# The Future of Death in America

Gary King
Institute for Quantitative Social Science
Harvard University

joint work with Samir Soneji

(talk at the Center for Population and Development Studies, Harvard University, 12/15/08)

# Citations

- Gary King and Samir Soneji. 2008. "The Future of Death in America"
- Gary King and Samir Soneji. 2008. "Eating Away Social Security's Financial Problems"
- Gary King and Federico Girosi. 2008. *Demographic Forecasting*, Princeton University Press.
- copies at `http://gking.harvard.edu`

# The Forecasting Task

- Time series: 25-50 annual mortality rates
- Cross-sections: 1 time series for each age, country, cause, sex, etc.
- Goal: Forecast each time series 25 years
- Challenges: reducing error by:
  - Pooling cross-sections
  - Including demographic knowledge (smooth over time and age)
  - Including biological knowledge (smoking, obesity)

# How (Some) Existing Mortality Forecasts Work

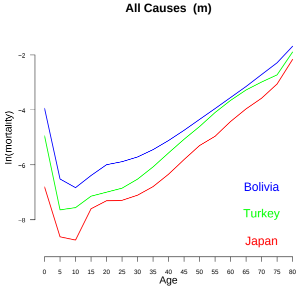
## Procedures:

- Develop private forecasts qualitatively (i.e., informally)
- Adopt a ‘toy’ statistical model
- Get data; produce tentative forecasts with the model
- Adjust model until forecasts fit private views
- Present forecasts, with statistical model as your “method”

## Meaning of procedures

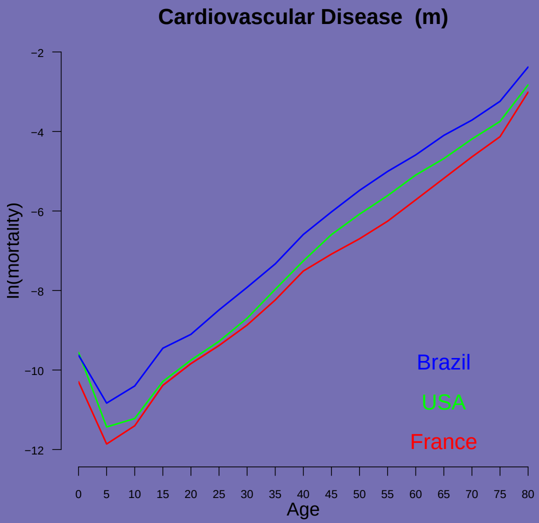
- Forecasts use qualitative information (good!)
- Statistical models add little (bad!)
- Method is invulnerable to being proven wrong
- We bring statistics to demography

# Existing Method 1: Parameterize the Age Profile

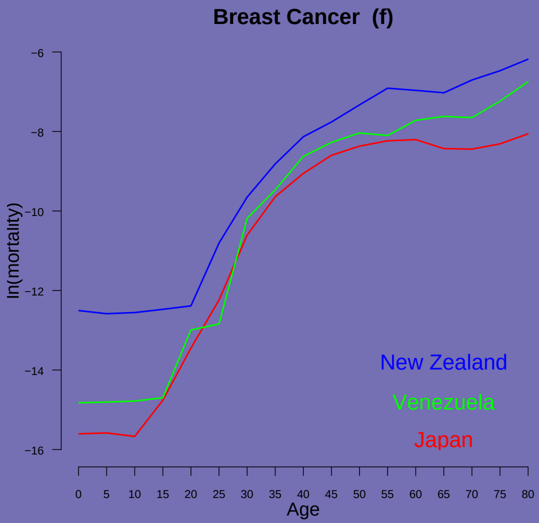



- Gompertz (1825): log-mortality is linear in age after age 20
  - Reduce age-specific rates to 2 parameters ($\mu_{\text{age}} = \gamma_1 + \gamma_2 \times \text{age}$)
  - Forecast only these 2 parameters ($\gamma_1, \gamma_2$)
  - Reduces variance, constrains forecasts
- Dozens of more general functional forms proposed since 1825
- But does it fit anything else?

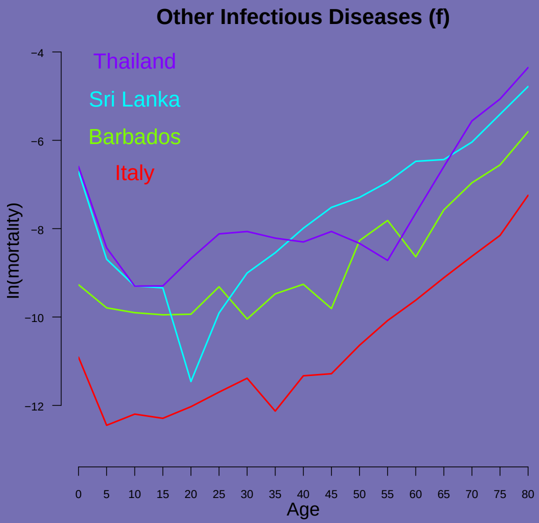
# Mortality Age Profile: The Same Pattern?



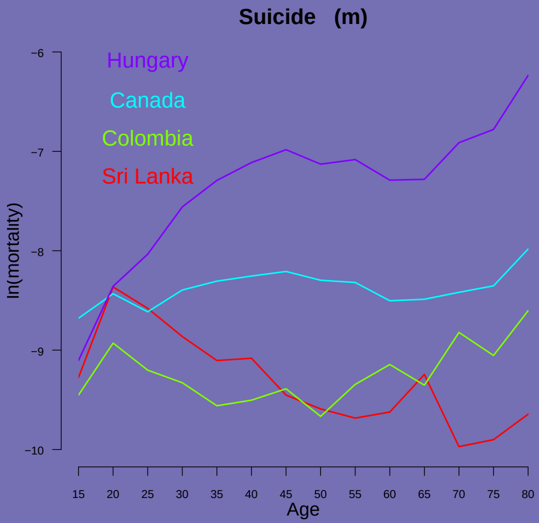
# Mortality Age Profile: The Same Pattern?



# Mortality Age Profile: The Same Pattern?



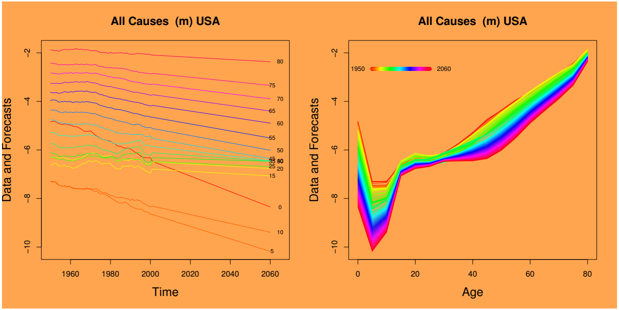
# Mortality Age Profile: The Same Pattern?



# Parameterizing Age Profiles Does Not Work

- No mathematical form fits all or even most age profiles
- Out-of-sample age profiles often unrealistic
- The key empirical patterns are qualitative:
    - Adjacent age groups have similar mortality rates
    - Age profiles are more variable for younger ages
    - We don't know much about levels or exact shapes
- Ignores covariate information

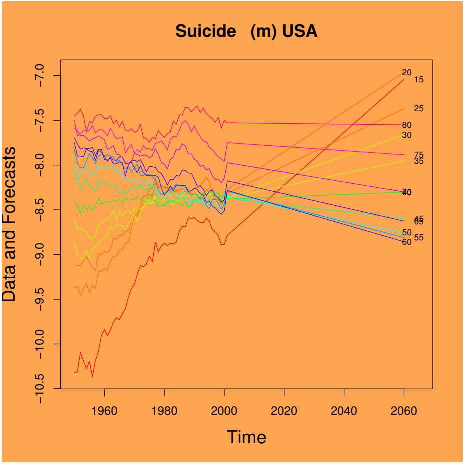
# Existing Method 2: Deterministic Projections

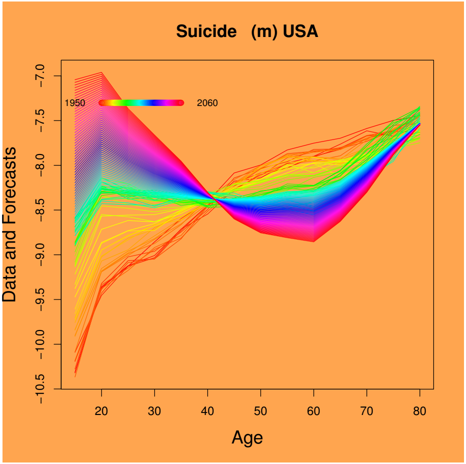



- Random walk with drift; Lee-Carter; least squares on linear trend
- Pros: simple, fast, works well in appropriate data
- Cons: omits covariates; forecasts fan out;
  age profile becomes less smooth
- Does it fit elsewhere?

