Socio-Economic Characteristics at the Small Area Level: New Models and Data for Policy Makers and for Needs Based Planning of Government Services

Ann Harding

Presentation to Department of Geography Seminar Series, University of California Santa Barbara, USA, 17 May 2007

National Centre for Social and Economic Modelling (NATSEM), University of Canberra
What are microdata and microsimulation models?

- Focus on individuals or households
- Start with large microdata sets (admin or sample survey)
- Primarily used to estimate impact of government policy change on these individuals or households
  - Impact on small sub-groups
  - Aggregate impact
  - Impact on government revenue or expenditure
Static models of taxes and transfers
Income tax and social security

- STINMOD model is now maintained by NATSEM for Australian government departments (Family and Community Services, Education Science and Training, Treasury, Employment and Workplace Relations)
- 13 years old
- STINMOD simulates all the major income tax and cash transfer programs (age pension, family payments etc)
- Used regularly in research – distributional impact of welfare state, impact of minimum wage rises, EMTRs
- Constructed on top of Income Surveys and Expenditure Surveys (2 versions)
### Income Tax Scale Steps

Year 2003-04 average parameters

<table>
<thead>
<tr>
<th>Step</th>
<th>Income</th>
<th>Marginal Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>0</td>
<td>0.00</td>
</tr>
<tr>
<td>Two</td>
<td>6000</td>
<td>0.17</td>
</tr>
<tr>
<td>Three</td>
<td>20000</td>
<td>0.30</td>
</tr>
<tr>
<td>Four</td>
<td>50000</td>
<td>0.42</td>
</tr>
<tr>
<td>Five</td>
<td>60000</td>
<td>0.47</td>
</tr>
<tr>
<td>Six</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Seven</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Eight</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Nine</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Ten</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

---

**OK**  **Refresh**  **Cancel**
## Outcome by Family Type and Income

**Estimated Change in Family Disposable Income - $ pw**

**2003-04 budget changes**

**Outcome:** ALL  
**Population:** All Recipients  
**Year 2003-04 Model Run**

<table>
<thead>
<tr>
<th>Weekly Taxable Income</th>
<th>Married no childr.</th>
<th>Married + children</th>
<th>Sole Parent</th>
<th>Single Adult</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 150</td>
<td>0.02</td>
<td>0.31</td>
<td>0.09</td>
<td>0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>150-299</td>
<td>0.26</td>
<td>0.66</td>
<td>0.01</td>
<td>0.33</td>
<td>0.29</td>
</tr>
<tr>
<td>300-449</td>
<td>0.42</td>
<td>1.17</td>
<td>2.08</td>
<td>2.79</td>
<td>1.52</td>
</tr>
<tr>
<td>450-599</td>
<td>2.05</td>
<td>4.16</td>
<td>4.77</td>
<td>4.70</td>
<td>4.05</td>
</tr>
<tr>
<td>600-749</td>
<td>5.14</td>
<td>4.15</td>
<td>4.50</td>
<td>4.03</td>
<td>4.32</td>
</tr>
<tr>
<td>750-899</td>
<td>5.51</td>
<td>5.09</td>
<td>4.11</td>
<td>4.00</td>
<td>4.59</td>
</tr>
<tr>
<td>900-1049</td>
<td>7.07</td>
<td>6.36</td>
<td>5.48</td>
<td>5.86</td>
<td>6.29</td>
</tr>
<tr>
<td>1050-1199</td>
<td>8.40</td>
<td>8.11</td>
<td>8.36</td>
<td>8.95</td>
<td>8.49</td>
</tr>
<tr>
<td>1200-1349</td>
<td>8.79</td>
<td>9.20</td>
<td>10.60</td>
<td>11.02</td>
<td>9.48</td>
</tr>
<tr>
<td>1350-1499</td>
<td>9.31</td>
<td>10.03</td>
<td>*</td>
<td>11.02</td>
<td>9.94</td>
</tr>
<tr>
<td>1500+</td>
<td>14.02</td>
<td>13.98</td>
<td>11.52</td>
<td>11.02</td>
<td>13.70</td>
</tr>
</tbody>
</table>

