age 53+, 2020
   - Policy outcomes, 2020

3. Replenishing cohort ages 51-52, 2022
   - Population age 53+, 2022
   - Policy outcomes, 2022

4. Replenishing cohort ages 51-52, 2024
   - Population age 53+, 2024
   - Policy outcomes, 2024

---

Replenishing Cohorts Module → Transitions Module → Policy Outcomes Module
Types of Simulation Experiments

• Alter initial characteristics of population
  – Decrease risk factors or disease prevalence

• Change policy module
  – Increase Medicare eligibility age, federal benefit levels, or Social Security claiming rules

• Intervene on transitions
  – Decrease likelihood of developing a disease, delay onset of a disease

• Alter characteristics of replenishing cohorts
Replenishing Cohort Module

• Initial conditions of simulated cohorts are estimated using information about how the mean of the marginal distribution is changing over time and the joint distribution of all variables at a point in time.
• Correlations between variables are held constant while the mean of the marginal distributions are allowed to change with trends.
• Health trends in the simulated cohorts are constrained to meet prevailing health trends in published data or other sources.
• Sampling weights are adjusted to match external estimates of population size (Census).
Handover

• Compare prevalence of chronic diseases and disabilities observed in HRS data from 1998-2014 to FEM forecasts (2010+)

• “Sanity check” on simulation results
  – Any discontinuities?
  – Do trends seem reasonable?
HRS -> FEM (Males, chronic diseases)
HRS -> FEM (Females, chronic diseases)
HRS -> FEM (Males, ADL/IADL)
HRS -> FEM (Females, ADL/IADL)
SHARE -> EU FEM (Males, chronic diseases)
SHARE -> EU FEM (Females, chronic diseases)
SHARE -> EU-FEM (Males, ADL/IADL)
SHARE -> EU-FEM (Females, ADL/IADL)
Validation

• Internal validity
  – Cross-validation for population statistics
  – ROC curves for individuals

• External validity (see technical appendices)
  – Compare to external sources for observed years

• External corroboration
  – Compare to other forecasts
Internal Validity - Crossvalidation

• Randomly split HRS into two groups
  – Estimate transition models on one group
  – Simulate the other group
  – Compare prevalence of disease between the two groups in observed years
## US FEM Internal Validity - Crossvalidation

<table>
<thead>
<tr>
<th>Condition</th>
<th>2000</th>
<th>HRS</th>
<th>p-value</th>
<th>2006</th>
<th>HRS</th>
<th>p-value</th>
<th>2012</th>
<th>HRS</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>11.9%</td>
<td>12.0%</td>
<td>0.77</td>
<td>16.9%</td>
<td>16.6%</td>
<td>0.62</td>
<td>21.9%</td>
<td>22.3%</td>
<td>0.63</td>
</tr>
<tr>
<td>Diabetes</td>
<td>14.1%</td>
<td>13.9%</td>
<td>0.57</td>
<td>19.4%</td>
<td>20.0%</td>
<td>0.32</td>
<td>24.6%</td>
<td>24.5%</td>
<td>0.91</td>
</tr>
<tr>
<td>Heart Disease</td>
<td>20.4%</td>
<td>19.9%</td>
<td>0.41</td>
<td>27.1%</td>
<td>26.2%</td>
<td>0.20</td>
<td>34.9%</td>
<td>32.9%</td>
<td>0.02</td>
</tr>
<tr>
<td>Hypertension</td>
<td>45.9%</td>
<td>44.4%</td>
<td>0.02</td>
<td>57.8%</td>
<td>57.1%</td>
<td>0.36</td>
<td>67.6%</td>
<td>67.1%</td>
<td>0.60</td>
</tr>
<tr>
<td>Lung Disease</td>
<td>7.7%</td>
<td>7.3%</td>
<td>0.37</td>
<td>10.5%</td>
<td>10.0%</td>
<td>0.24</td>
<td>13.0%</td>
<td>12.2%</td>
<td>0.20</td>
</tr>
<tr>
<td>Stroke</td>
<td>6.6%</td>
<td>6.5%</td>
<td>0.72</td>
<td>9.2%</td>
<td>8.8%</td>
<td>0.37</td>
<td>12.7%</td>
<td>11.5%</td>
<td>0.03</td>
</tr>
</tbody>
</table>
EU FEM – Crossvalidation (mortality)

Variable: died at revision 20

Source: EUFEM (1) and Harmonized SHARE (0)
EU FEM – Crossvalidation (cancer)

Variable: cancre at revision 20

Source: EUFEM (1) and Harmonized SHARE (0)
EU FEM – Crossvalidation (diabetes)

Variable: diabe
at revision 20

Source: EUFEM (1) and Harmonized SHARE (0)
EU FEM – Crossvalidation (heart disease)

Variable: hearte at revision 20

Source: EUFEM (1) and Harmonized SHARE (0)
EU FEM – Crossvalidation (hypertension)

Variable: hibpe at revision 20

Source: EUFEM (1) and Harmonized SHARE (0)
EU FEM – Crossvalidation (lung disease)

Variable: lunge at revision 20

Source: EUFEM (1) and Harmonized SHARE (0)
EU FEM – Crossvalidation (stroke)

