Dynamic Cohorts

Tom Fiddaman
ISDC, 2017
Contents

- Motivation
- Aging Chains
- Workarounds
- Dynamic cohorts
- Implications & Extensions
World 3

Forecast error @ 45 years:
Population 7%
Population 65+ 44%

Population 0 To 14
- Deaths 0 to 14
- Maturation 14 to 15

Population 15 To 44
- Deaths 15 to 44
- Maturation 44 to 45

Population 45 To 64
- Deaths 45 to 64
- Maturation 64 to 65

Population 65 Plus
- Deaths 65 plus

- Total fertility
- Reproductive lifetime
- Life expectancy
- Population time
- Equilibrium time

World 3 Forecast error @ 45 years:
Population 7%
Population 65+ 44%
What are we looking for?

**Maximize quality**
- Accuracy
- Operational correspondence with policies
- Speed
- Transparency

**Minimize effort**
- Construction
- Initialization
- Calibration
- Reuse
Why disaggregate?

- Components of interest have different dynamics
- A priori aggregation is hard
- Correspondence with measurements
- Representation of policies
State vs. Categorical Representation

- Healthy People
- Sick People
- Really Sick People
- Unhealthy People

Health (Regenerating) vs. Degrading
Aging Chains

• Advantages
  – The obvious approach in a “flat” language (no arrays)
  – Visible
  – Works when you don’t need explicit age interpretation

• Limitations
  – Dispersion
  – Abrupt dynamics
  – Transition time ≠ age difference
  – Lots of work, especially for coflows
Dispersive System

Input Pulse

100th Order Norm SD = 10%

3rd Order

10th Order

widgets/Month

Time (Month)

input: 3rd Order
output: 3rd Order
output: 10th Order
output: 100th Order
Abrupt Age Shifts in Dynamics

Mortality Rate

fraction/year

Age 70  Age 75  Age 80  Age 85

5 Year  Annual
Discrete Time

• Set time step (DT) = cohort duration

+ No dispersion

— No flexibility to represent fast dynamics
— No ability to test for stability
Shifting

- Instead of moving people continuously, move them at discrete times
  
  + No dispersion
  + Flexible time step

—Sawtooth behavior
Continuous Cohortimg

- Use hidden internal states, one per time step (DT)

  + No dispersion

  — Potentially heavy computational burden

Diagramming Aging Chains

Stuff → Aging Out → Aging In → Stuff → Aging Out → Aging In → Stuff

Attribute → Attribute Out → Attribute In → Attribute → Attribute Out → Attribute In → Attribute

Rate → Average Attribute → Loss Rate → Average Attribute → Loss Rate → Average Attribute

Attribute Loss Rate → Attribute Loss Rate → Attribute Loss Rate → Attribute Loss Rate
Agents

- Model individuals
- Age is a state of the person, not a stock of people
  - No dispersion
  - Any nonlinear behavior can be represented
    - Discrete, stochastic behavior
    - Big computational and cognitive burdens

*I have a map of the United States ... actual size.*

*It says, “Scale: 1 mile = 1 mile.” I spent last summer folding it.*

– Steven Wright
Dynamic Cohorts

• Moving things through age categories is hard work and causes dispersion.

• So, don’t move them!

• Instead, maintain a dynamic list of cohorts, with age as a state.
Ingredients

• **Create cohorts “born” within an interval**
  – Rationale: people born at the same time have similar attributes and experiences
  – Not conceptually different from disaggregation by gender, region, vehicle type

• **Accumulate age (or calculate it from the birth date)**

• **Represent internal dynamics of the group**

• **Track until you lose interest**
  – Too few members
  – Age > maximum age of interest

• **Calculate aggregates for feedback to the rest of the model**
Implementation in Ventity

Population
- Births
- Global parameters

Actions
- Create
- Delete

Cohort
- Population
- Age
- Dynamics

Collection of Cohorts
- Aggregates
Cohort Entity
Population Entity
Cohort Life Cycles