**TOTAL** | 5.99 | 8.31 | 2.24 | 2.75 | 4.64
Disposable income of sole parents with one child aged 8+, 2006-07
Impact of 2005 ‘welfare to work’ budget changes
EMTRs of sole parents with one child aged 8+, 2006-07 *

* EMTR of 65% means that person keeps 35 cents from an additional dollar of earnings
Dynamic models
Dynamic microsimulation modelling

- **Simulates** the events that happen to ordinary Australians over their lifetime
- Starts in 2001 with 180,000 people (1% of the Australian population – the Census sample)
- Models individuals (the **micro** level)
- Uses regression equations to model human behaviour over time (**dynamic**)
- APPSIM (Australian Population and Policy Simulation Model) currently under devt – won ARC grant last year with 13 govt depts as research partners
- Earlier version was DYNAMOD3
DYNAMOD3’s Simulation Cycle

Australia 2001

Simulation Clock

Immigration

New Quarter?

Macro-economics

Couple Dissolution

Pregnancy and Births

Deaths

Emigration

Disability & Recovery

New Year?

Couple Formation

Annual Earnings And Savings

Labour Force

Education

Simulation Clock

Australia 2050
Spatial models
**Characteristics of available datasets**

<table>
<thead>
<tr>
<th></th>
<th>National sample surveys</th>
<th>Census of Popn &amp; Housing</th>
<th>?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Population detail</strong></td>
<td>High</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td><strong>Geographic detail</strong></td>
<td>Low</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>
Synthetic Spatial Microdata

Solution:

Combine the information-rich survey data with the geographically disaggregated Census data

Using ‘spatial microsimulation’ (synthetic estimation) to:

create detailed unit record data for small areas (synthetic spatial microdata)
Constructing small-area estimates

**SMALL AREA DATA**
- 2001 Census data at SLA level:
  - XCP data for SLAs

**UNIT RECORD DATA** (SOURCE)
- 1998-99 Household Expenditure Survey

**UNIT RECORD DATA** (AMENDED)
- Updated to 2001 Enhanced income

**REWEIGHTING USING LINKING VARIABLES**

**SMALL-AREA ESTIMATES**
1) Unit record dataset
2) Set of weights for each SLA

Major data preparation task

Iterative process to identify a set of variables suitable for reweighting
What is ‘reweighting’?

turning the national household weights in the HES survey file into ...  ... household weights of small-areas

<table>
<thead>
<tr>
<th>Unit record</th>
<th>Household ID</th>
<th>Weekly income</th>
<th>Weekly rent</th>
<th>Other variables</th>
<th>Household weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>.</td>
<td>1029</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>11</td>
<td>4</td>
<td>.</td>
<td>157</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>11</td>
<td>4</td>
<td>.</td>
<td>157</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>11</td>
<td>4</td>
<td>.</td>
<td>157</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>11</td>
<td>0</td>
<td>.</td>
<td>1003</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>11</td>
<td>0</td>
<td>.</td>
<td>1003</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>10</td>
<td>4</td>
<td>.</td>
<td>70</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>12</td>
<td>4</td>
<td>.</td>
<td>70</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>12</td>
<td>0</td>
<td>.</td>
<td>703</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>12</td>
<td>0</td>
<td>.</td>
<td>703</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>13964</td>
<td>6892</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>NSW SLA1</th>
<th>NSW SLA2</th>
<th>NSW SLA3</th>
<th>Other SLAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2.45</td>
<td>13.54</td>
<td>16.38</td>
<td>0</td>
</tr>
<tr>
<td>2.45</td>
<td>13.54</td>
<td>16.38</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3.27</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3.27</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12465</td>
<td>25853</td>
<td>27940</td>
<td>0</td>
</tr>
</tbody>
</table>

Num of households in Aust: 7,122,000

Num of households in small areas: 12465 25853 27940
Linkage variables available in the 2001 Census and 1998-99 HES