# The same pattern?

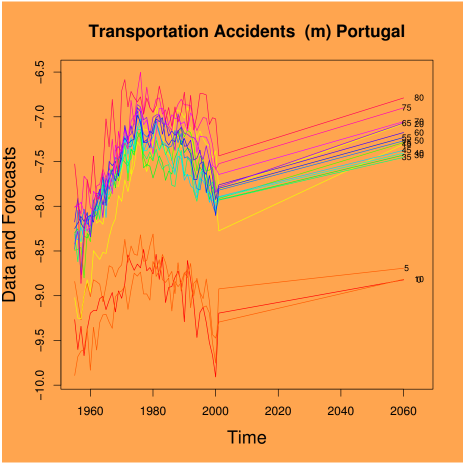
## Random Walk with Drift ≈ Lee-Carter ≈ Least Squares

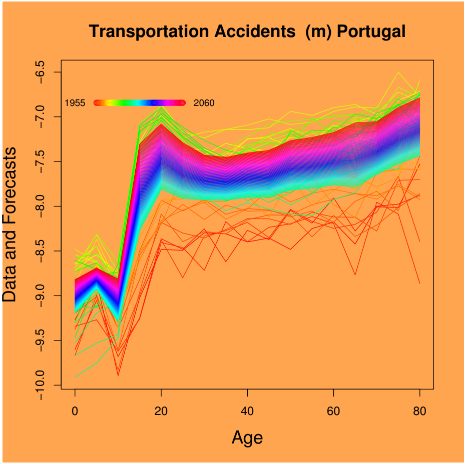


# The same pattern?

## Random Walk with Drift ≈ Lee-Carter ≈ Least Squares



# Deterministic Projections Do Not Work

- Linearity does not fit most time series data
- Out-of-sample age profiles become unrealistic over time

# Existing Method 3: Stacked Regression
(Murray and Lopez, 1996)

- Model mortality over countries ($c$) and ages ($a$) as:

$$m_{cat} = \mathbf{Z}_{ca,t-\ell}\boldsymbol{\beta}_{ca} + \epsilon_{cat} \quad , \quad t = 1, \ldots, T$$

- $\mathbf{Z}_{ca,t-\ell}$ : covariates lagged $\ell$ years.
- $\boldsymbol{\beta}_{ca}$ : coefficients to be estimated
- Equation by equation estimation: huge variances
- Pool over countries: $\boldsymbol{\beta}_{ca} \Rightarrow \boldsymbol{\beta}_{a}$
  - Small variance (due to large $n$)
  - large biases (due to restrictive pooling over countries),
  - considerable information lost (due to no pooling over ages)
  - same covariates required in all cross-sections
- (It always seems ok to pool over variables outside your own field.)

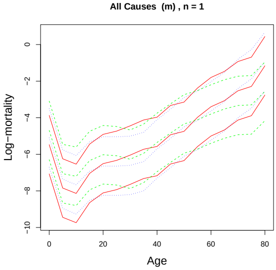
# The New Approach

- Start with separate equation-by-equation regressions
- Use Bayesian priors to smooth across age, time, age$\times$time, etc.
- Put priors on E(mortality), not coefficients
- No arbitrary normalizations
- Different covariates allowed in each regression
- Only one smoothing parameter to represent demographic information
- $\rightsquigarrow$ An easy-to-use software program, YourCast

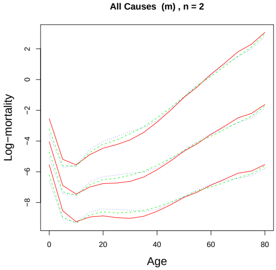
# Formalizing (Prior) Indifference (so no cooking the books)

equal color = equal probability

Level indifference

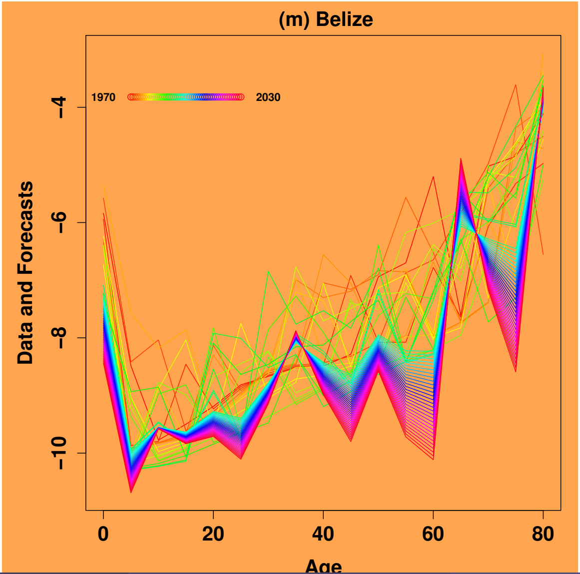



Level and slope indifference



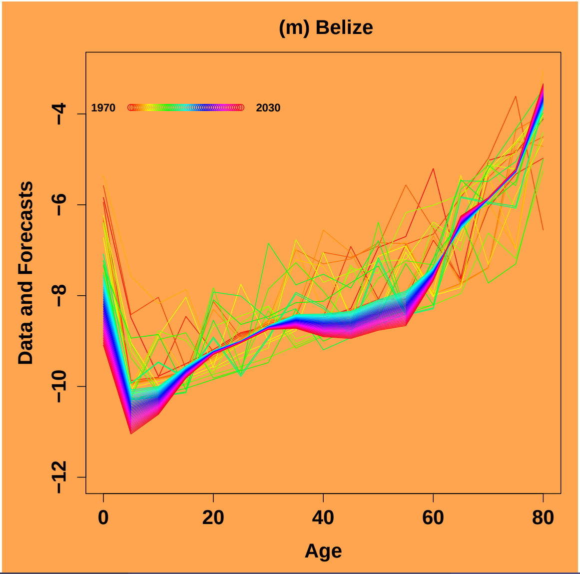
# Mortality from Respiratory Infections, Males

## Least Squares



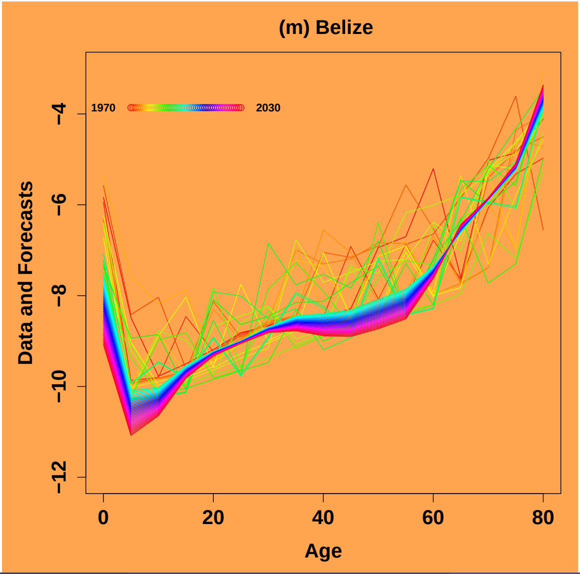
# Mortality from Respiratory Infections, males, $\sigma = 2.00$

## Smoothing over Age Groups



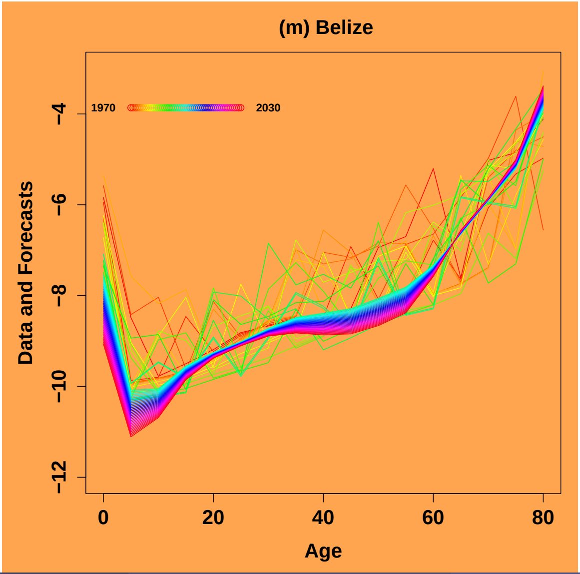
# Mortality from Respiratory Infections, males, $\sigma = 1.51$

