The Higher Education Loan Programme (HELP/HECS) – Microsimulation Modelling of Individual Repayment Prospects

Michael O'Neill

(Paper by Michael O'Neill and Susan Antcliff)

Outline

1. Microsimulation modelling of incomes
2. The HELP Scheme
3. Review of HELP models and data
4. Model and challenges in implementation
5. Model performance and shortcomings
6. Application to other actuarial problems

1. Microsimulation modelling of incomes

- Simulating the behaviour of individual units:
  1. static or dynamic
  2. deterministic or stochastic
- Dynamic and stochastic income models:
  – Champernowne (1953): Markov chains for probability of transition between classes
  – Lillard and Willis (1978): Standard earnings function using demographics and environmental covariates

  Income = F(demographics, time) + error

2. The HELP Scheme

- Total revenue
  - up front 30%
  - deferral 55-60%
  - voluntary 10-15%

Income Range	Repayment Rate
Below $41,595	Nil
$41,595–$46,333	4.0%
$46,334–$51,070	4.5%
$51,071–$53,754	5.0%
$53,755–$57,782	5.5%
$57,783–$62,579	6.0%
$62,580–$65,873	6.5%
$65,874–$72,492	7.0%
$72,493–$77,247	7.5%
$77,248 and above	8.0%

- Scheme design:
  - non-linear income-contingent repayments
  - very long time frames
  - changing policy parameters

3. Review of HELP Models and Data

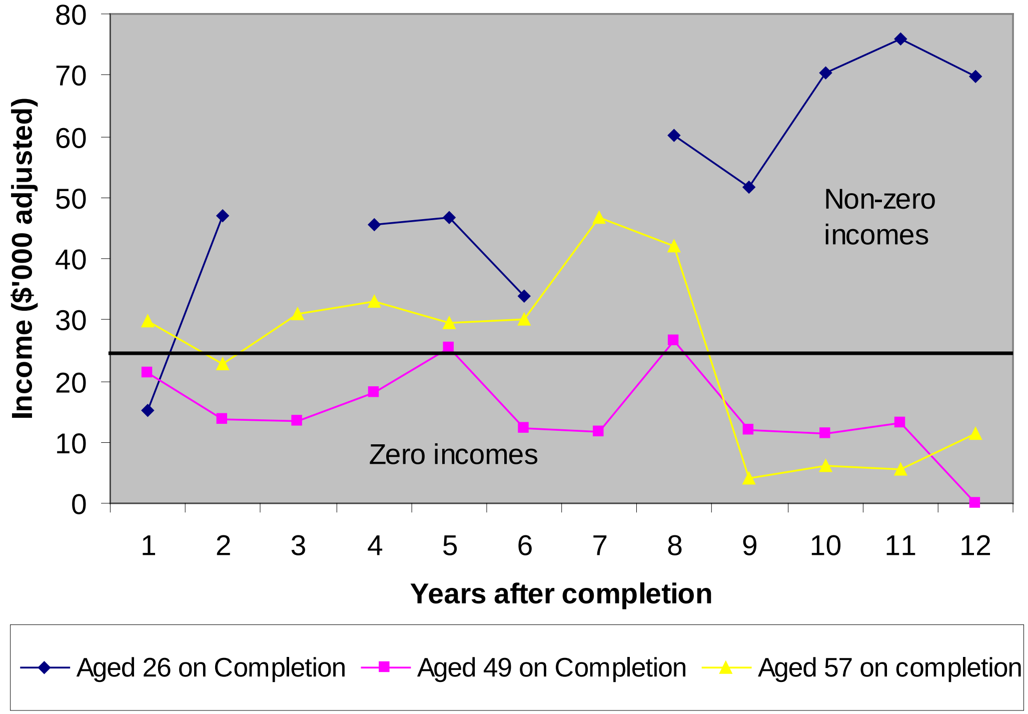
- *Speculative Model (1994)*
  - Public economic data and decrements
- *Cell Based Model (1995)*
  - 6 demographic, 3 repayment categories
  - Reliant on stable transition probabilities.
- *Microsimulation Model Version 1 (1998)*
  - Monte Carlo Markov Chain, memory-less, excessive volatility
  - Time to repayment and level of doubtful debt overestimated
- *Microsimulation Model Version 2 (2004)*
  - 10 years longitudinal income data
  - Lifetime income profiles

3. Review of HELP Models and Data

- ATO assessable incomes since 1993/94
- "Non-zero income" = Income which might potentially give rise to a repayment
- Groups:
  1. Never earn a non-zero income
  2. Always earn non-zero incomes
  3. Oscillate between zero and non-zero incomes