- **Person level variables**
  - Age
  - Sex
  - Social marital status
  - Country of birth
  - Level of schooling
  - Non-school qualifications
  - Educational institution attending
  - Study status
  - Hours worked
  - Individual income
  - Occupation
  - Labour force status
  - Year of arrival
  - Relationship in household

- **Family level variables**
  - Family type
  - Family income

- **Household level variables**
  - Dwelling structure
  - Tenure type
  - Household income
  - Household type
  - Household size
  - Number of dependents
  - Number of cars
  - Rent paid
  - Mortgage repayments
Application 1: Analysis of Specific Population Sub-Groups

- Allows – for small areas:
  - identification and analysis of specific socio-demographic groups and characteristics
  - analysis at various population levels: e.g. persons, income units, households

- Examples – children in low income families; children in jobless families; unskilled youth, those in housing stress
Percentage of households in unaffordable housing

- 0.88% - 6.05%
- 6.06% - 9.48%
- 9.49% - 14.00%
- 14.01% - 29.27%
Age profile of those in poverty in postcode in metro Sydney

- 25-34 years: 27% (Aust average 20.3%)
- 35-44 years: 38% (Aust average 35.1%)
- 45-54 years: 16% (Aust average 17.7%)
- 55-64 years: 15% (Aust average 12.3%)
- 65+ years: 2% (Aust average 10.5%)
- < 25 years: 2% (Aust average 4.1%)
Application 2: Predict spatial impact of a policy change

- Spatial microdata now linked with NATSEM’s existing microsimulation models to model the immediate distributional/revenue impact of a policy change
  - link synthetic spatial output to STINMOD and model changes to the tax and transfer system for small geographic areas
  - Currently modelling changes in Commonwealth Rent Assistance, income tax, social security and family payments
Where did the $5bn of 2005-06 tax cuts go?

- Updated 2001 population numbers to 2005-06 using ABS estimates pop’n growth by SLA
- Updated household incomes and rules of govt programs to 2005-06

<table>
<thead>
<tr>
<th>2004-05</th>
<th>2005-06</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax threshold</td>
<td>Tax rate</td>
</tr>
<tr>
<td>$6,000</td>
<td>0.17</td>
</tr>
<tr>
<td>$21,600</td>
<td>0.3</td>
</tr>
<tr>
<td>$58,000</td>
<td>0.42</td>
</tr>
<tr>
<td>$70,000</td>
<td>0.47</td>
</tr>
</tbody>
</table>
Estimated average tax cut per household per week, Sydney SLAs, 2005-06

$3.50 - $9.50 (lightest)
$9.51 - $13.70
$13.71 - $19.30
$19.31 - $34.10 (darkest)
Estimated average dollar tax cut per household per week, by regions, 2005-06

<table>
<thead>
<tr>
<th>State/Territory</th>
<th>Capital city</th>
<th>Major urban</th>
<th>Regional town</th>
<th>Rural town</th>
<th>Rural area</th>
<th>All regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>NSW</td>
<td>$14.6 (1.3)</td>
<td>$9.1 (0.8)</td>
<td>$8.8 (0.8)</td>
<td>$9.0 (0.8)</td>
<td>$8.0 (0.7)</td>
<td>$12.2 (1.1)</td>
</tr>
<tr>
<td>VIC</td>
<td>$12.6 (1.1)</td>
<td>$9.7 (0.8)</td>
<td>$8.3 (0.7)</td>
<td>$8.4 (0.7)</td>
<td>$8.5 (0.7)</td>
<td>$11.5 (1.0)</td>
</tr>
<tr>
<td>QLD</td>
<td>$11.0 (1.0)</td>
<td>$9.0 (0.8)</td>
<td>$8.8 (0.8)</td>
<td>$7.1 (0.6)</td>
<td>$9.4 (0.8)</td>
<td>$9.9 (0.9)</td>
</tr>
<tr>
<td>ACT</td>
<td>$16.0 (1.4)</td>
<td>na</td>
<td>na</td>
<td>na</td>
<td>$11.7</td>
<td>$16.0 (1.4)</td>
</tr>
<tr>
<td>All four</td>
<td>$13.3 (1.2)</td>
<td>$9.1 (0.8)</td>
<td>$8.6 (0.7)</td>
<td>$8.5 (0.7)</td>
<td>$8.6 (0.7)</td>
<td>$11.5 (1.00)</td>
</tr>
</tbody>
</table>

