

# Modelling Small Area Population Movements

**Anna Munday**

**Dr John Henstridge**

*Data Analysis Australia*

# Outline

- **Background**
- **Forecasting populations**
- **The DAA migration model**
- **Some results**

# Background

- Population forecasts are used for making strategic and long term decisions
- Migration is the key to forecasting regional Australia
- Some of our clients require forecasts for small or regional areas
- A need to better understand small area migration

# Forecasting population

## → Simple demographics

- People get older
- Some die
- More are born
- Some move away
- Some move in

## → Measure using statistics

## → Combine using mathematics

# What a model looks like

→ Consider five year age groups

→ And five year time periods

→ E.g. from 1996 to 2001

→ Match Census years

# Let most people age

→ Move to the next age group

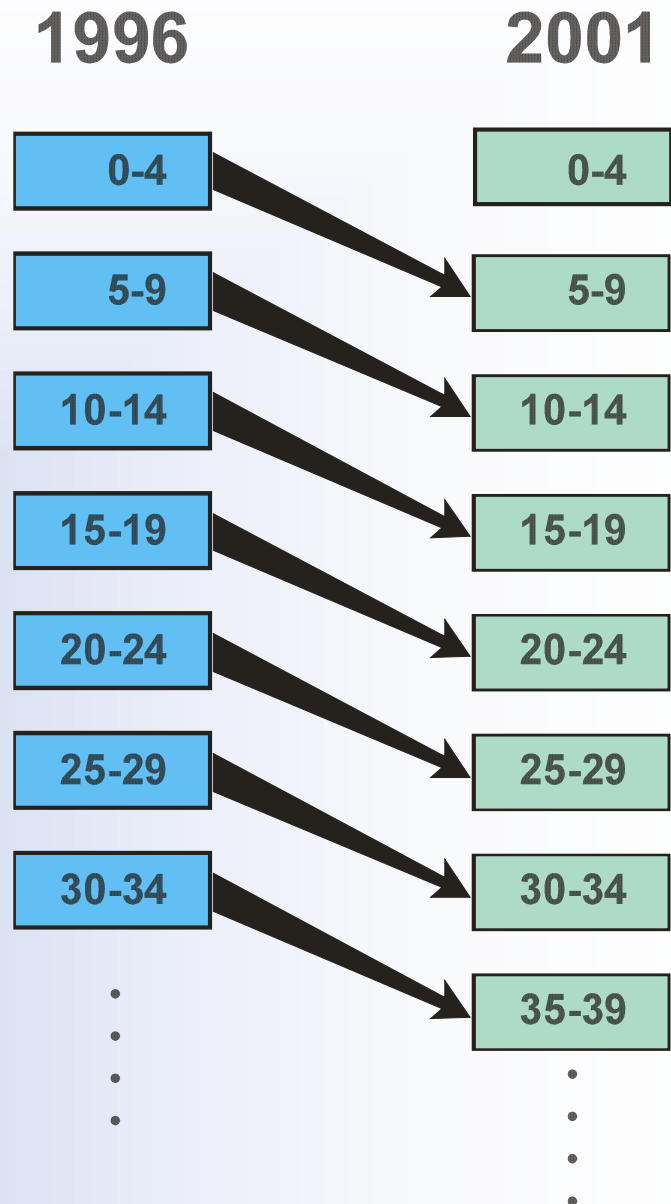
→ The one exact thing

→ Some do not make it

