Demographic Microsimulation and Long-Range Regional Population Forecasting

Dmitry Messen and Himanshu Joshi

Houston-Galveston Area Council

www.h-gac.com

Introductions

- Houston-Galveston Area Council
  - Regional Planning Agency for the Houston Area
  - Metropolitan Planning Organization
    - 8-County Region: 5.7 million people
- Long-range (30+ years) Transportation Planning
  - Population, employment, and land use forecasting
  - Modeling of future travel flows
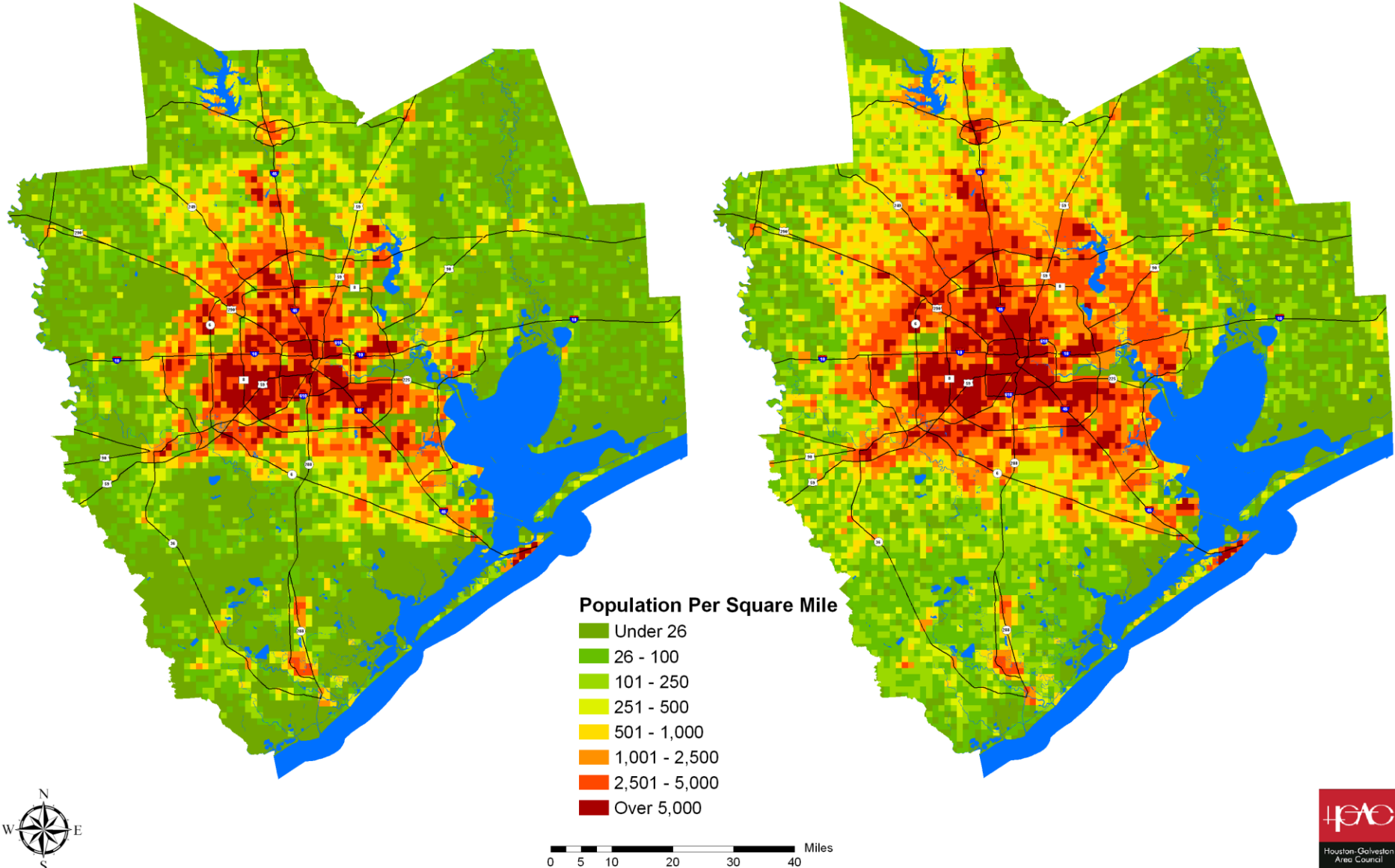 - Socioeconomic Modeling Group
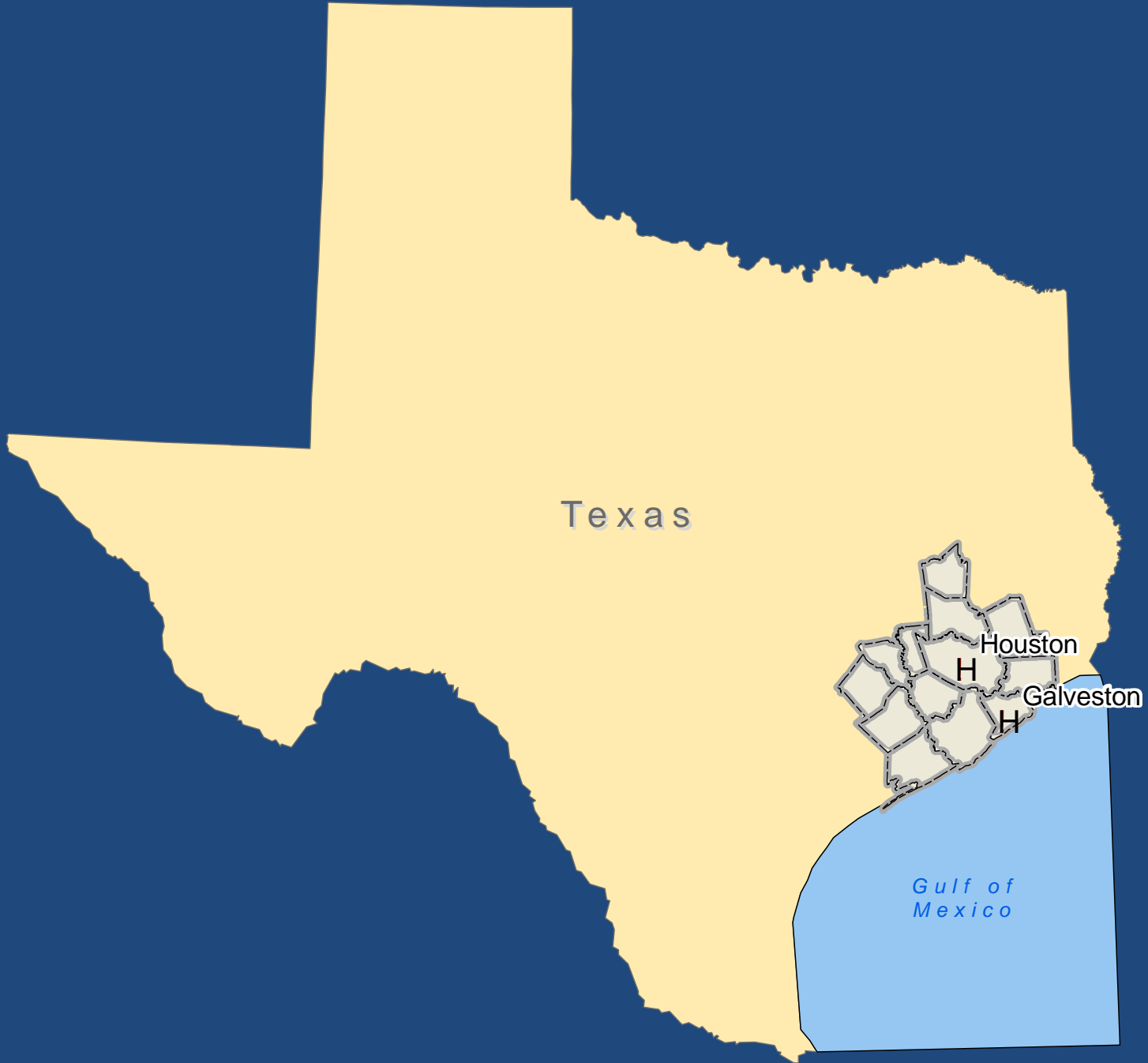
Texas