Note: Regions are aggregates of SLAs. Figures in bracket are multiples of $11.5. Convergent SLAs only.
HOUSEMOD

- Spatial model (SLA)
- Models receipt of Commonwealth Rent Assistance
  - Means-tested assistance to low income private renters
- Can change rules of CRA and predict spatial impact
- Has been extended to add:
  - public renters as well
  - plus projections for 20 years
% in unaffordable housing
Application 3: Where to put govt offices? (access channel planning)

- The Centrelink CuSP Model (Customer Service Projection Model)

- Centrelink needed an evidence based methodology to help:
  - match services available to customers’ needs and preferences
  - deliver the service via the most suitable channel and
  - in most efficient way.

- CuSP model assists Centrelink strategic decision-making by:
  - producing projections of Centrelink customers and channel use
  - over the next 5 years
  - for small areas
  - and under alternative scenarios about the future
Projected % changes in customer numbers – 2002-07
Application 4: Forecasting current & future need for services

- CAREMOD model simulates current characteristics of older Australians at a detailed regional level (SLA)
  - Imputing functional status and thus likely need for different types of care

- Also new ARC grant 2007-2009 for examining spatial implications of population ageing over next 20 years (esp. for needs-based planning of govt services) – with four states and territories
% needing high level institutional care

PERCENTAGE OF POPULATION AGED 55 YEARS AND OVER WITH CARE MODALITY 5

Quintiles (Percent)
- 2.1 - 7.9
- 7.9 - 9.0
- 9.0 - 10.0
- 10.0 - 11.2
- 11.2 - 38.6
Where do self-funded retirees live
Other forthcoming ARC funded & ARCRNSISS initiatives

- By end 2006, estimates of poverty, housing stress & smoking expenditure in 03-04 to be available via web to ARCRNSISS members
  - At labour market area level
  - At SLA level
  - Based on new 03-04 Household Expenditure Survey

- Developing index of social exclusion for children

- Would like to link the SLA estimates to administrative data about usage of government services
  - Eg do poor children use public health services more or less than children from affluent families

- Continue to refine the technology
Evidence based policy making

- Growing demand for decision support tools
  - Reduce risk to policy makers making billion dollar decisions
  - Assess distributional implications of policy change before implemented
  - Improve predictive capacity & strategic planning

- NATSEM has now constructed dozens of microsimulation models, based on ABS or admin microdata

- Exciting new developments are
  - Spatial microsimulation models
  - Health and housing models
  - Next generation of dynamic models

- For free copies of all publications as released, email hotline@natsem.canberra.edu.au
Selected references

Spatial Microsimulation (spatial estimates of poverty, disadvantage etc)


Child Social Exclusion Index (small area index of social exclusion specifically developed for children)

CuSP Model (spatial model for service delivery and access issues)

CAREMOD (spatial model of care needs)


HOUSEMOD (spatial model of housing assistance and housing issues)


DYNAMOD (dynamic microsimulation model - now being replaced by APPSIM)


* Means available on NATSEM website at www.natsem.canberra.edu.au
Selected references

**STINMOD applications (static tax-benefit model)**

**CHILDMOD (static child support model)**

**NSW Hospitals Model (spatial model of socio-economic status and hospital usage and costs)**

**MediSim (static model of the Pharmaceutical Benefits Scheme)**

**HEALTHMOD (under development: model of hospital, medical and pharmaceutical sectors)**
Lymer, s, Brown, Payne, and Harding, 2006, ‘Development of HEALTHMOD: a model of the use and costs of medical services in Australia, 8th Nordic Seminar on Microsimulation Models, Oslo, June (http://www.ssb.no/english/research_and_analysis/conferences/misi/)