## Smoothing over Age Groups



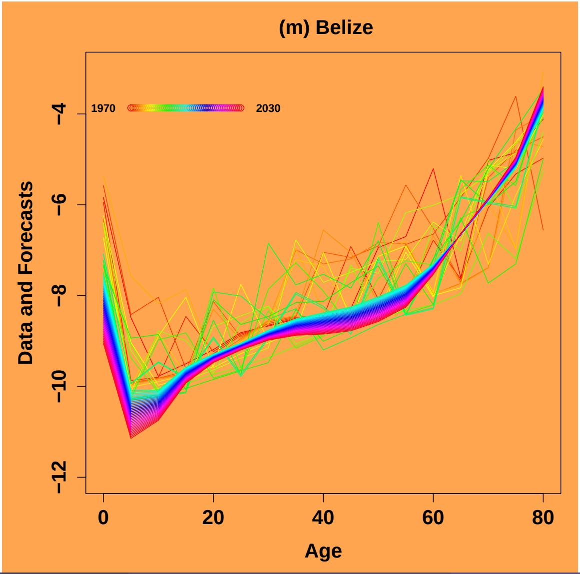
# Mortality from Respiratory Infections, males, $\sigma = 1.15$

## Smoothing over Age Groups



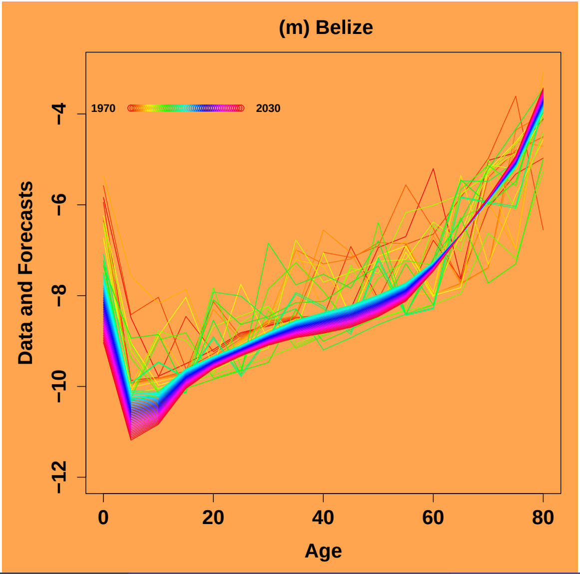
# Mortality from Respiratory Infections, males, $\sigma = 0.87$

## Smoothing over Age Groups



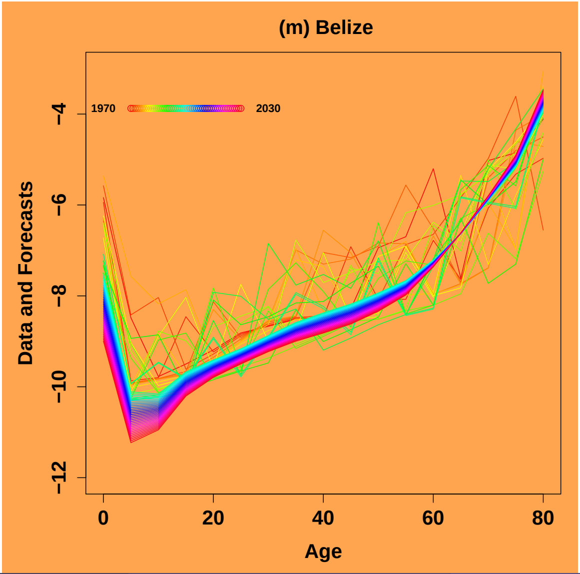
# Mortality from Respiratory Infections, males, $\sigma = 0.66$

## Smoothing over Age Groups



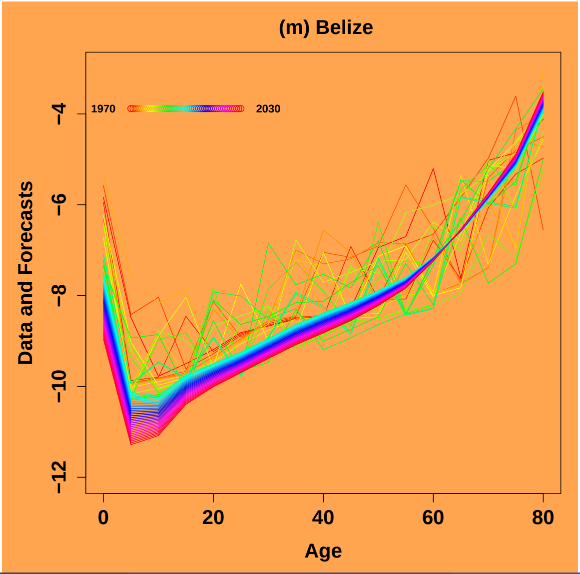
# Mortality from Respiratory Infections, males, $\sigma = 0.50$

## Smoothing over Age Groups



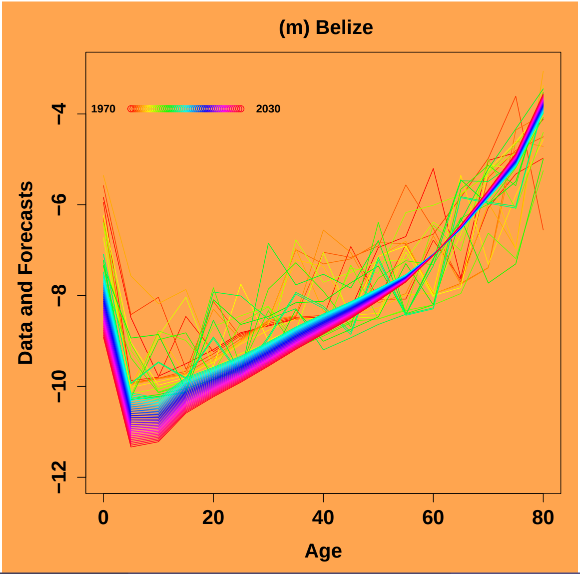
# Mortality from Respiratory Infections, males, $\sigma = 0.38$

## Smoothing over Age Groups



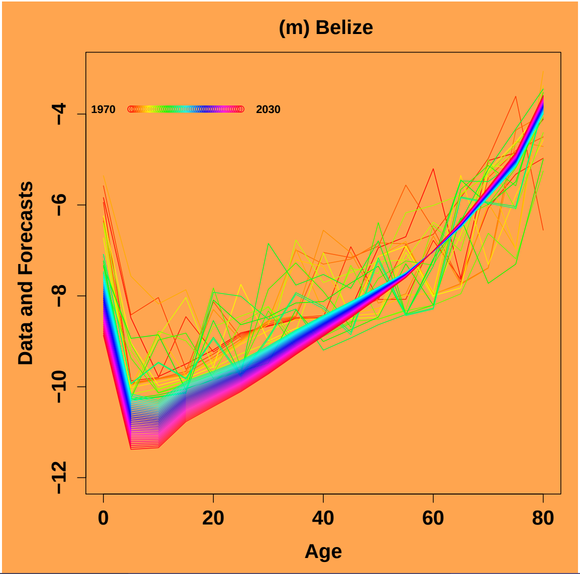
# Mortality from Respiratory Infections, males, $\sigma = 0.28$

## Smoothing over Age Groups



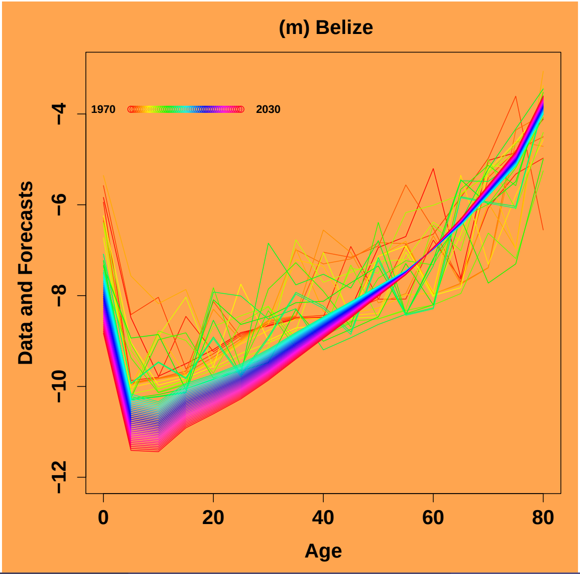
# Mortality from Respiratory Infections, males, $\sigma = 0.21$

## Smoothing over Age Groups



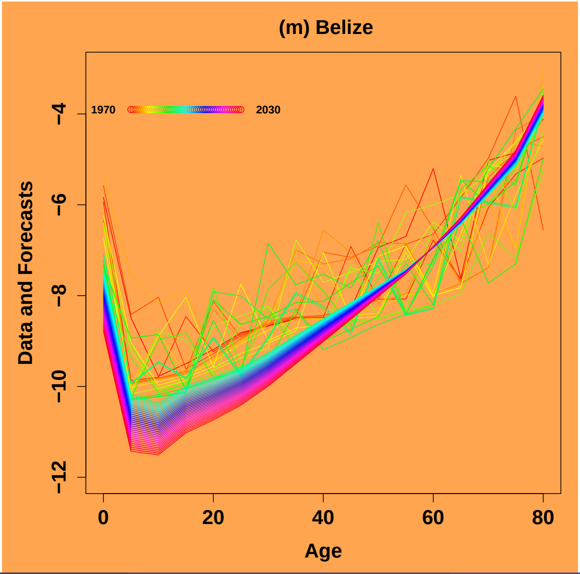
# Mortality from Respiratory Infections, males, $\sigma = 0.16$