→ Age specific mortality

→ Well measured

→ Very predictable



# Add in births

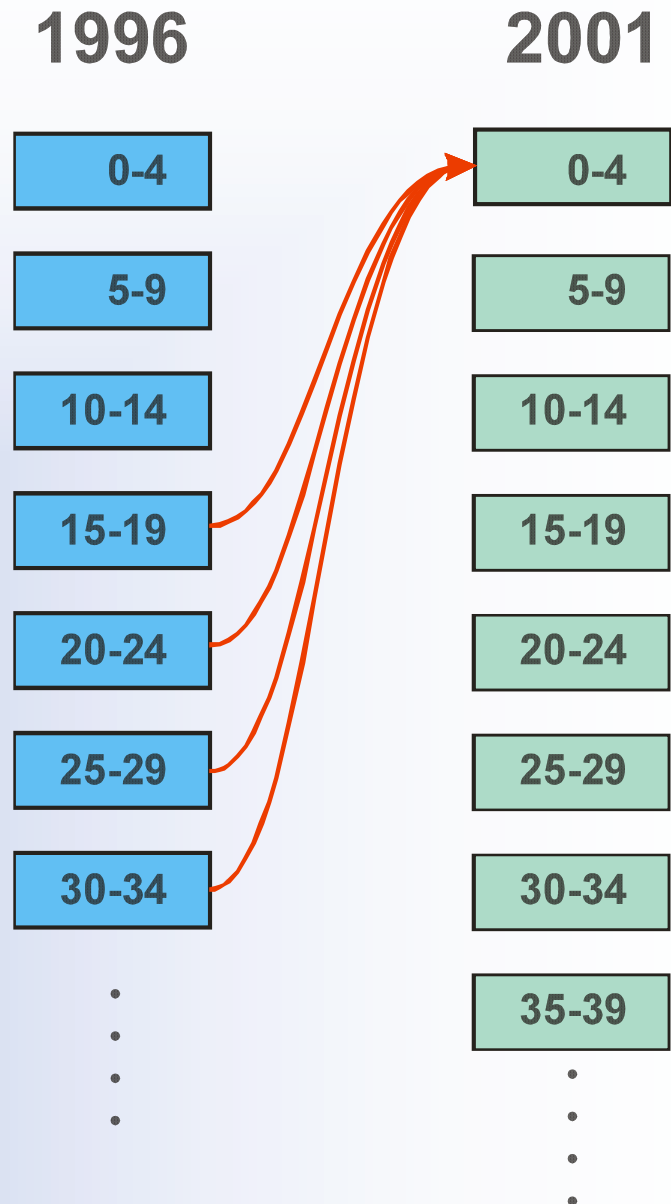
→ People produce babies

- Age specific birth rates
- Linked to adult females

→ Well measured

- Births registered

→ Reasonably predictable



# Subtract emigration Add immigration

→ Both also depend on  
age

→ Intrastate

→ Interstate

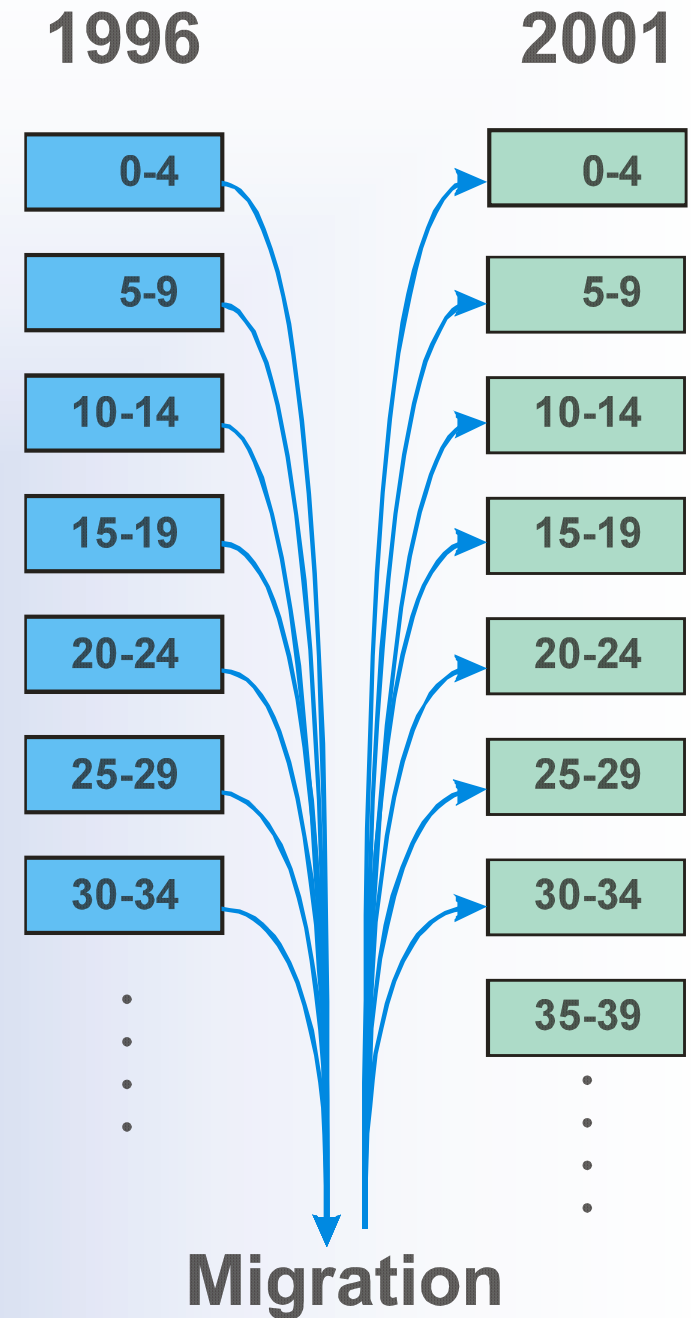
→ Overseas

→ Poorly measured

→ Only from Census

→ Poorly predicted