Houston

H

Galveston

H

Gulf of
Mexico

Population Density

2005

2035

Population Per Square Mile

- Under 26
- 26 - 100
- 101 - 250
- 251 - 500
- 501 - 1,000
- 1,001 - 2,500
- 2,501 - 5,000
- Over 5,000

Miles
0   5   10   20   30   40

Population Modeling

- Suite of Models
  - Macro-simulation (standard deterministic aggregate cohort-component)
  - Micro-simulation (disaggregated probabilistic)
  - Same data platform
    - Person-level records synthesized from aggregate data
- Segmentation: 824 bins
  - 103 age classes (individual years, 0 through 102)
  - 4 ethnicities (Black, Hispanic, Other, White/Anglo)
  - 2 Sexes

Population Modeling

- Temporal step: 1 year

- Processes
  - Aging
  - Surviving
  - Giving birth
  - Migration
    - Domestic IN (from another region)
    - Domestic OUT (to another region)
    - International IN (from a foreign country)

Data Sources

- 2007 Base-year population (TXSDC)

- Fertility rates (TXSDC)

- Survival rates (TXSDC)

- Migration rates
  - Constructed from ACS PUMS 2005 and 2007
  - Stratified only by age (5-year age "block" rates)
  - Foreign in-migration rates adjusted downward (to account for latent foreign out-migration)
  - Future rates discounted by 5% each year (to account for the growing population base)
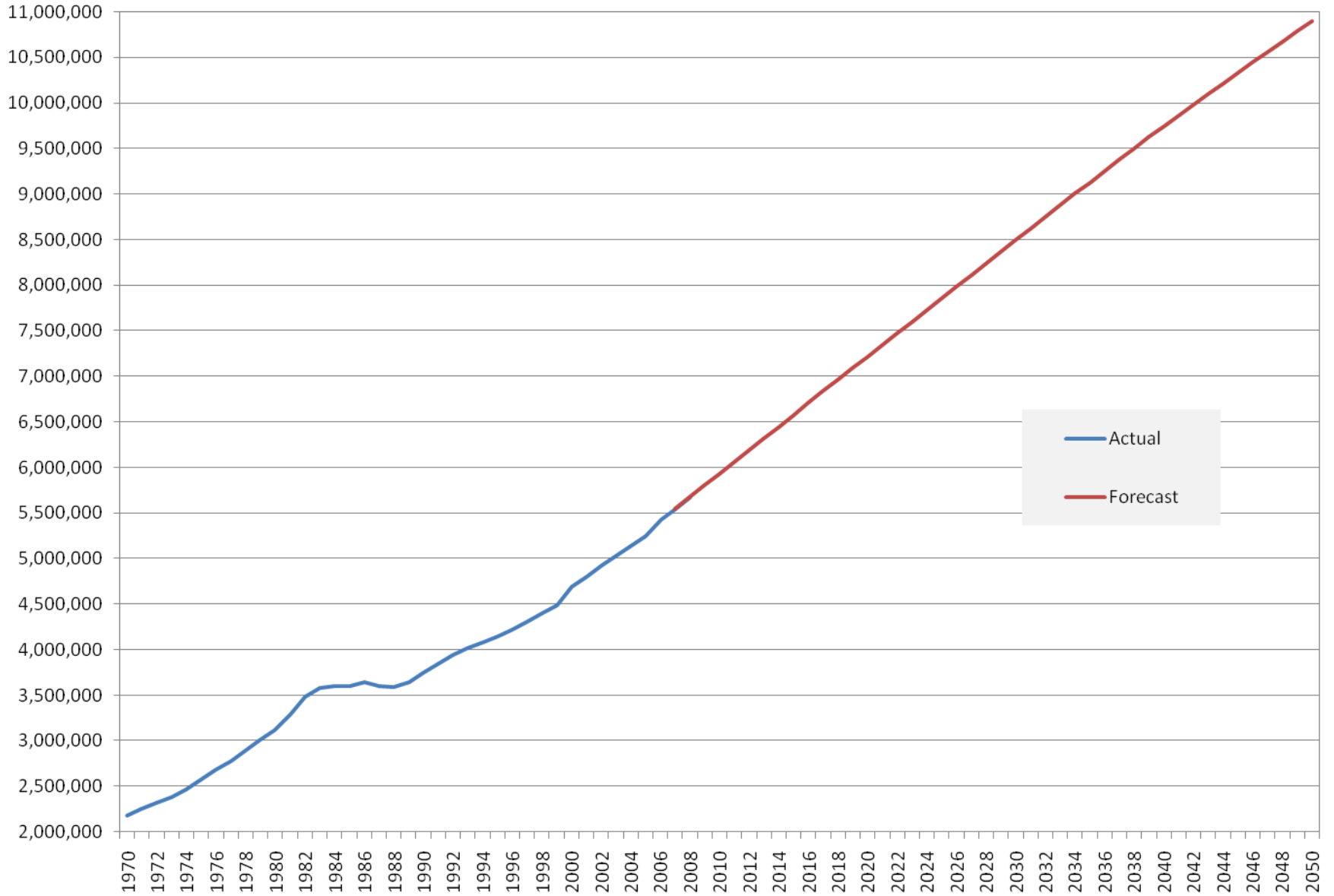
Model Mechanics