3. Review of HELP Models and Data

Figure 1: Examples of intermittent income profiles – Males



3. Review of HELP Models and Data

Figure 2: Examples of intermittent income profiles – Females

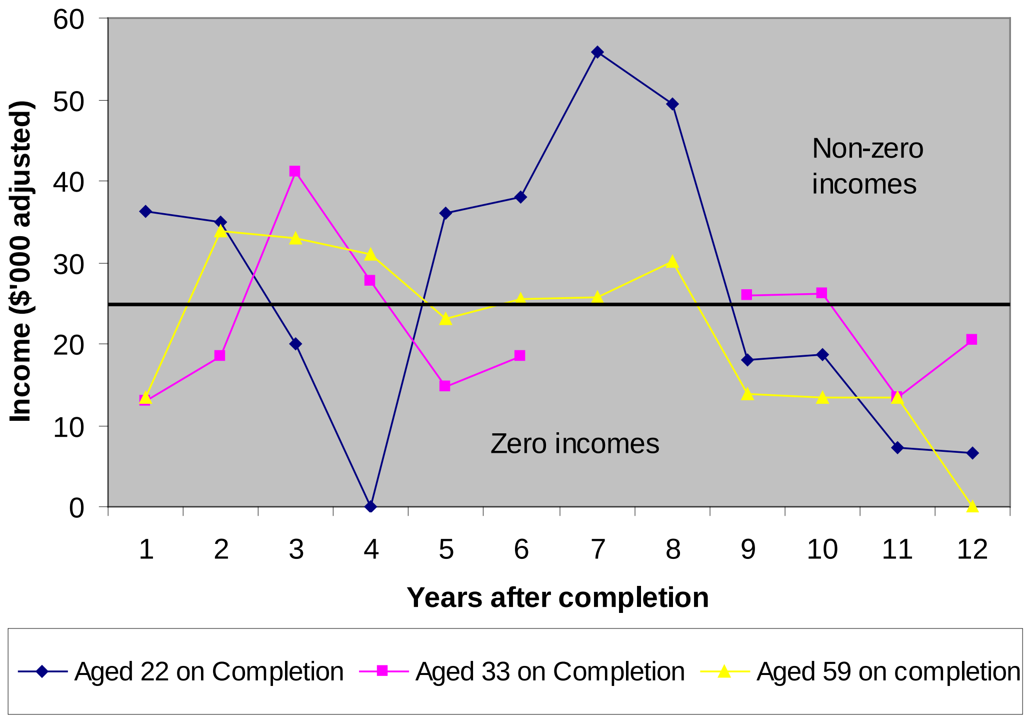


Figure 3: Promotional income profiles

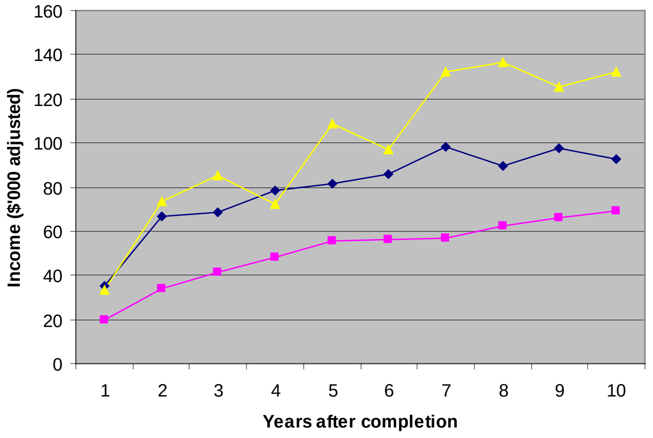


Figure 4: Flat income profiles

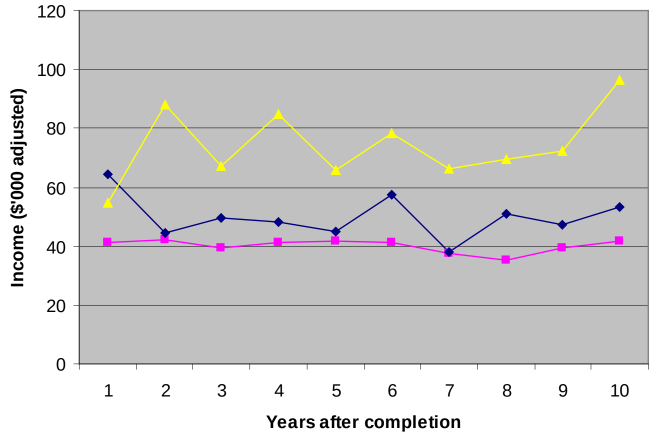
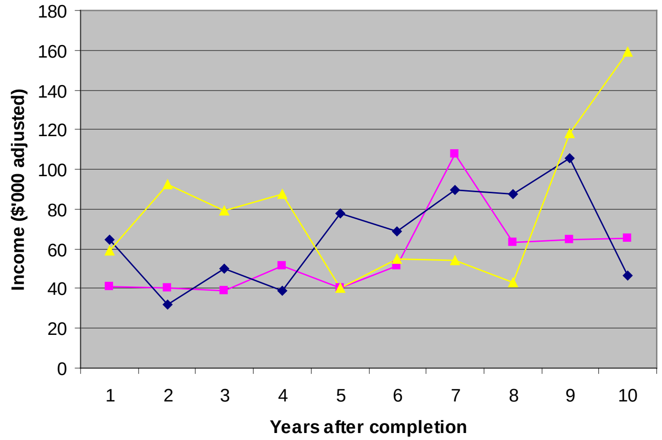