## Smoothing over Age Groups



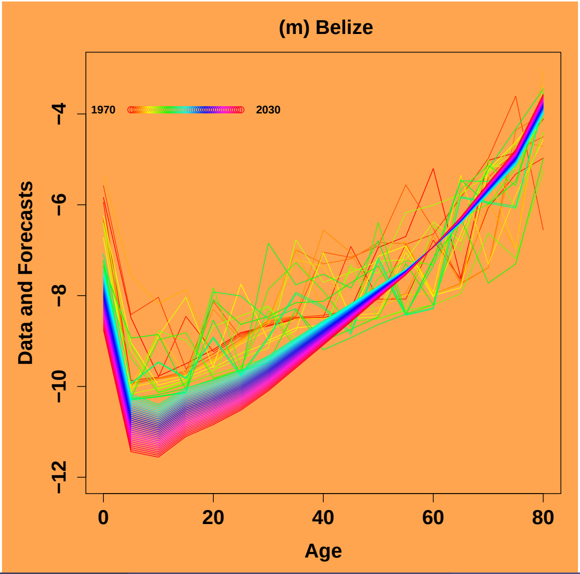
# Mortality from Respiratory Infections, males, $\sigma = 0.12$

## Smoothing over Age Groups



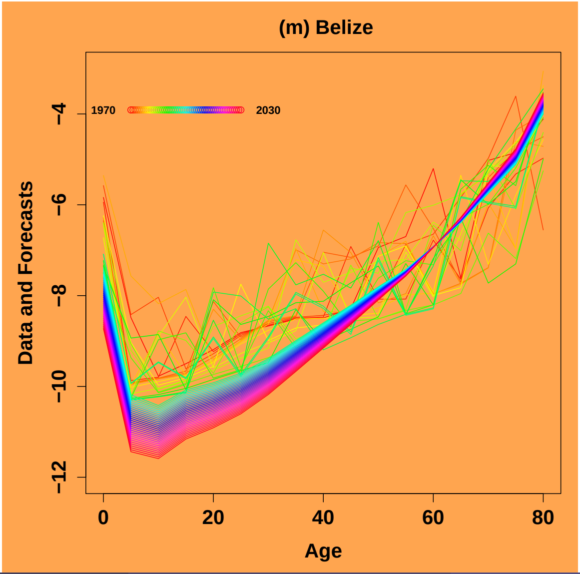
# Mortality from Respiratory Infections, males, $\sigma = 0.09$

## Smoothing over Age Groups



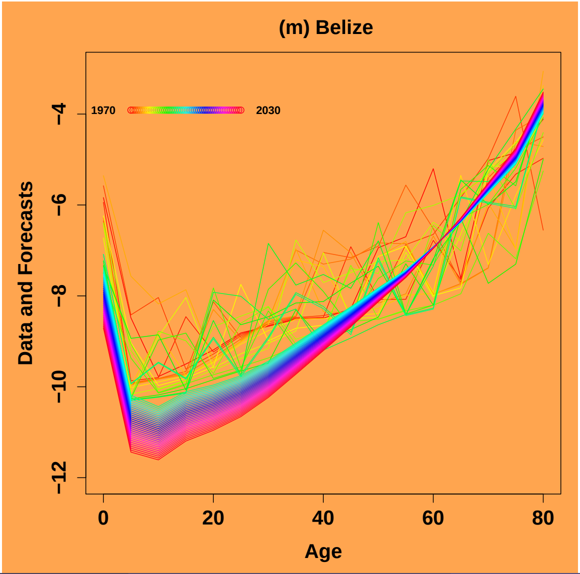
# Mortality from Respiratory Infections, males, $\sigma = 0.07$

Smoothing over Age Groups



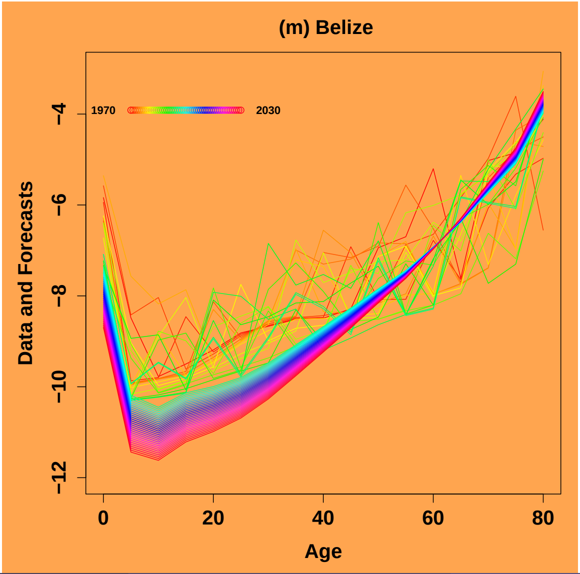
# Mortality from Respiratory Infections, males, $\sigma = 0.05$

## Smoothing over Age Groups



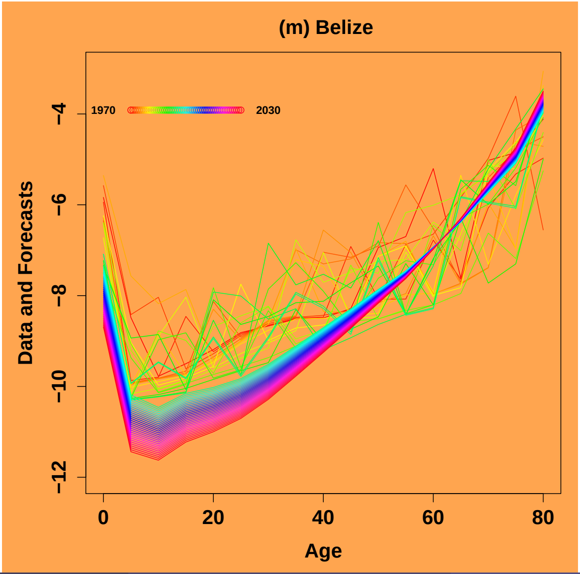
# Mortality from Respiratory Infections, males, $\sigma = 0.04$

## Smoothing over Age Groups



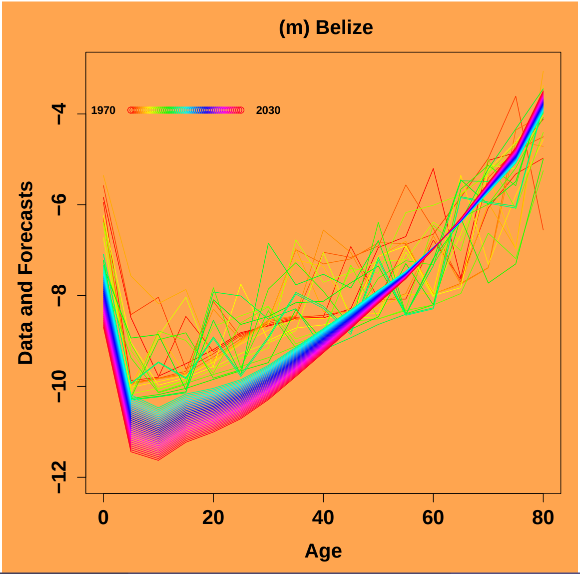
# Mortality from Respiratory Infections, males, $\sigma = 0.03$

## Smoothing over Age Groups



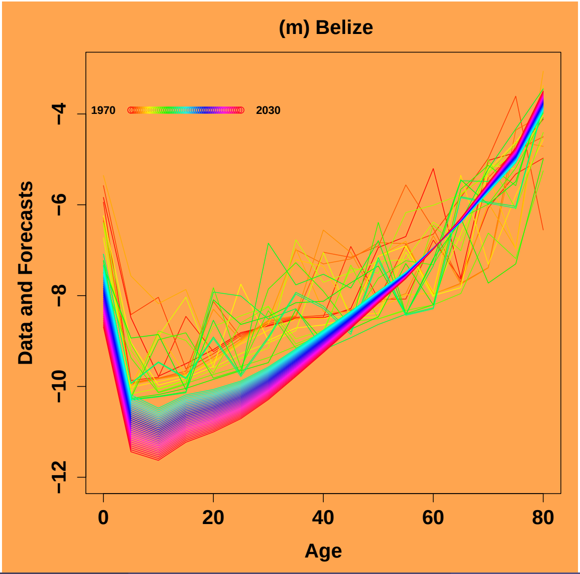
# Mortality from Respiratory Infections, males, $\sigma = 0.02$

