Rethinking Mortality

Using a State-Based *Dynamic Probabilistic Model* Leveraging *National-Scale Health Data*

Pramodya Samarasinghe

Australian National University

1. Introduction

Goal: Improve prediction of annual mortality (q_x)

$$q_x = \Pr(\,death\ before\ age\ x+1 \mid age\ x\,)$$

Why predicting q_x matters: q_x is used in retirement saving systems
If q_x is *underestimated* → people outlive savings; public costs rise.
If q_x is *overestimated* → people lock in low income and leave money unspent.

Gap in predicting q_x
The benchmark, **Australian Life Tables (ALT)**, use age + gender only → hides large health differences → cross-subsidies (where one group subsidises another).

Project Impact
1. Use health data to produce accurate, calibrated q_x for retirees.
2. Show downstream impact on life expectancy (LE), withdrawal plans, and Age Pension reliance.

Why now?

(Old): Pension
- Employer guarantees payment
- Risk is born by the employer

→

(New): Superannuation
- You have a pot of savings
- **THE RISK IS WITH YOU**
- Good estimate of q_x → better planning

2. Dataset & Challenges

Datasets:
- **PLIDA** (mortality and demographic data)
- **MBS** (medical service claims)
- **PBS** (prescription and medication dispensing records).

Access & novelty: Analysis performed ***inside the Australian Bureau of Statistics' DataLab***, providing rare, audited access to ***linked national health and mortality data***.

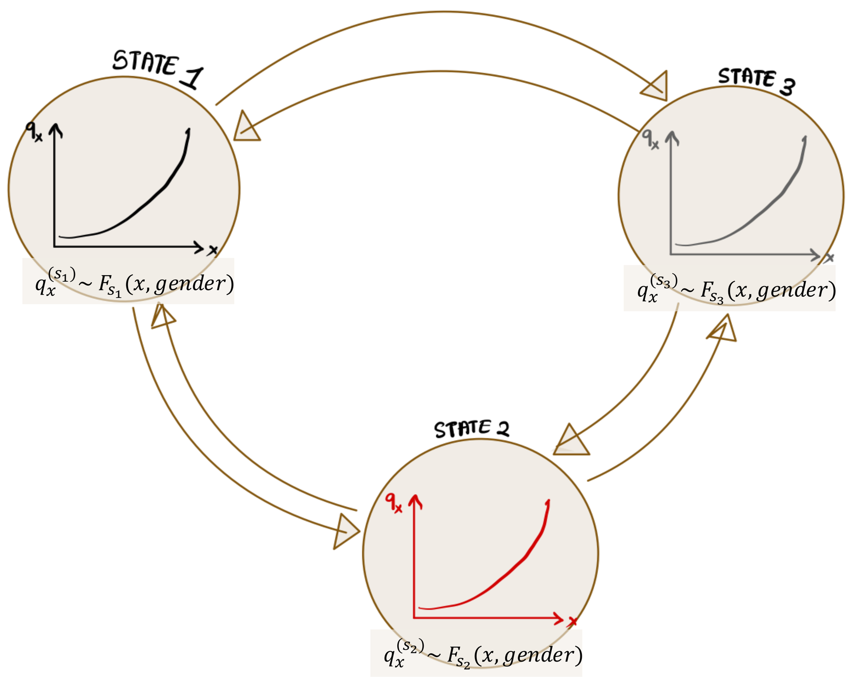
Key challenges
1. **No diagnosis information** → infer **health states** from procedure/medication patterns.
2. **Privacy & deployment** → model at **state/cluster** level (fairness & confidentiality), not per-person.
3. **Short window** → need **age-band pooling** for stable transitions.
4. **Compute & validation** → prioritise **interpretable** model that's valuation-ready.

Executed in a secure, isolated environment with strict software limits, no internet access, and frequent slowdowns from virtualization and security maintenance.