Figure 5: Highly variable income profiles



4. Model and challenges in implementation

1. Income incidence model: projects when an individual will have a non-zero income

   F(demographics, income history)

2. Income progression model: projects amount of incomes

   F(demographics, income history, projected income incidence)

4. Model and challenges in implementation

Income incidence model:

- Males and females modelled separately
- Discrete probability of never earning (zero-inflation)

$$\Pr\left(\sum_{j=i+1}^{\infty} Inc_j = 0\right) = f\left(I(x \le 25), I(x \ge 45), x, x^2 \mid i\right)$$

where x is the age at completion

i is the number of years since completion

Inc_j is the projected income in year i

$f(x|i)$ is used to denote a generic function of a linear combination of x, conditional on i (a GLM)

4. Model and challenges in implementation

– Probability of earning a non-zero income i years following completion

$$\Pr\left(Inc_i \geq 0\right) = \begin{cases} f\left(y, y^2\right), & i = 1 \\ f\left(y, y^2, I\left(Inc_1 > 0\right)\right), & i = 2 \\ f\left(y, y^2, i, (x+i), I(i>4), I(i>8), I\left(Inc_{i-1} > 0\right), I\left(\sum_{j=1}^{i-1} Inc_j > 0\right), \left(\sum_{j=1}^{i-1} Inc_j > 0\right) \Big/ (i-1)\right), & i \geq 3 \end{cases}$$

where *y* is the age in projection year *i*

4. Model and challenges in implementation

Income progression model:

– Probability of assignment to the regression group

$$\Pr(\text{Regression}_{\text{i}}) = \begin{cases} f(g, y, y^2, d), & i = 0 \\ f(g, y, y^2, d, Inc_1), & i = 1 \\ f(\hat{R}_i^{\,2}, \hat{\beta}_i, \hat{\mu}_i, \hat{\sigma}_i^{\,2}, g, y, y^2, d \mid i), & i \geq 2 \end{cases}$$

where *g* is an indicator variable for gender

d is the duration of study

$\hat{\mu}_i, \hat{\sigma}_i^{\,2}, \hat{R}_i^{\,2}, \hat{\beta}_i$ are the mean and variance, and the significance and slope of the log-linear regression.

4. Model and challenges in implementation

$$Inc_{i,k} = a_k + b_k \times \ln(i + \lambda_k) + \varepsilon_k$$

for individual k, where λ_k = 1 or 10, capturing major differences in steepness

– Point-estimate parameters

$$\sigma = \begin{cases} f(g, x, x^2, d), & i = 0 \\ f(g, x, x^2, d, Inc_1), & i = 1 \\ f(g, x, x^2, d, \hat{\mu}_i, \hat{\sigma}_i), & i \geq 2 \end{cases} \qquad \mu = \begin{cases} f(g, x, x^2, d \mid \sigma), & i = 0 \\ f(g, x, x^2, d, Inc_1 \mid \sigma), & i = 1 \\ f(g, x, x^2, d, \hat{\mu}_i \mid \sigma), & i \geq 2 \end{cases}$$

where $\hat{\mu}_i, \hat{\sigma}_i$ are the estimated mean and standard deviation fitted to the data.

4. Model and challenges in implementation

– Regression parameters

$$\alpha = \begin{cases} f(g,x,x^2,d), & i=0 \\ f(g,x,x^2,d,Inc_1), & i=1 \\ f(g,x,x^2,d,\hat{\alpha}_i,\hat{\beta}_i,\hat{\sigma}_i), & i\geq 2 \end{cases} \qquad \beta = \begin{cases} f(g,x,x^2,d\mid\alpha,\sigma), & i=0 \\ f(g,x,x^2,d,Inc_1\mid\alpha,\sigma), & i=1 \\ f(g,x,x^2,d\mid\alpha,\sigma), & i\geq 2 \end{cases}$$

$$\sigma = \begin{cases} f(g,x,x^2,d\mid\alpha), & i=0 \\ f(g,x,x^2,d,Inc_1\mid\alpha), & i=1 \\ f(g,x,x^2,d,\hat{\sigma}_i\mid\alpha), & i\geq 2 \end{cases}$$

where $\hat{\alpha}_i,\hat{\beta}_i,\hat{\sigma}_i$ are the estimated intercept, slope and standard deviation of the regression fitted to the data.