## Smoothing over Age Groups



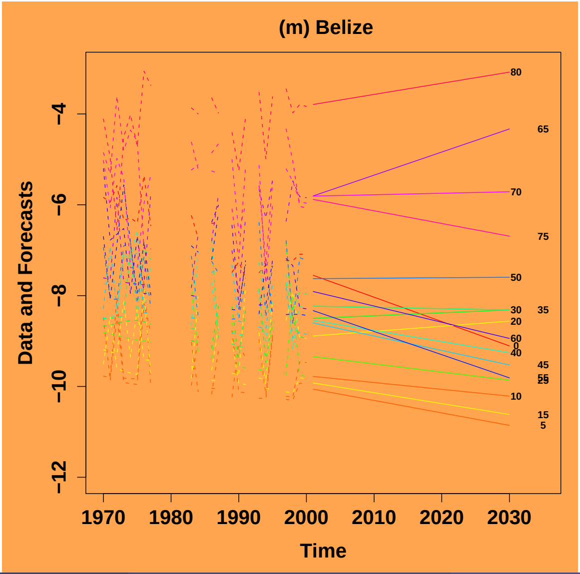
# Mortality from Respiratory Infections, males, $\sigma = 0.01$

## Smoothing over Age Groups



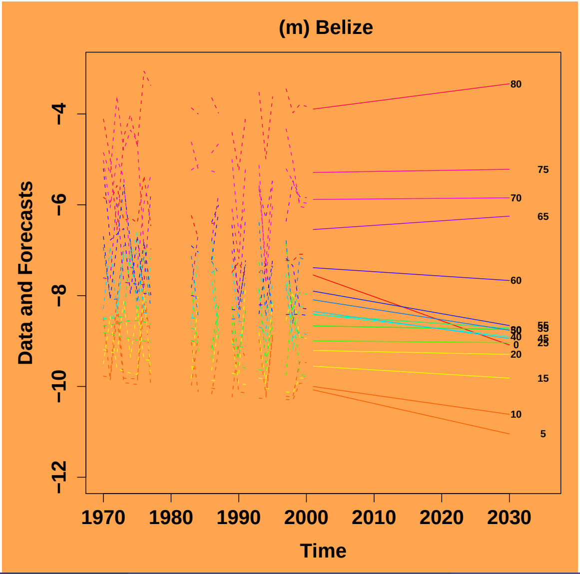
# Mortality from Respiratory Infections, males

## Least Squares



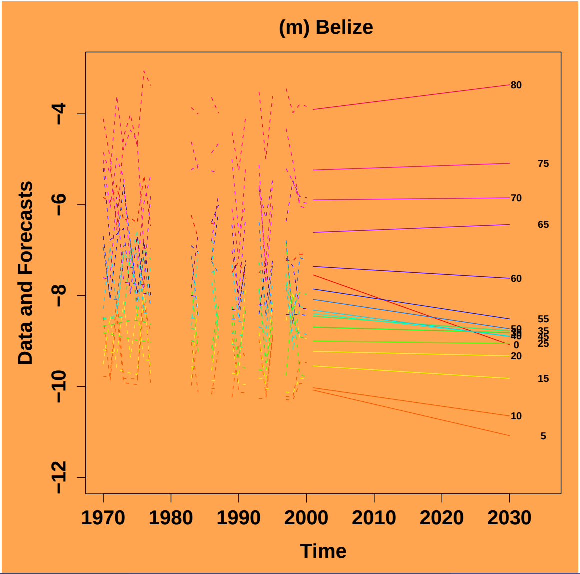
# Mortality from Respiratory Infections, males, $\sigma = 2.00$

## Smoothing over Age Groups



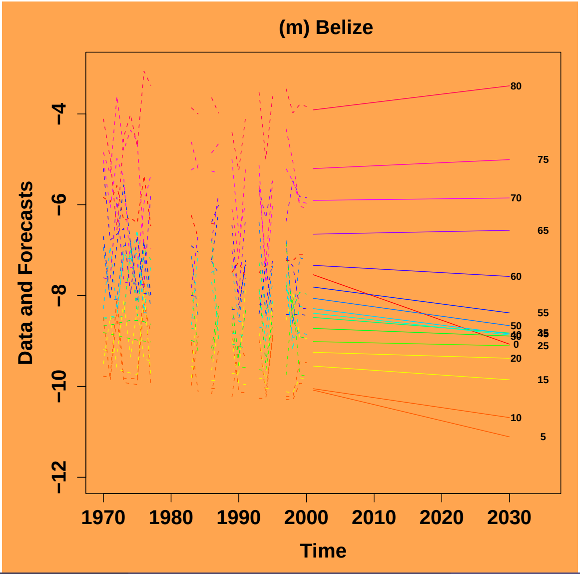
# Mortality from Respiratory Infections, males, $\sigma = 1.51$

## Smoothing over Age Groups



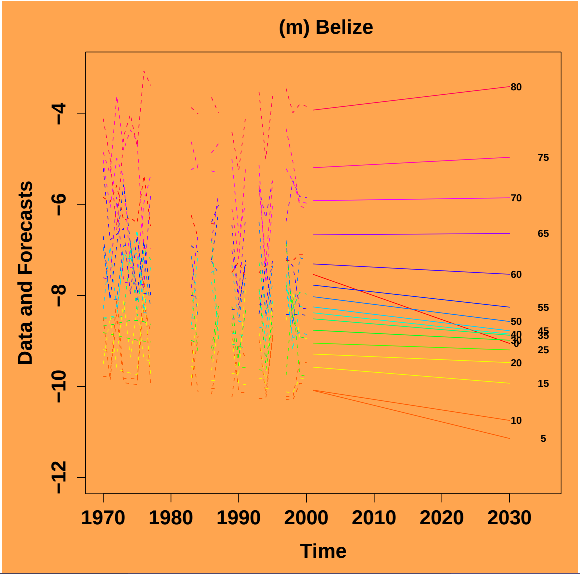
# Mortality from Respiratory Infections, males, $\sigma = 1.15$

## Smoothing over Age Groups



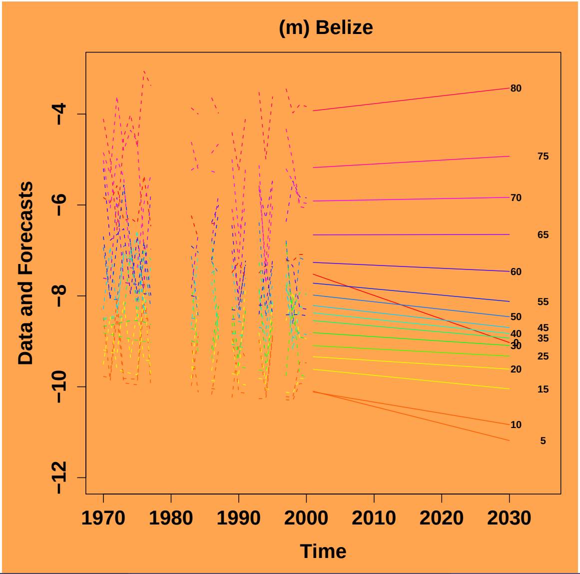
# Mortality from Respiratory Infections, males, $\sigma = 0.87$

## Smoothing over Age Groups



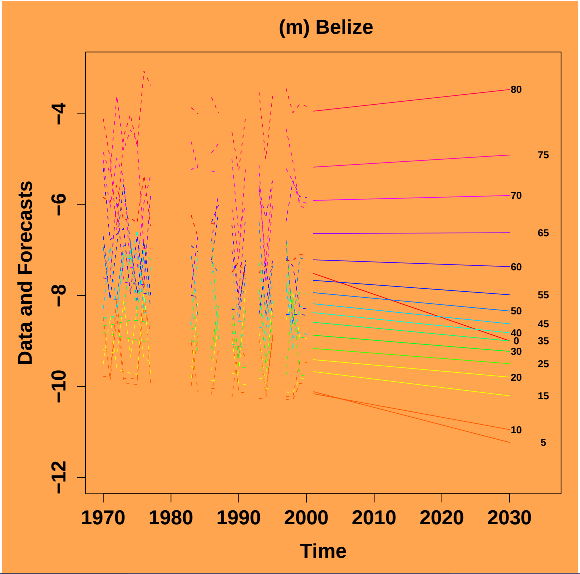
# Mortality from Respiratory Infections, males, $\sigma = 0.66$