A. Base-year population

B. Deaths are generated by applying (1-survival rates) to (A)

C. Survivals are generated by removing (B) from (A)

D. Births are generated for females from (C)

E. Domestic and/or foreign in-migrants are generated by applying the domestic in-migration rate to (A)

F. Domestic out-migrants are generated by applying the domestic out-migration rate to (C)
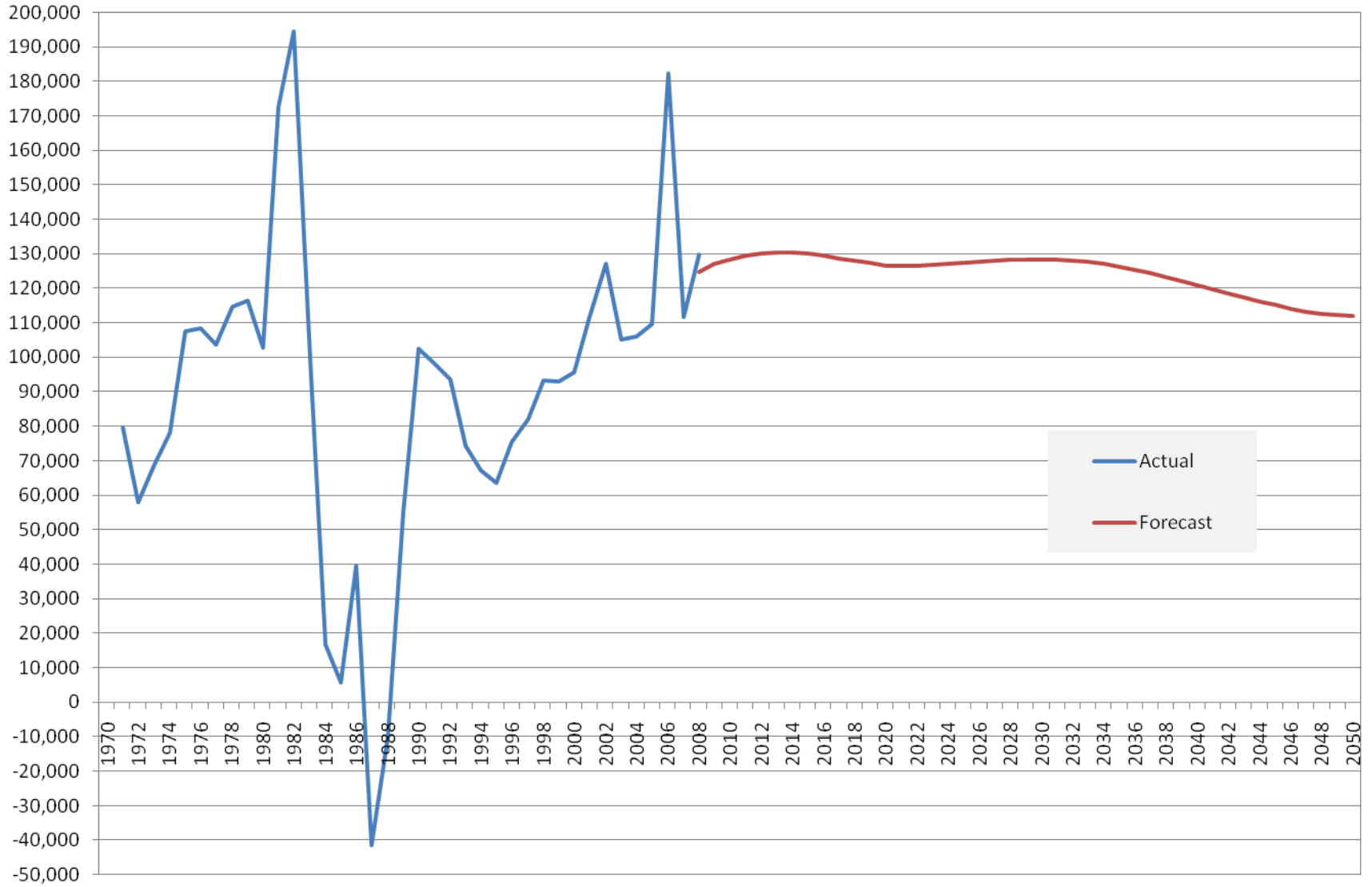
G. Age survived (C) by 1