Projected nominal assessable incomes

Figure 6: Males with history

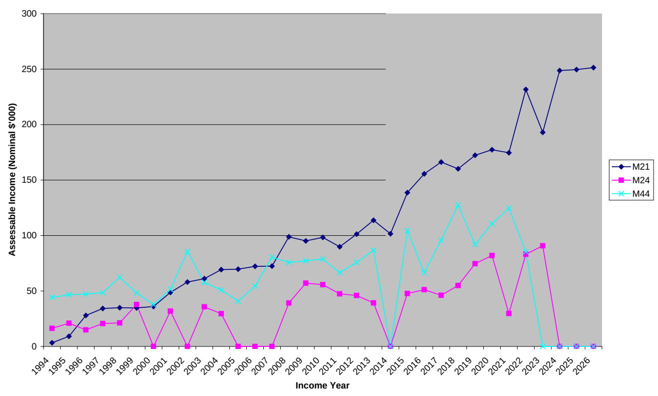


Figure 7: Females with history

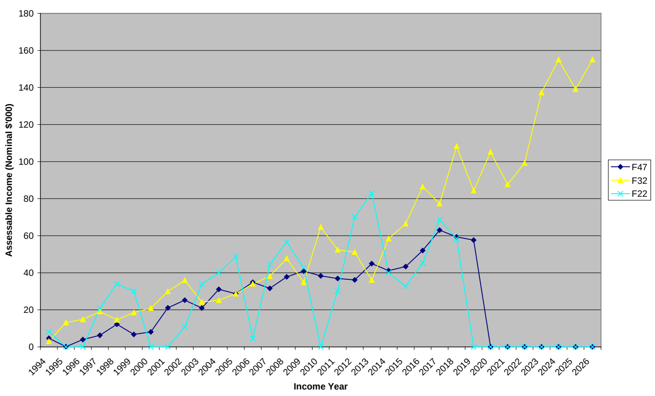
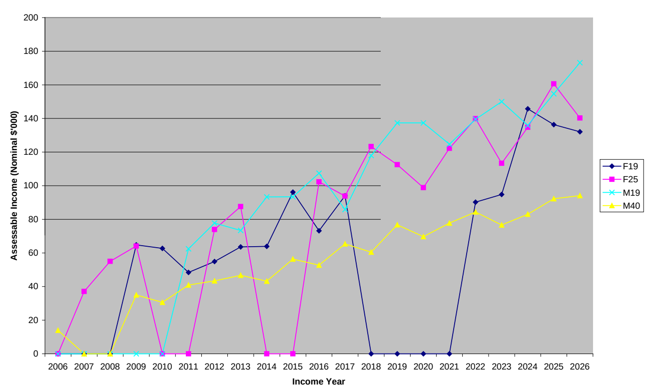


Figure 8: No income history



5. Model performance and shortcomings

Figure 9: Deviation between revenue projected and actual revenue

Financial Year	Model	Percentage Deviation
1995-96	Cell based	11.2%
1996-97	Cell based	10.3%
1997-98	Cell based	-4.5%
1998-99	Microsimulation v.1	-18.5%
1999-00	Microsimulation v.1	1.2%
2000-01	Microsimulation v.1	-5.4%
2001-02	Microsimulation v.1	7.1%
2002-03	Microsimulation v.1	-4.9%
2003-04	Microsimulation v.1	12.5%
2004-05	Microsimulation v.2	13.4%*
2005-06	Microsimulation v.2	0.9%
2006-07	Microsimulation v.2	0.5%
2007-08	Microsimulation v.2	-2.1%

* The 2004-05 result was substantially affected by the reduction in the discount available on voluntary repayments from 15% to 10% with effect from 1 January 2005, resulting in a large bring forward of voluntary repayments to the second half.

5. Model performance and shortcomings

1. Scheme maturity
2. Stationarity / stability of the conditional distributions
3. Inappropriate functional forms
4. Macro-economic feedback

6. Application to other actuarial problems

1. Where there is a non-linear element and longitudinal outcomes are of interest, eg:
   - income tested social security payments; and
   - across year income averaging provisions in the tax system.
2. Design of health insurance products including excess, taking into account variability in health status and trends in usage with age, and segmenting by usage.

Contact details

Michael O'Neill FIAA 2007
BActS (Hons) / LLB 2004
Investment Analyst
Cannae Capital Partners
+61 2 8023 4603
www.cannae.com.au

Susan Antcliff FIAA
BSc (Hons)
Actuary
Australian Government Actuary
+61 2 6263 4189
www.aga.gov.au