## Smoothing over Age Groups



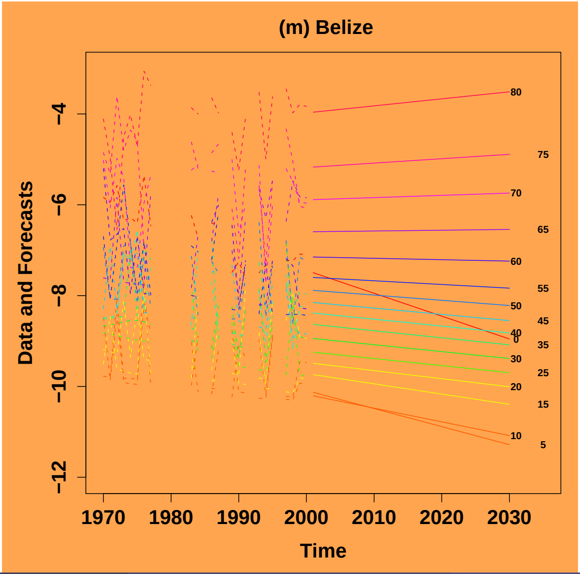
# Mortality from Respiratory Infections, males, $\sigma = 0.50$

## Smoothing over Age Groups



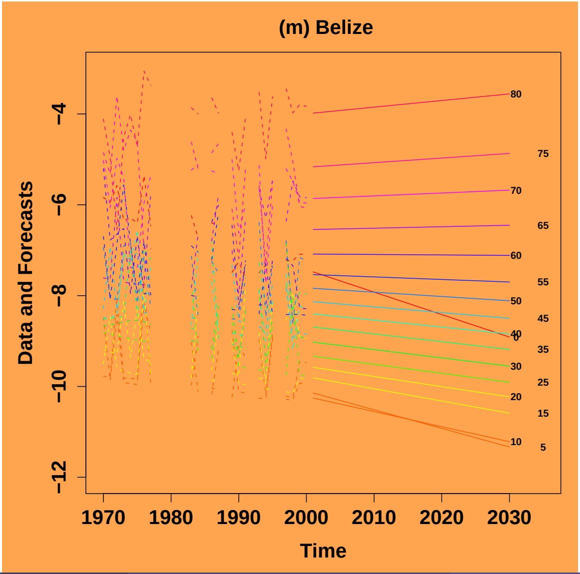
# Mortality from Respiratory Infections, males, $\sigma = 0.38$

## Smoothing over Age Groups



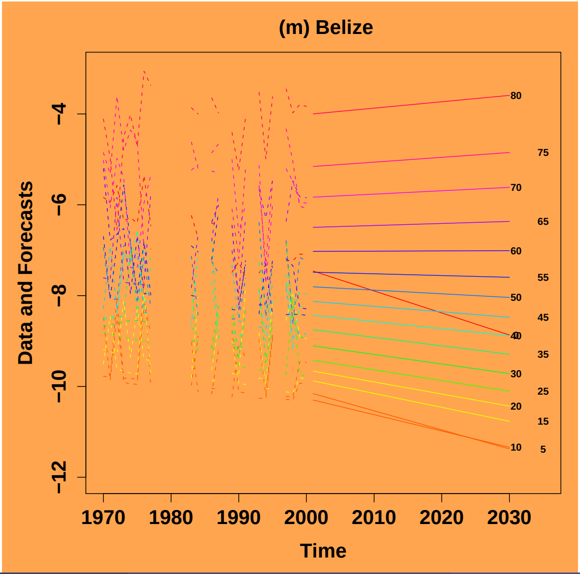
# Mortality from Respiratory Infections, males, $\sigma = 0.28$

## Smoothing over Age Groups



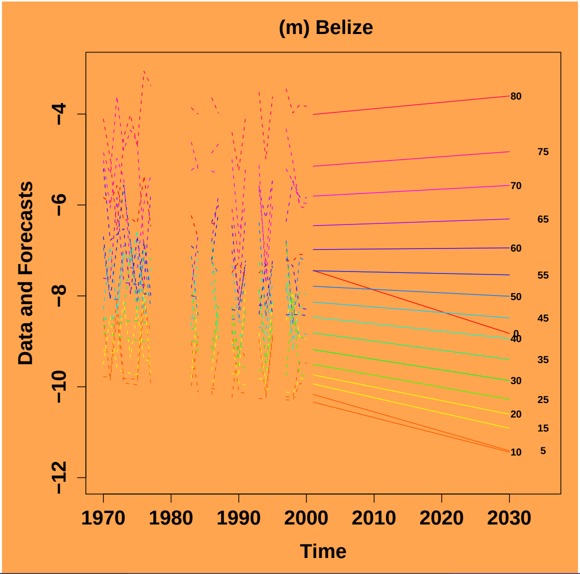
# Mortality from Respiratory Infections, males, $\sigma = 0.21$

## Smoothing over Age Groups



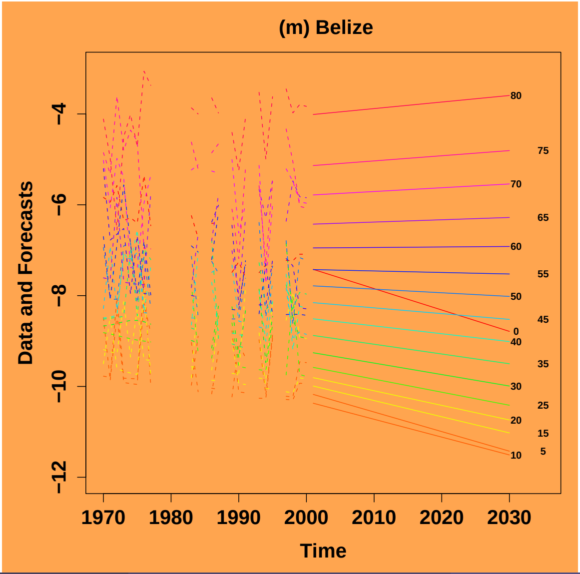
# Mortality from Respiratory Infections, males, $\sigma = 0.16$

## Smoothing over Age Groups



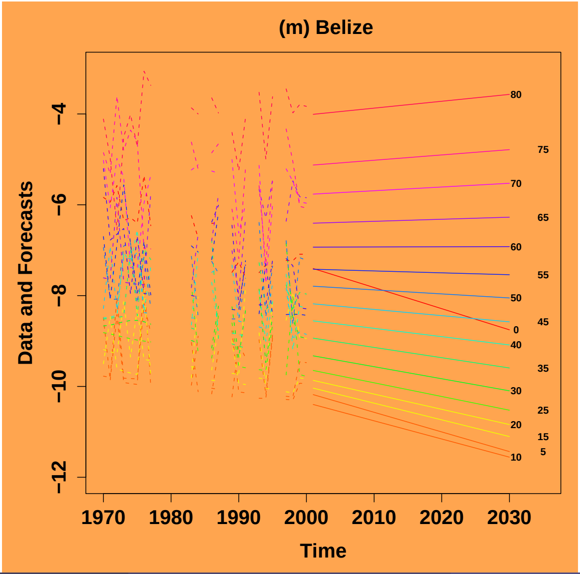
# Mortality from Respiratory Infections, males, $\sigma = 0.12$

## Smoothing over Age Groups



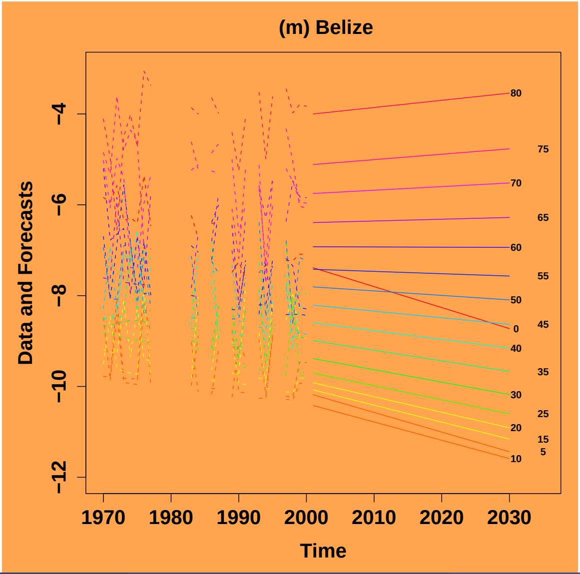
# Mortality from Respiratory Infections, males, $\sigma = 0.09$

## Smoothing over Age Groups



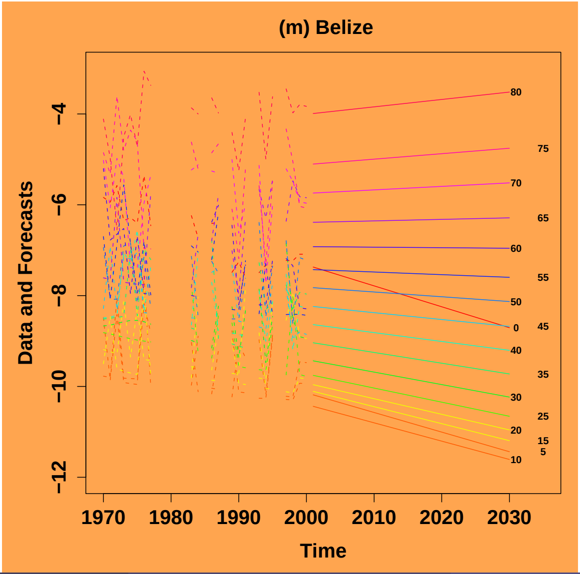
# Mortality from Respiratory Infections, males, $\sigma = 0.07$