Parameters corresponding with Japan in Eberlein & Thompson
## Initialization

### Table 1 Cohort Initialization Data

<table>
<thead>
<tr>
<th>Enabled</th>
<th>Time</th>
<th>CohortID</th>
<th>gender</th>
<th>region</th>
<th>Creation Time</th>
<th>Life Years</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRUE</td>
<td></td>
<td>F A0</td>
<td>F</td>
<td>Japan 1yr</td>
<td>2009</td>
<td></td>
<td>501613</td>
</tr>
<tr>
<td>TRUE</td>
<td></td>
<td>F A1</td>
<td>F</td>
<td>Japan 1yr</td>
<td>2008</td>
<td></td>
<td>512203</td>
</tr>
<tr>
<td>TRUE</td>
<td></td>
<td>F A2</td>
<td>F</td>
<td>Japan 1yr</td>
<td>2007</td>
<td></td>
<td>522909</td>
</tr>
<tr>
<td>TRUE</td>
<td></td>
<td>F A3</td>
<td>F</td>
<td>Japan 1yr</td>
<td>2006</td>
<td></td>
<td>530882</td>
</tr>
<tr>
<td>TRUE</td>
<td></td>
<td>F A4</td>
<td>F</td>
<td>Japan 1yr</td>
<td>2005</td>
<td></td>
<td>536693</td>
</tr>
</tbody>
</table>
Precision

Population over 90, Japan

- Continuous Cohorting – 3200 stocks
- Dynamic Cohorting – 240 stocks
Dynamic Cohort in Comparison

• **Advantages**
  – Low computational burden
  – Clear mapping of agent to group
  – Simple internal dynamics
  – Easy debugging

• **Differences**
  – Data-model matching
  – Initialization in equilibrium

• **Limitations**
  – Nonlinearity in group dynamics
  – Heterogeneity in group members
Projections at Different Resolutions
1, 5, 10 year cohorts

Population

Population over 90

(no significant difference)
Two Ways to Fail

• **Insufficient detail**
  – Cohorts too wide
  – Neglected heterogeneity

• **Omitted dynamics**
  – Age-Period-Cohort effects
  – Time-varying rates
  – Internal dynamics of groups
Example: Infection Model x Cohort

Introduce a disease into an age-disaggregated population, with

- Infection onset proportional to age, and
- Infection-induced mortality
Infected Population
Example: Vehicle Fleet x Model Year

- **Cohort Effect**
  - FleetID
  - Model threshold mileage for reliability
  - Mileage Scrapping
  - Driving

- **Age Effect**
  - Model Year
  - Nominal Age
  - effect of age on condition
  - Model threshold age for reliability

- **Period Effect**
  - Period Effect
  - effect of mileage on condition
  - mileage per vehicle
  - vehicle condition index

- **Experience Effect**
  - Experience Effect
  - loss rate
  - Model loss rate from accidents
  - Model loss rate from failure
Vehicle Utilization Effect on Scrap Rate

- High Driving demand
- Subsequent retirements of high-mileage vehicles
- Remaining fleet is young

Loss Rate
Fraction/Yr

Time
2000 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100
0 0.01 0.02 0.03 0.04 0.05 0.06 0.07 0.08
Applications

• **Living things**
  – People – health, education
  – Fish

• **Perishables**
  – Pharmaceuticals
  – Food

• **Capital**
  – Vehicles
  – Buildings

• **Services**
  – Loans
  – Bonds
  – Contracts
Bottom Line 1

• Cohorts are everywhere
• Cohorts are just a special case of aggregation questions we face every time we model
• We need tools that make it easy to quickly build and test alternative specifications
• Managers have always had an appetite for tactical detail

• Big data availability is growing

• We need ways to fill the appetite and exploit the data *without losing the essential insights of feedback and accumulation.*
THANK YOU
Cohort Duration Aggregation

\[ \tau \]

\[ (2 + \delta \tau) \tau \]
Age-Period-Cohort Effects

• Vital Rate = F(Age, Period, Cohort)
• E.g., Mortality Rate = F( Age, Current Time, Birth Year )
  – Age: aging process effect on mortality
  – Current time: medical technology and risk
  – Birth year: cohort has common experience

• Statistical challenge: Age = Current Time – Birth Year

• Dynamic challenge: Cohorts have common dynamics and experience
• Vital Rate = F(Age, Period, Cohort, Experience)