→ Linked to economy



# Put it together

→ Complex when combined

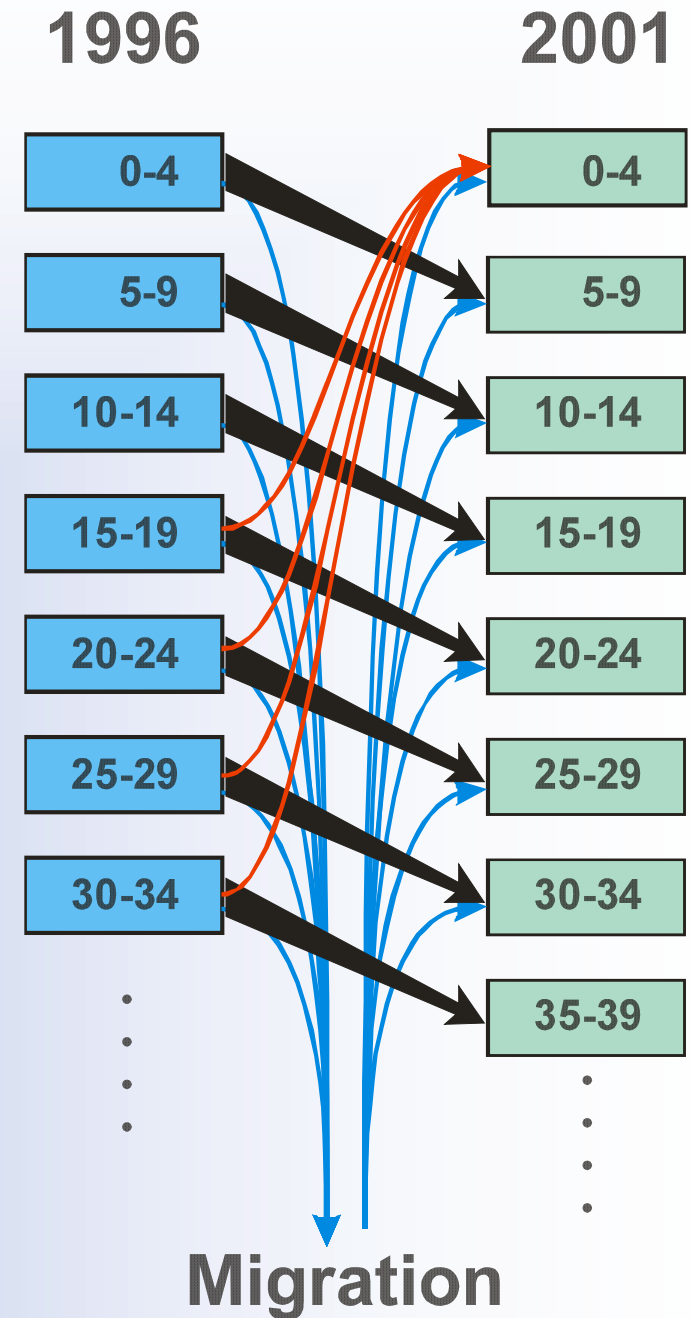
→ Still need to think of gender

→ Do for each area

→ Australia

→ State

→ Local area



# The DAA migration model

→ **Aim: Model small area migration**

→ **Data**

→ Custom tables from ABS – count of movements for each age-sex cohort between each pair of SSDs between 1996 and 2001

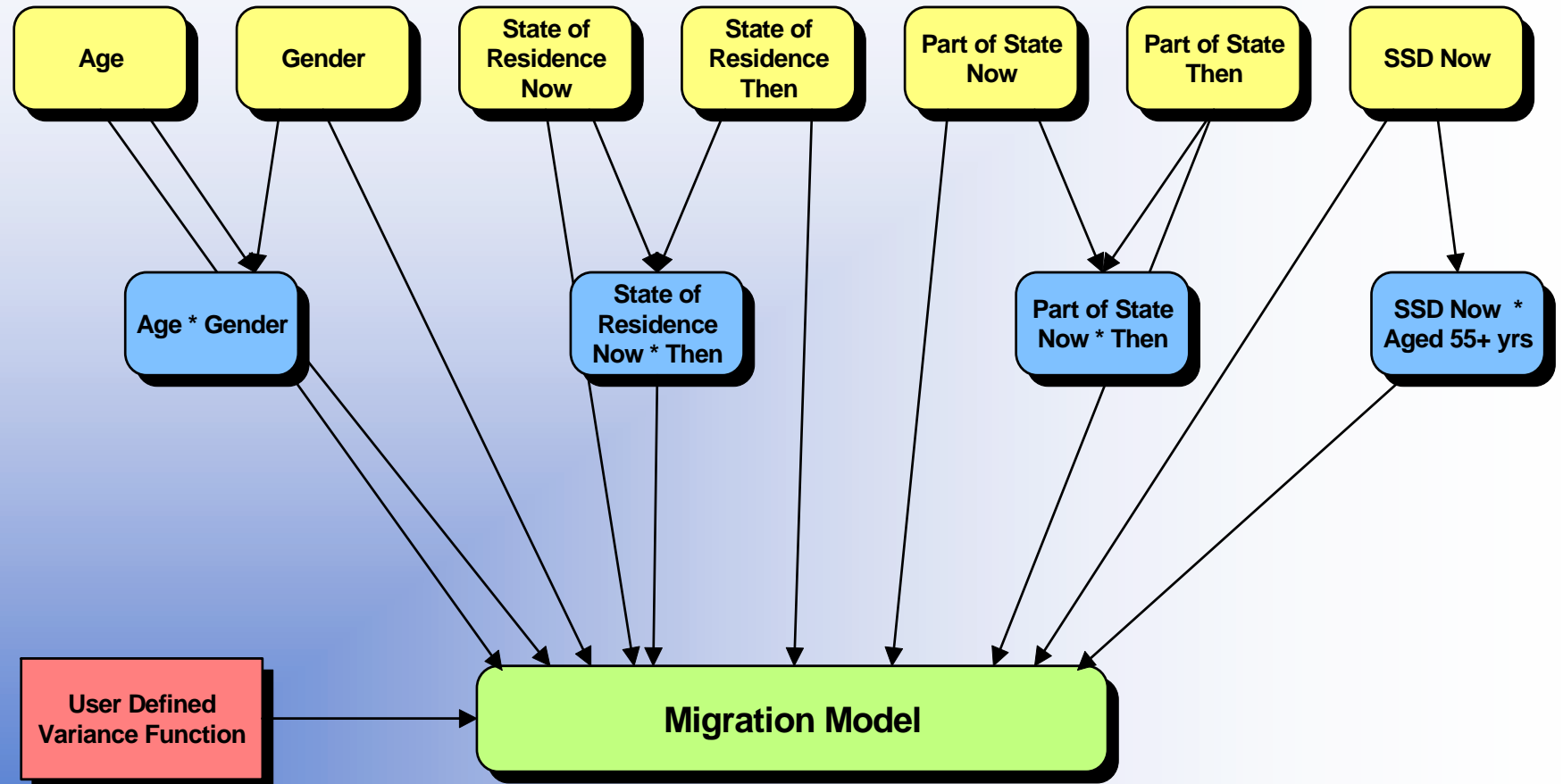
→ Population counts from 1996 ABS Census