Models

- Macro-simulation
  - Rates are used as deterministic shares (simple multiplication)
- Micro-simulation
  - Rates are used as event (death, birth, migration) probabilities
  - Monte-Carlo method
    - A randomly drawn number (between 0 and 1) from a uniform distribution is compared with the rate
    - If the rate is, for example, 0.1 and the random number is 0.05 then the event "happens"
- Models Implemented in SAS
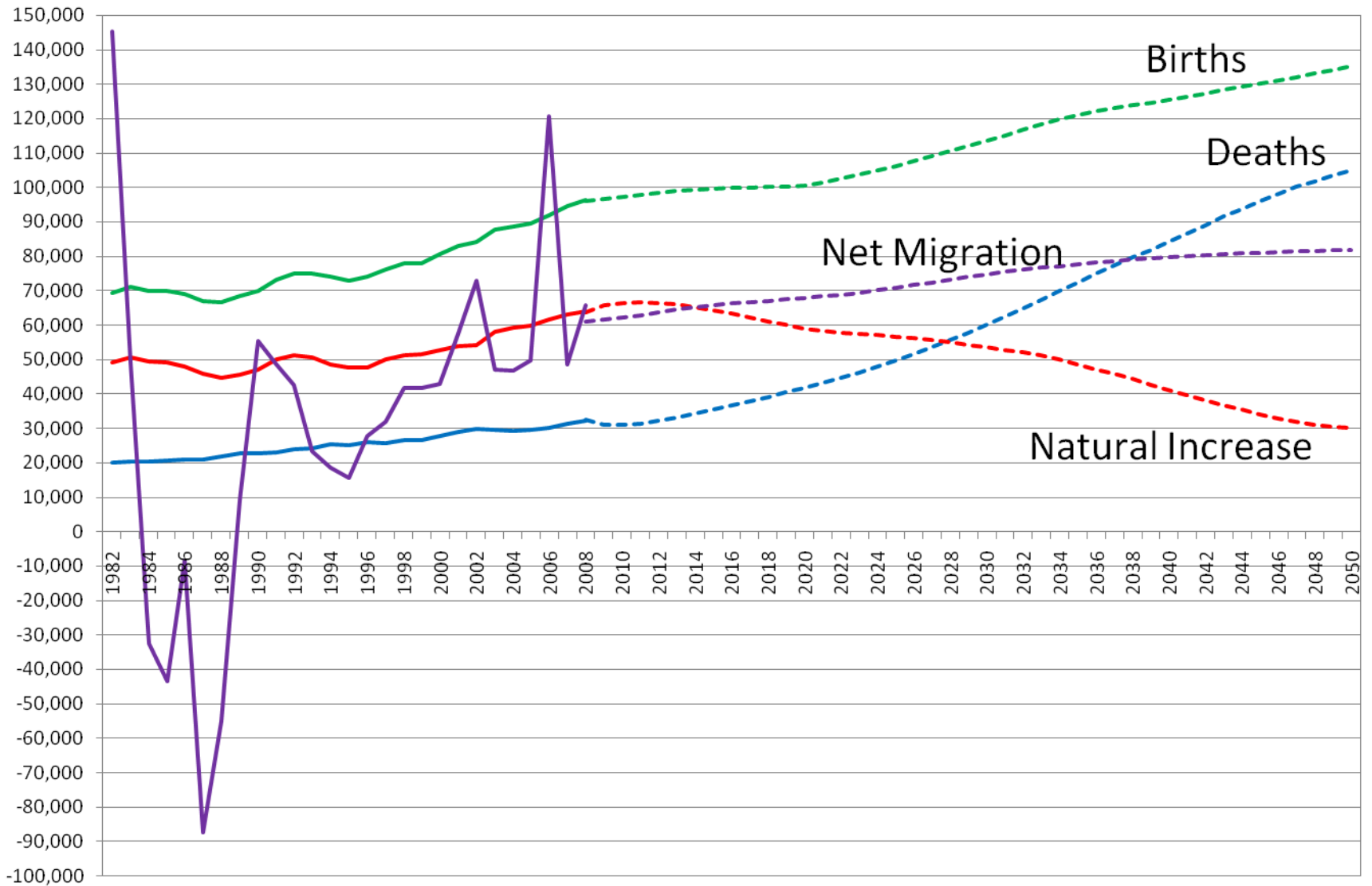 - Few seconds for macro-simulation, 4 hours for micro-simulations

Total Population

Annual Population Change

Components of Population Change

Forecasts Comparison

Legend:
- TXSDC Scenario 1.0 (red)
- TXSDC Scenario 00-07 (orange)
- TXSDC Scenario 00-04 (yellow)
- UH IRF (dark blue)
- Global Insight (green)
- H-GAC (purple)
- Woods&Poole (light blue)
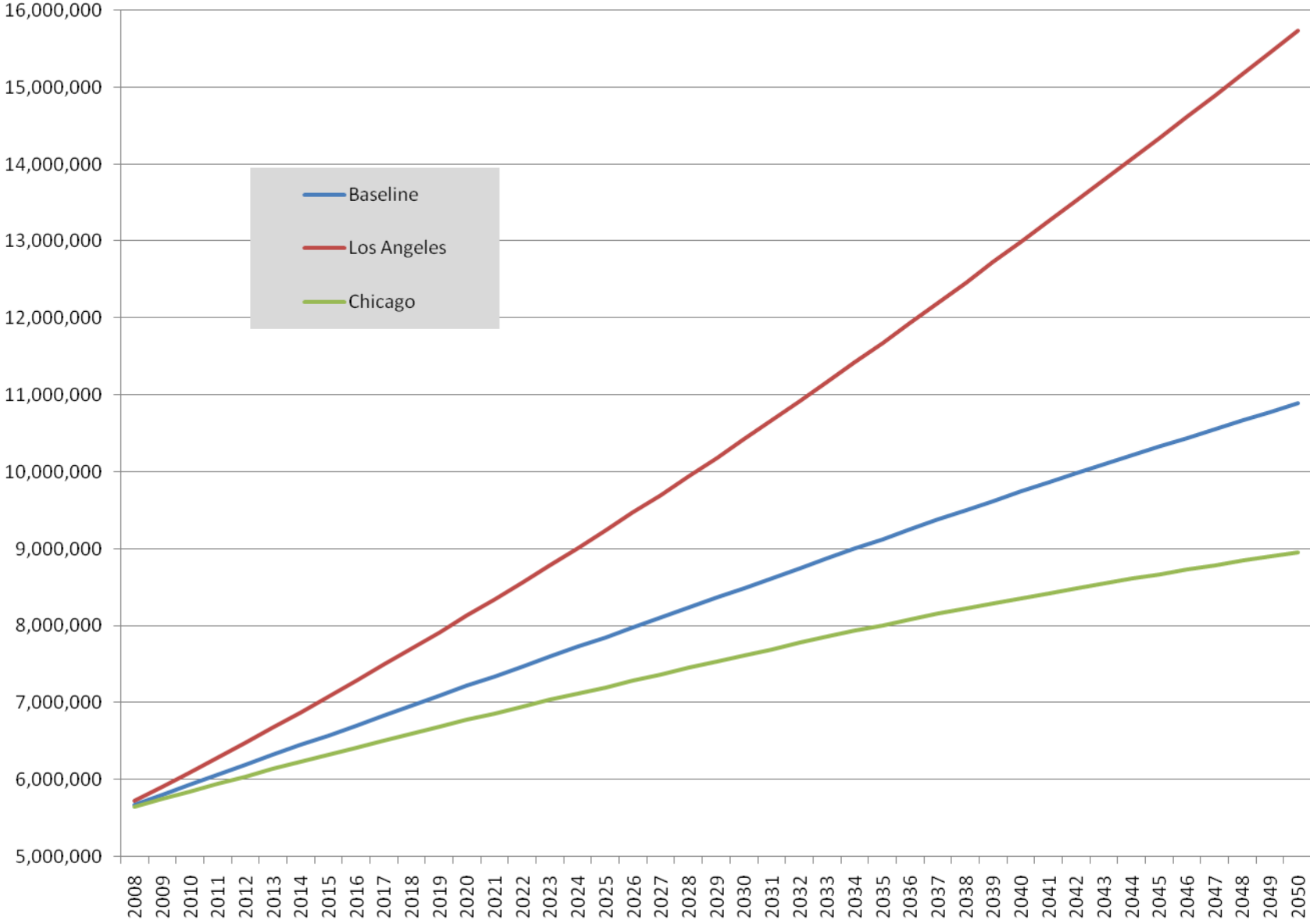
Historical Perspective