3. Proposed Model

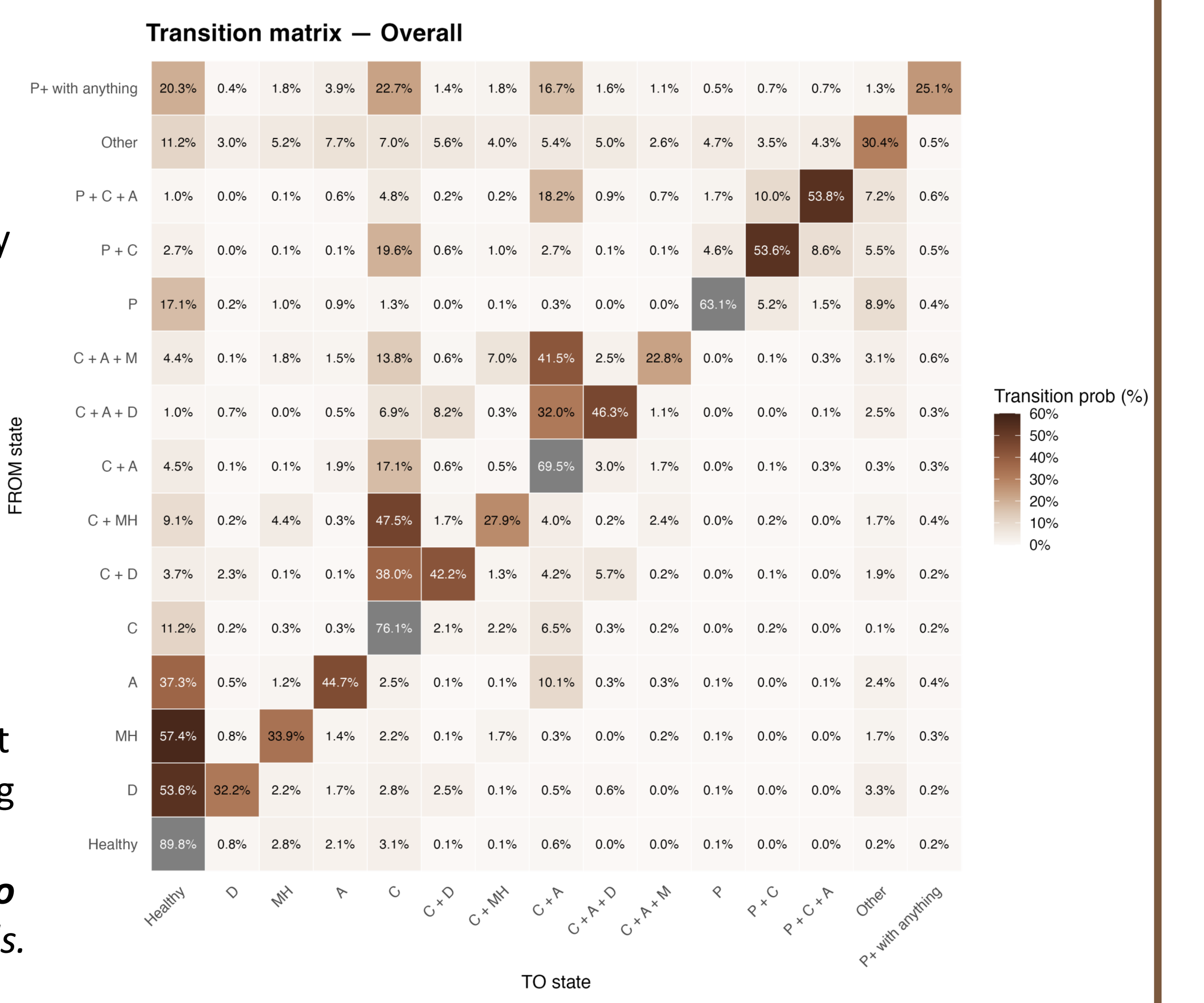
Fitted and identified two key issues with standard models (K-means, Decision Trees, Survival Trees, GLMs): no clear target variable and inability to model dynamic health transitions.

Proposed Markov Chain Model
- Total of **15** most common states.
- Each state denotes a combination of diseases, primarily;
  P: Parkinson's, C: Cardiac, A: Antithrombotic, MH: Mental Health, PM: Pain Medication, D: Diabetes
- Death is not a state; transitions (q_x) are modelled by **age, state**, and **gender**.
  For state (s): $q_x^{(s)} = \dfrac{deaths(\,state \wedge age \wedge gender)}{exposure(state \wedge age \wedge gender)}$
- Accounts for the dynamic nature of the health status of individuals by considering **transition probabilities between states.**
- *Transition probabilities* are calculated overall and split by age bins (gender was not prominent).



Key Assumptions:
- **Death-first annual cycle**: at most one death/transition per year.
- **Markov property**: next year depends only on current health state.
- **Age-banded transitions:** transitions by age bins 50–70/ 70–90/ 90–110.

Key Findings:
- **All models outperform the ALT** evidencing health status is a clear determinant of mortality.
- Transitions indicate that the probability of an individual to ***remain in their current state is highest.***
- ***Mental health conditions*** *are likely to recover* within a year.
- Using the age bin transitions, it is clear that ***younger** ages* have higher likelihood of ***healing*** while ***older*** ones have higher chance in development of ***more diseases.***
- The ALT assumes a static nature, thus it underpredicts life expectancy for young ages and over predicts for old.
- The transitions cause ***life expectancy to deviate significantly*** from static models.