→ **Many parameters (198\*198\*18\*2\*...)**

→ **Aim to reduce the number of parameters**

→ **Fitted a Poisson GLM**

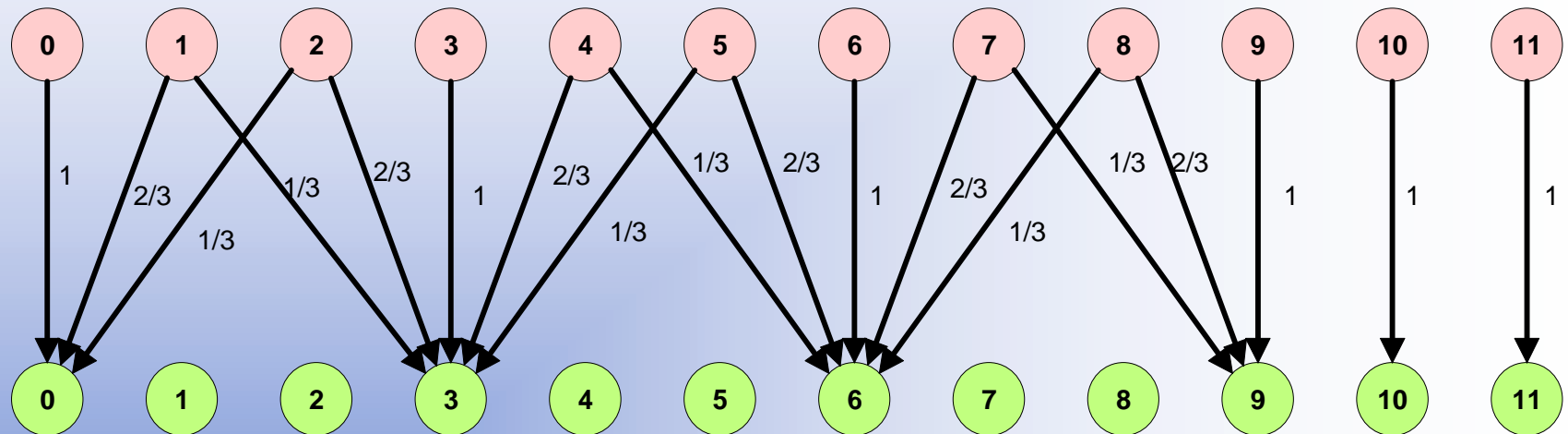
# Model structure



# Issues

- Large but sparse dataset
- Confidentialising process
- Moving as a family/household
- Increase variance – not appropriate to apply GLM technique with the usual variance function