## Smoothing over Age Groups



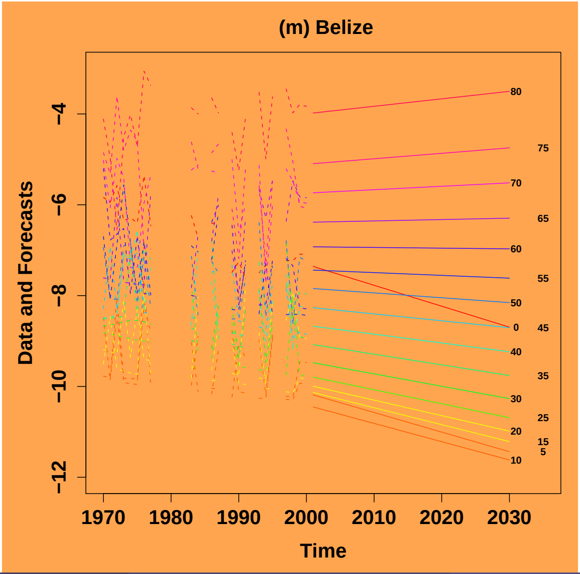
# Mortality from Respiratory Infections, males, $\sigma = 0.05$

## Smoothing over Age Groups



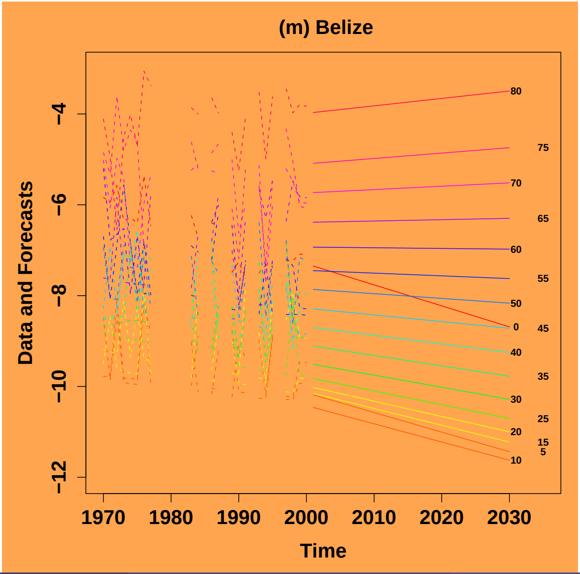
# Mortality from Respiratory Infections, males, $\sigma = 0.04$

## Smoothing over Age Groups



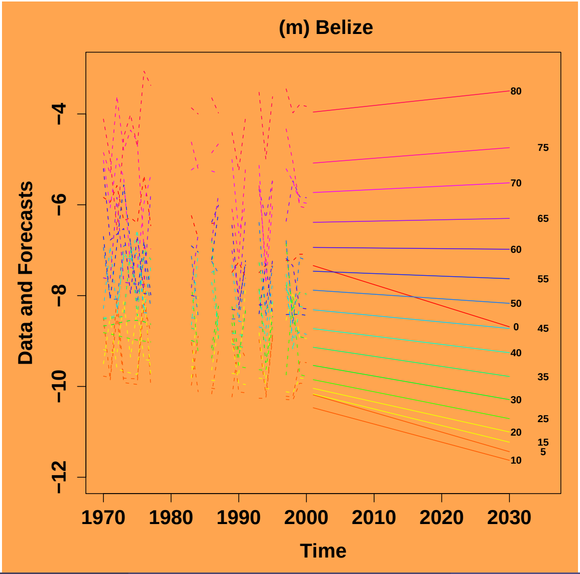
# Mortality from Respiratory Infections, males, $\sigma = 0.03$

## Smoothing over Age Groups



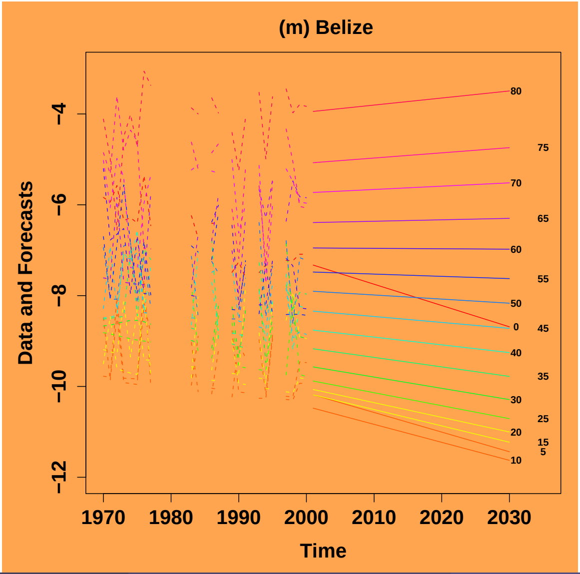
# Mortality from Respiratory Infections, males, $\sigma = 0.02$

## Smoothing over Age Groups



# Mortality from Respiratory Infections, males, $\sigma = 0.01$

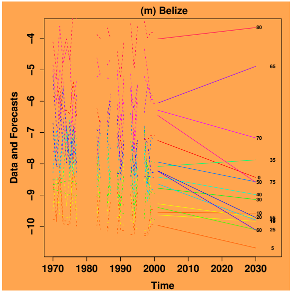
## Smoothing over Age Groups

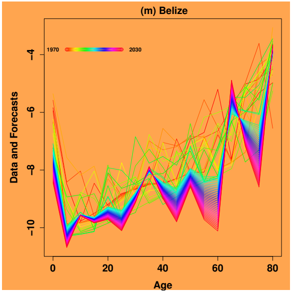


# Smoothing Trends over Age Groups

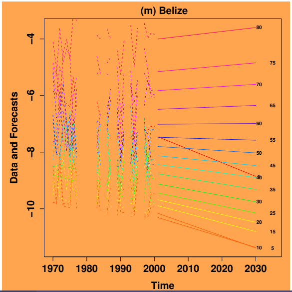
Log-mortality in Belize males from respiratory infections

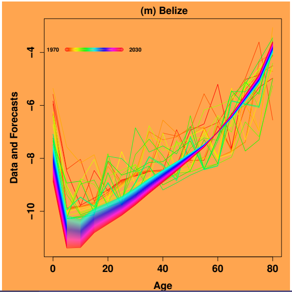
Least Squares




Smoothing
Age Groups

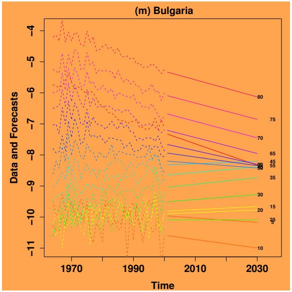


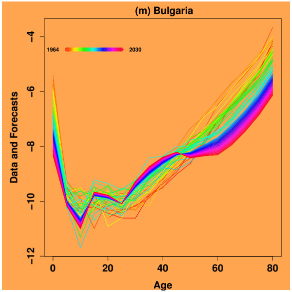



# Smoothing Trends over Age Groups and Time

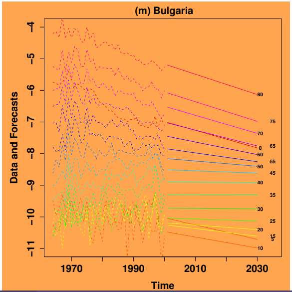
## Log-Mortality in Bulgarian males from respiratory infections