- U.S. Regions that ever reached 5.7 million
    - Pre-1950   New York
    - 1950s   Chicago, Los Angeles
    - 1970s   Washington DC
    - 1980s   San Francisco
    - 1990s   Philadelphia
    - 2000s   Boston, Dallas, Houston, Atlanta
- Excluding New York, there are only 3 regions with at least 30 years of growth after they had reached 5.7 million
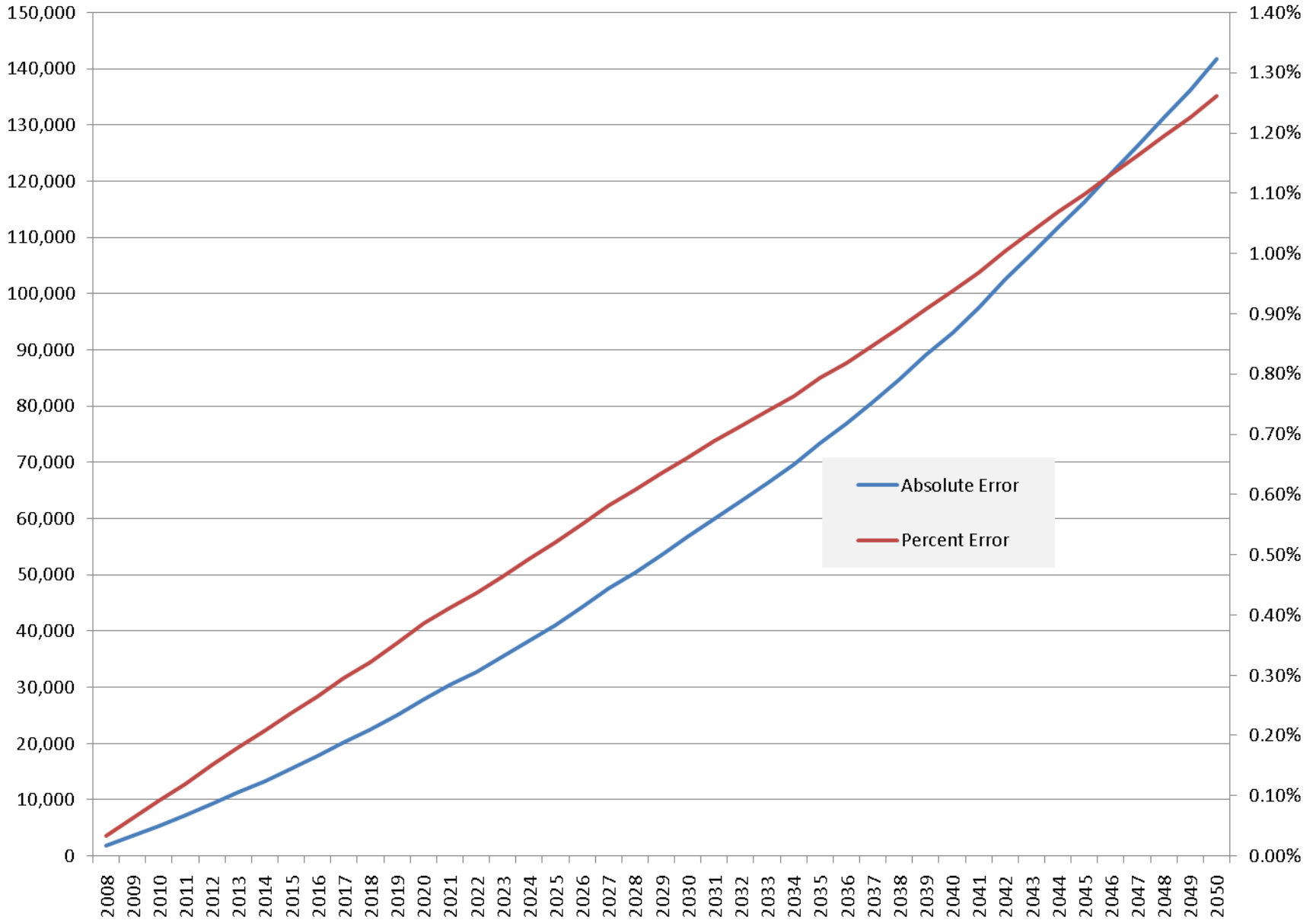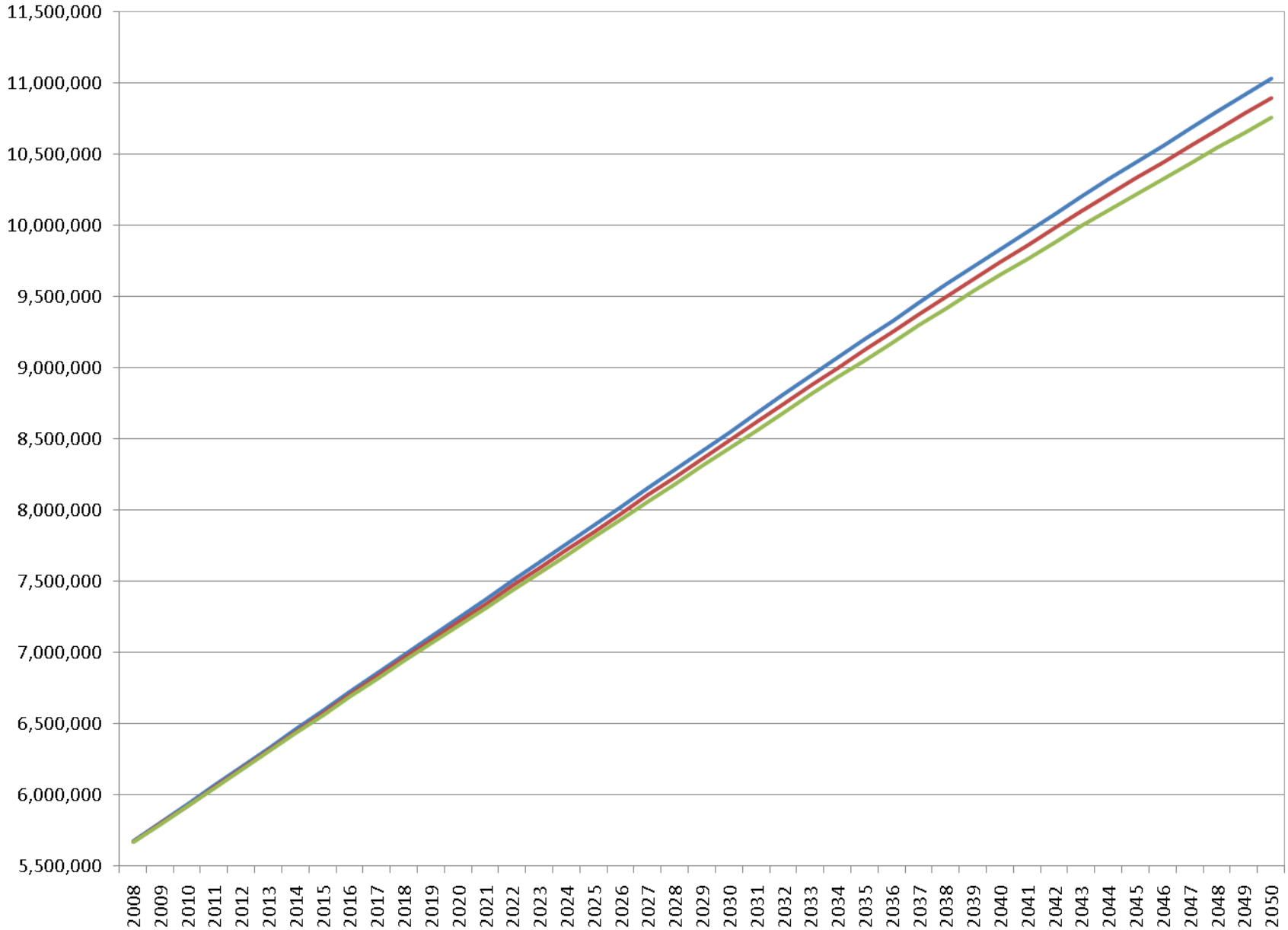
Scenarios