# Confidentialising process



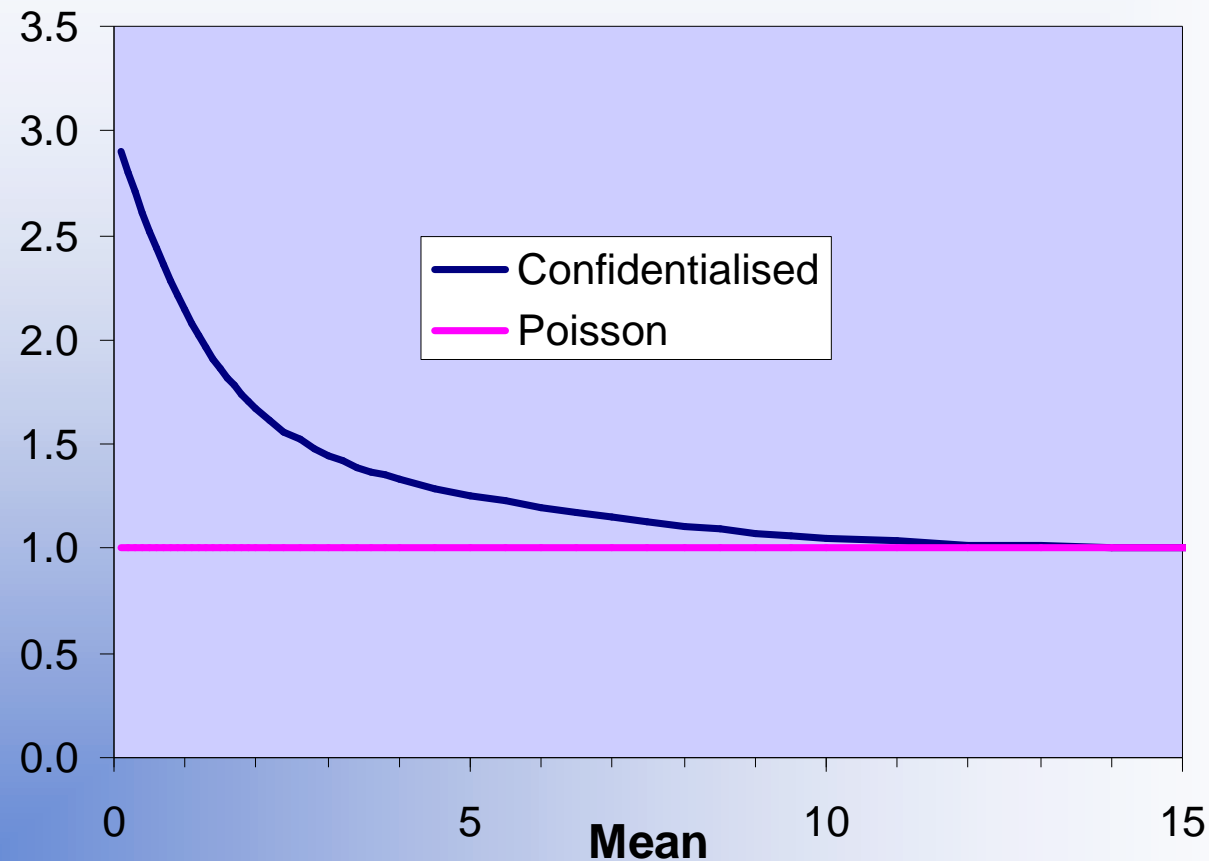
→ Small values shifted to multiple of 3

→ Irreversible process

→ Increases variance

# Variance function

Variance/Mean



→ Empirically calculated

→ Application only for small values means no simple form

# Why use GLIM?

→ **\$OWN directive allows non-standard links and variance function**