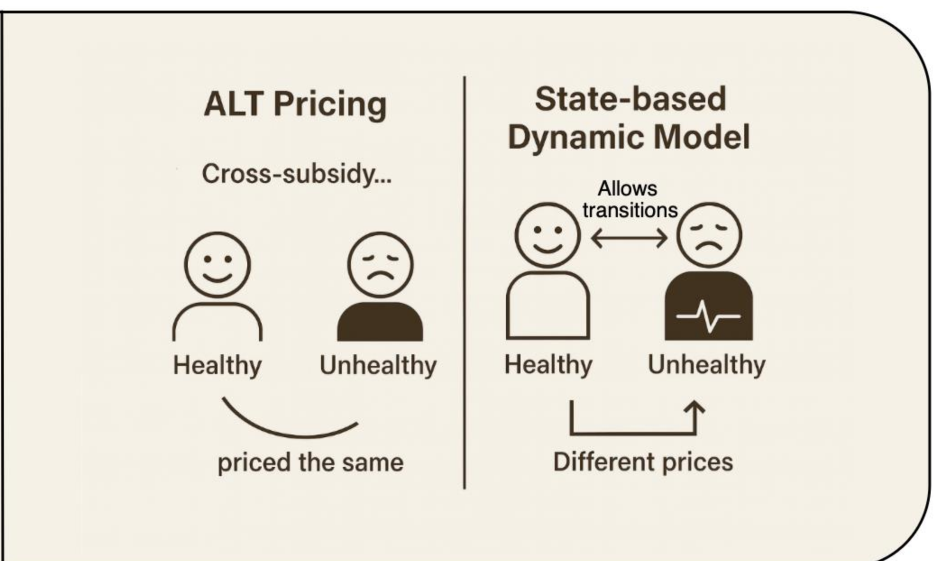


The model incorporates state-to-state transitions, distributing mortality risk across health trajectories rather than over-emphasising near-death labels.

4. How this impacts your retirement

Benefits of proposed model:

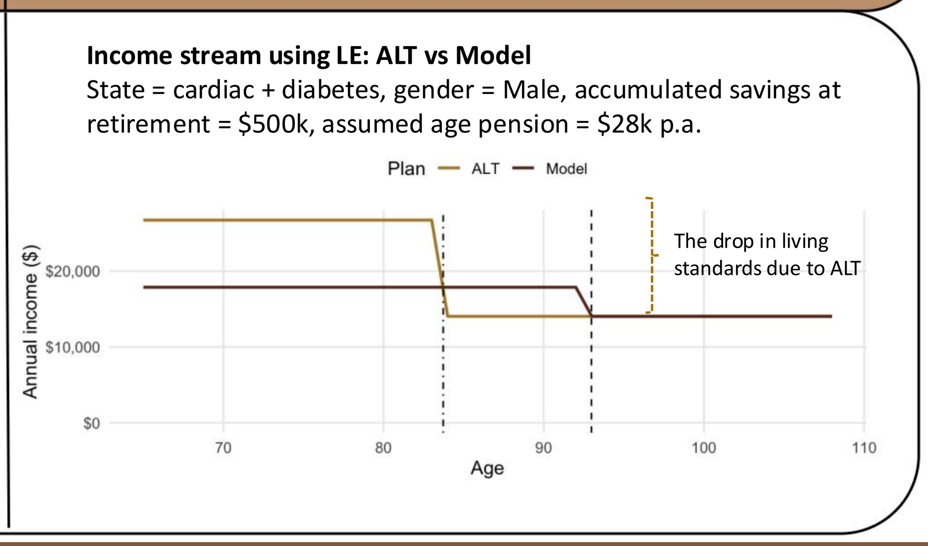
Fairer retirement options



Reduced Pension dependence

	ALT	Model
Probability outliving savings	9.05%	**4.61%**
Expected period by which you outlive savings	4.31 years	**3.57 years**

Steadier living standard



For simplicity, assume one performs **constant withdrawal till life expectancy (LE)** from their superannuation balance.

LE defines retirement planning horizon:
underestimating q_x or **LE** risks savings running out;
overestimating means lower living standards.

Better q_x → safer, steadier income.

The model remains interpretable, scalable and ***outperforms the ALT in predicting annual mortality*** q_x.