Microsimulation

- Random component + path dependence = each forecast is slightly different

- 100 model runs (14 days); 5.5 million records in the base year

- Compute summary statistics

- Interpret range/2 as "error"
  - Upper and lower bounds

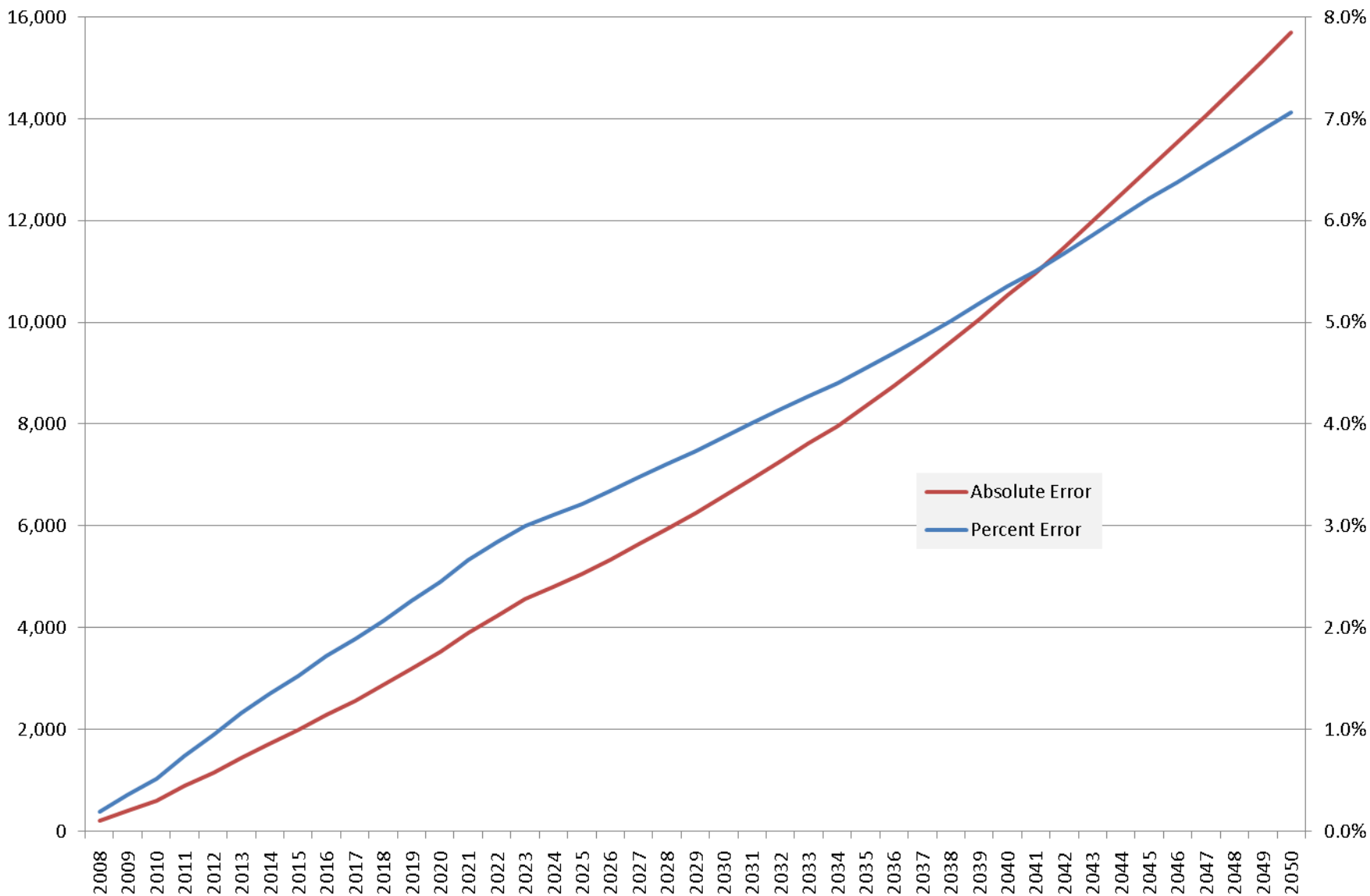Base Year Pop = 5.5 mln; 100 model runs

Microsimulation "Bounds" on the Macrosimulation Forecast

Base Year Size Issue

- Are 100 model runs enough?
- Hypothetical, much smaller region
  - 100,000 pop in the base year
  - Draw a random sample
  - 500 models runs

Base Year Pop = 0.1 mln; 500 model runs

Alternative Virtual Realities

- Jeff Tayman's paper on the accuracy of demographic forecasts
- "Actual" as just one of the possible outcomes
- Human biology and behavior
  - Predictable—only in the sense that we can quantify probabilities of certain outcomes
  - But never predetermined
- Microsimulation and lottery

Future Directions

- Simulation experiments
  - Sensitivity to base pop size and number of runs
- Detailed conditional events probabilities
  - Family dynamics
    - Formation and dissolution (implications for the housing markets)
    - Timing of births
  - Migration
    - Reasons, Methods
    - Role of size and employment growth

Thank you!

Questions?