→ Log link, non-standard variance and deviance functions

→ **Memory efficient**

→ Avoided massive file structures

- ~600 parameters
- >1,100,000 cells
- Control over iterations

→ **Efficiently explore different models**

# Advantages of the model

## → Absolute consistency

→ Modelling movements

## → Smoothing process

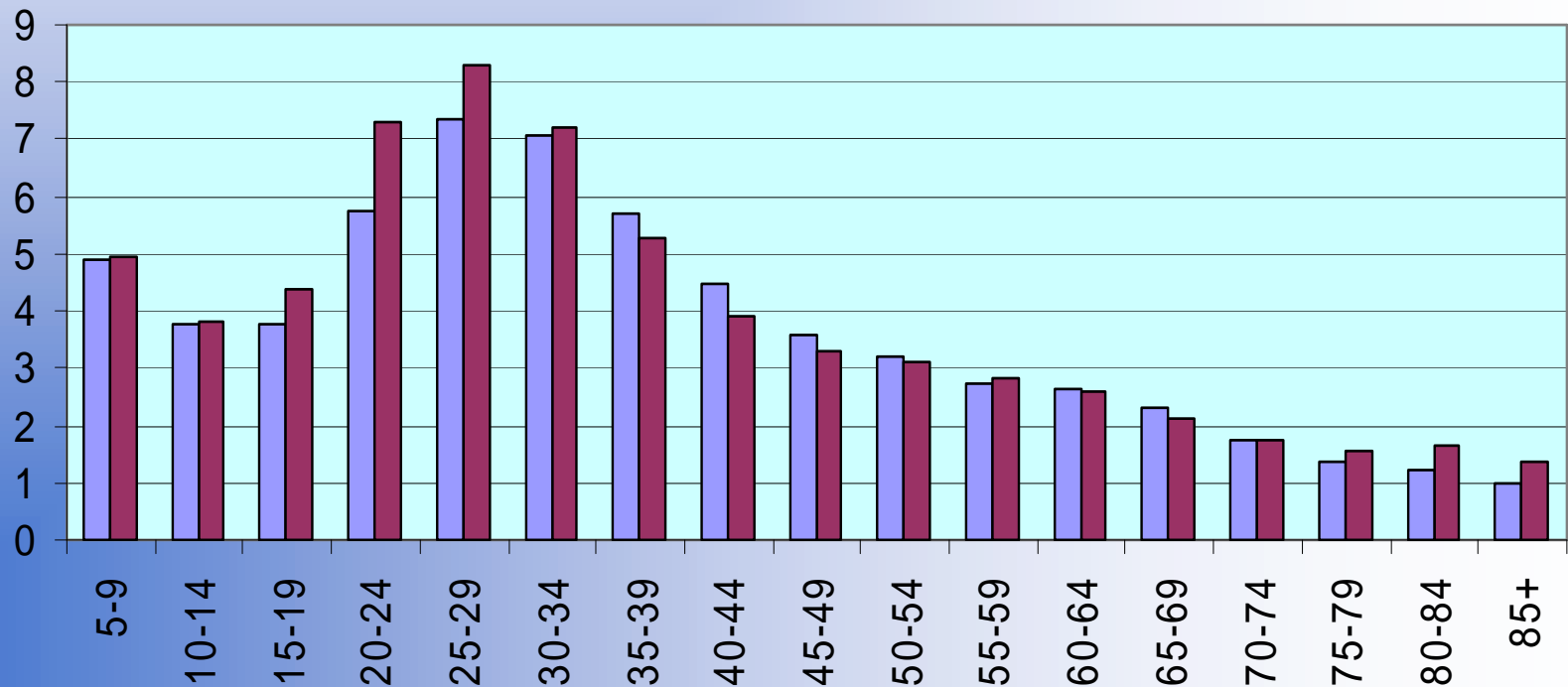
→ Maintain richness of the data but with some smoothing