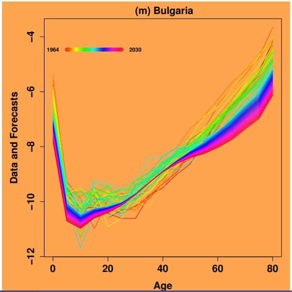
Least Squares




Smoothing
Age and Time

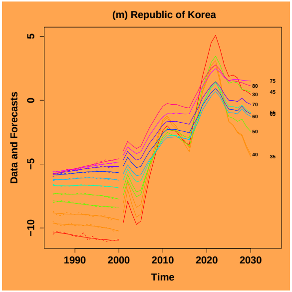


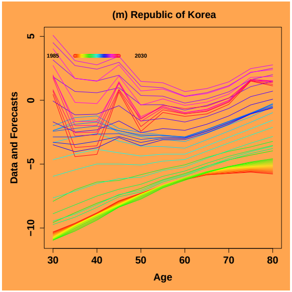



# Using Covariates (GDP, tobacco, trend, log trend)

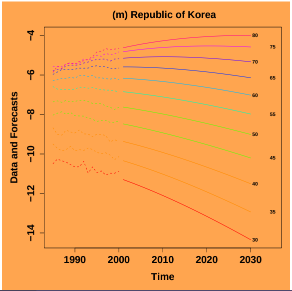
## Lung cancer in Korean Males

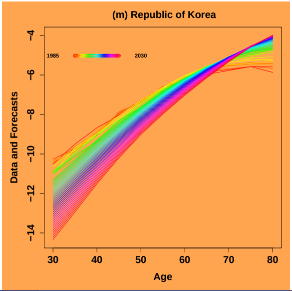
Least Squares




Smooth over age, time, age/time

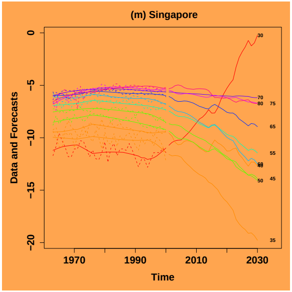


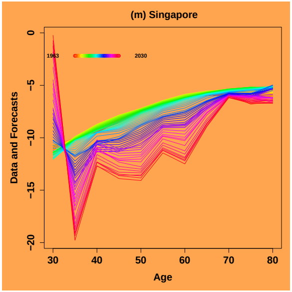



# Using Covariates (GDP, tobacco, trend, log trend)

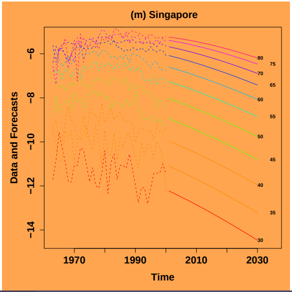
## Lung cancer in Males, Singapore

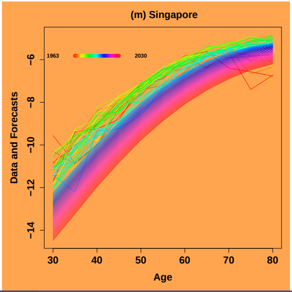
Least Squares



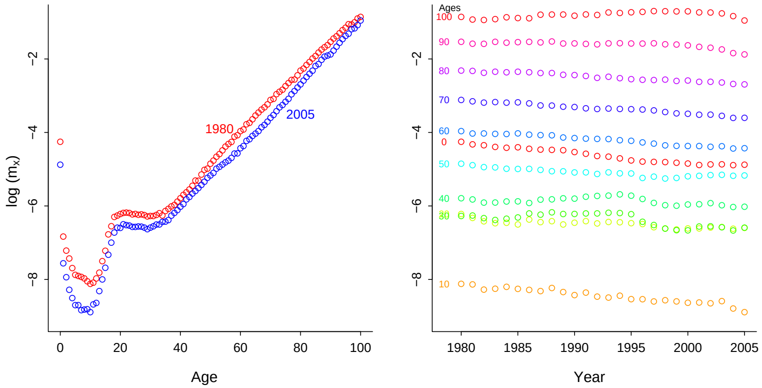

Smooth over age, time, age/time



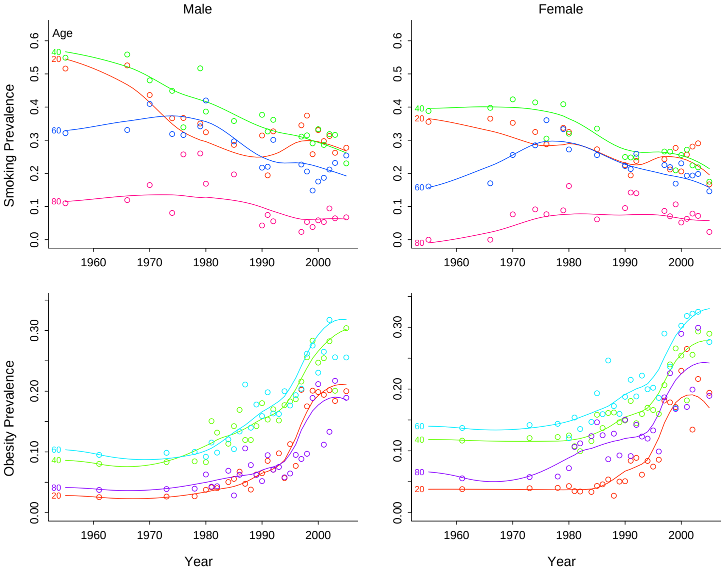



# U.S. Male Mortality over Age and Time

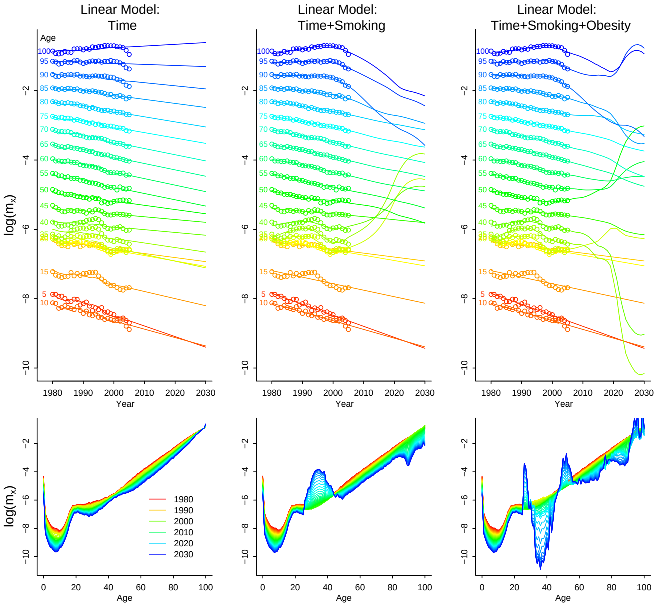
## Demographic Facts: Smoothness in both dimensions



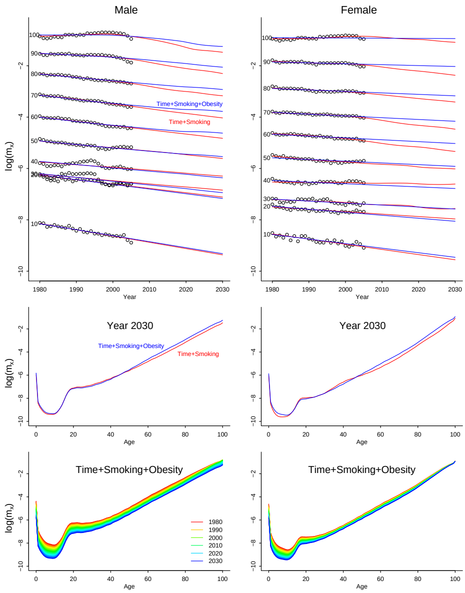
# Biological Risk Factors in U.S. Data



# Linear Models: Biology vs. Demography

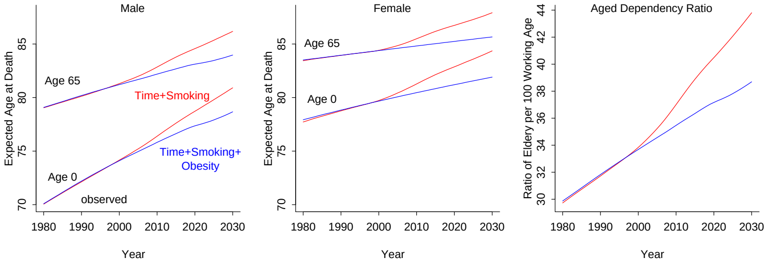


# Forecasts: Biology *and* Demography




- Forecasts retain smoothness over age and time
- After age 50, age-specific mortality increases when adding obesity.
- 2030 forecast for 70-year-olds (per 100,000PYs). Males: 2,290 deaths with obesity; 1,775 without; Females: 1,558 with, 1,144 without.
- At ages over 90, model forecasts converge faster for females than males.

# Life Expectancy and Aged Dependency Ratios




- Male Life Expectancy (±25 years)
  - Past: +5.1 years ⇝ 75.2
  - Future: +5.7 years ⇝ 80.9 (excluding obesity)
  - Future: +3.5 years ⇝ 78.7 (a 35% drop)
- Female Life Expectancy (±25 years)
  - Past: +2.7 years ⇝ 80.4
  - Future: +3.9 years ⇝ 84.4 (excluding obesity)
  - Future: +1.8 years ⇝ 81.9 (a 53% drop)
- ⇝ 1/2 trillion dollar difference for Social Security

# For papers, software, etc.

http://GKing.Harvard.edu