Variable: stroke at revision 20

Source: EUFEM (1) and Harmonized SHARE (0)
EU FEM – Crossvalidation (any ADL)

Variable: anyadl
at revision 20

Source: EUFEM (1) and Harmonized SHARE (0)
EU FEM – Crossvalidation (any IADL)

Variable: anyiadl
at revision 20

Year

Source: EUFEM (1) and Harmonized SHARE (0)
EU FEM – Crossvalidation (log(BMI))

Variable: logbmi at revision 20

Source: EUFEM (1) and Harmonized SHARE (0)
EU FEM – Crossvalidation (ever smoke)

Variable: smokev at revision 20

Source: EUFEM (1) and Harmonized SHARE (0)
Internal Validity - ROC curves

- 2004-2014 US-FEM (to assess equivalent of 10 year risk)
- 2007-2013 EU-FEM
US FEM – 10 year mortality

Mortality

Area under ROC curve = 0.8096
US FEM – 10 year dementia

Area under ROC curve = 0.8846
US FEM – 10 year cancer

Area under ROC curve = 0.6028
US FEM – 10 year lung disease

Chronic Lung Disease

Area under ROC curve = 0.7145
EU FEM – 6 year mortality

2007-2013 mortality

Minimal

Full

Area under ROC curve = 0.8002

Area under ROC curve = 0.8336
EU FEM – 6 year cancer

2007-2013 incident cancer

Minimal

Full

Area under ROC curve = 0.5960

Area under ROC curve = 0.6116
EU FEM – 6 year diabetes

2007-2013 incident diabetes

Minimal

Full

Area under ROC curve = 0.6135

Area under ROC curve = 0.6992
EU FEM – 6 year heart disease

2007-2013 incident heart disease

Minimal

Full

Area under ROC curve = 0.6567

Area under ROC curve = 0.6819
EU FEM – 6 year hypertension

2007-2013 incident hypertension

Minimal

Full

Area under ROC curve = 0.6231

Area under ROC curve = 0.6552
EU FEM – 6 year lung disease

2007-2013 incident lung disease

Minimal

Full

Area under ROC curve = 0.5722

Area under ROC curve = 0.6166
EU FEM – 6 year stroke

2007-2013 incident stroke

Area under ROC curve = 0.6449

Area under ROC curve = 0.6811
External Corroboration - US FEM Compared to US Census Forecasts

US Population 51+ (millions)

- Census (2012 projection)
- Census (2014 projection)
- US FEM

US FEM Technical Appendix

• **US FEM technical appendix**
  
  - [https://healthpolicy.box.com/v/FEMTechdoc](https://healthpolicy.box.com/v/FEMTechdoc)
  - [https://healthpolicy.box.com/v/estimatesFEM](https://healthpolicy.box.com/v/estimatesFEM)
Our Current Approach to Uncertainty

• “Bootstrap everything”
• Re-sample and re-estimate all transition and cost models from HRS, MEPS, and MCBS, taking into account the sampling strategy of the original data
• Re-simulate with these estimates
Simulation Implementation

• Data processing in SAS and Stata
• Estimation in Stata
• Simulation in C++

• Prefer a Linux environment, but we support users on Mac OS and Windows (using Linux emulation)
Today’s Outline

Background and Motivation

Simulation Methods

Data Requirements
Gateway to Global Aging (g2aging.org) harmonizes several of the HRS-like studies

- United States (HRS)
- Mexico (MHAS)
- England (ELSA)
- 20+ European Countries and Israel (SHARE)
- Costa Rica (CRELES)
- Korea (KLoSA)
- Japan (JSTAR)
- Ireland (TILDA)
- China (CHARLS)
- India (LASI)
Other HRS Sister Studies Not Yet Harmonized

- Brazil (ELSI)
- Indonesia (IFLS)
- Malaysia (MARS)
- New Zealand (HART)
- North Ireland (NICOLA)
- South Africa (HAALSI)
- Scotland (HAGIS)
- China, Ghana, India, Mexico, Russian Federation, and South Africa (WHO-SAGE)
Data with similar structure can also be adapted

• United States (Panel Survey of Income Dynamics)
• Singapore (SCHS)
• Taiwan (TLSA)
Different collaboration models

• Based on US code
  – Embed a programmer at USC (Singapore FEM, Los Angeles County FEM)
  – Regular development calls (Mexico FEM)
  – Less regular development calls (Italian FEM)

• Independent development, but collaborative on research
  – Japanese FEM, Compas
Transition Model Data

• Longitudinal panel
  – Ideally: Regular follow-up time, nationally representative, through death

• Samples large enough to support the analyses of interest
  – Subpopulations, rare diseases
Cost Model Data Requirements

• Different approaches:
  – Longitudinal panel
  – Cross-sectional snapshot
  – Claims data
Replenishing Cohort Data Requirements

• Trends for replenishing cohorts for status in future years:
  – Risk factors (smoking, BMI, etc.)
  – Chronic diseases
  – Demographics (population size, education, etc.)
  – Economic status (work status, earnings, etc.)
Other Data Requirements

• Benefit algorithms if applicable
  – Public pensions, disability programs, etc.
• All-cause mortality adjustment
• Immigration forecasts