→ Still considers issues such as age-specific migration

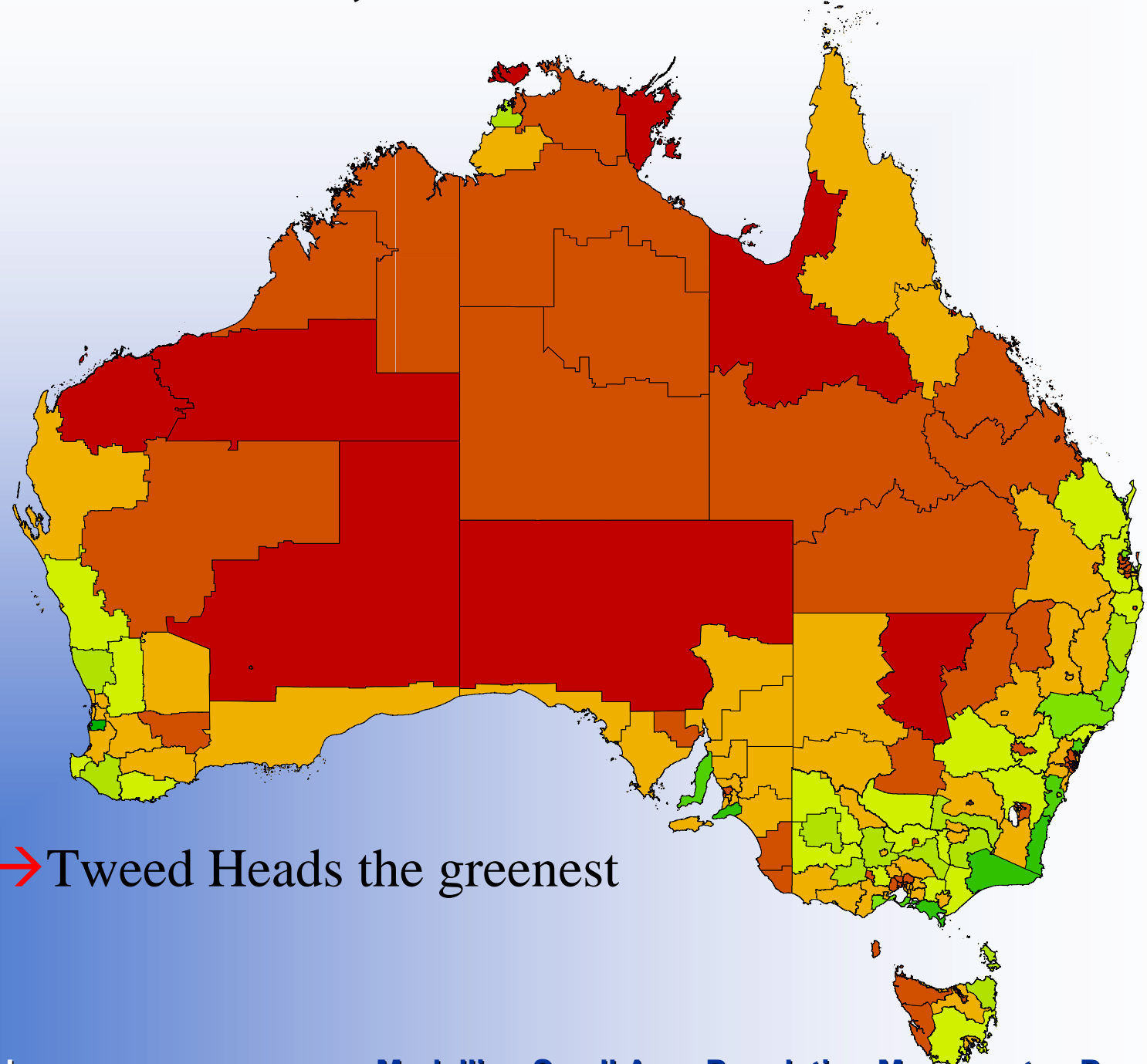
# Movements by age

- People move for work and family
- Children move with their parents

Relative Movements by Age



# Seniors, relative to others



→ Tweed Heads the greenest