Key Contributions

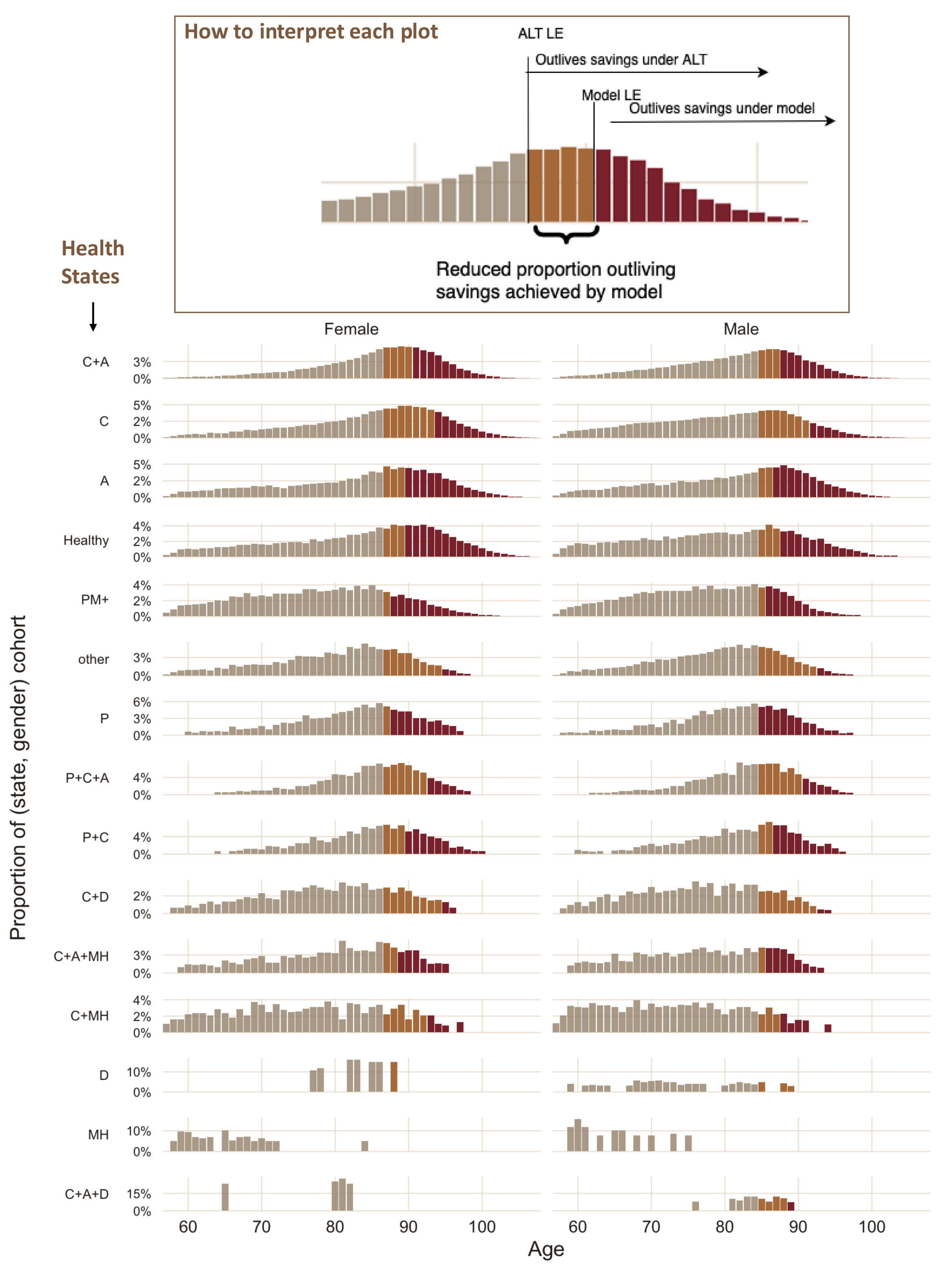
ALT:
Age + gender only. Within each gender, ALT gives the **same LE for every health state.**

Our model:
Health state-specific. Mortality (q_x^s) and transitions vary **by state**; **LE shifts by state** revealing real heterogeneity.

Direction of change:
ALT tends to **underestimate longevity**. If withdrawals are set on ALT, people are **more likely to outlive savings**; the model supports **safer, steadier drawdowns.**

Gender gap:
Females live longer than males in both ALT and our model; the **state effect applies to both genders** (two columns of the graph show consistent improvement).

Proportion outliving life expectancy: ALT vs Proposed Model



5. Limitations

- **Short modelling window**→ data limited to 2011–2016; may miss long-term patterns
- **End-of-life bias** → transitions closer to death may dominate, overstating severe health states (impact is reduced due to transitions)
- **Aggregated modelling** → operates on clusters, not individuals; limits precision.
- **Data constraints** → MBS/PBS lack diagnosis fields; rely on proxies like drug use or procedures.
- **Computational limits** → virtual machine restrictions reduced scope for large-scale validation and tuning.

6. Summary & Benefits

Prediction of q_x
- State-based q_x improves accuracy over ALT.
- Captures real health differences within age and gender.

Qualitative strengths
- **Dynamic:** models recovery and deterioration via transitions.
- **Simple to explain:** states map to questionnaire-ready items, **interpretable** model outcomes
- Scalable and auditable with clear inputs and outputs.

Actuarial implications
- More accurate LE estimates for planning and pricing.
- Smoother, safer drawdowns with fewer retirees outliving savings.
- Lower unexpected Age Pension reliance and better targeting of funds.

Takeaway:
Better q_x gives accurate LE, stable retirement income, and clearer policy guidance.