

# Mapping Significant Economic Disadvantage in New South Wales: TECHNICAL APPENDIX

## SMALL AREA ESTIMATION METHOD AND VALIDATION

OCTOBER 2019

By Yogi Vidyattama and Robert Tanton

## ABOUT NATSEM

The National Centre for Social and Economic Modelling (NATSEM) was established on 1 January 1993, and supports its activities through research grants, commissioned research and longer-term contracts for policy analysis and model development and maintenance.

In January 2014, the Institute for Governance and Policy Analysis (IGPA) at the University of Canberra was established to harness the research strengths of NATSEM and the ANZSOG Institute for Governance (ANZSIG). The aim of this Institute is to create and sustain an international class research institution for the study and practice of governance and public policy. The Institute has a strong social mission committed to the production of leading-edge research and research driven education programs with genuine public value and, by implication, policy impact. The establishment of IGPA has created exciting opportunities for the development of cutting-edge research in public policy analysis through combining expertise in qualitative and quantitative methods, micro-simulation and policy modelling and evaluation.

NATSEM is one of three research centres within IGPA. Policy changes often have to be made without sufficient information about either the current environment or the consequences of change. NATSEM aims to be a key contributor to social and economic policy debate and analysis by undertaking independent and impartial research of the highest quality, including supplying valued commissioned research services.

NATSEM is one of Australia's leading economic and social policy research centres and is regarded as one of the world's foremost centres of excellence for micro-data analysis, microsimulation modelling and policy evaluation. In keeping with IGPA's core mission, NATSEM's research activities aim to have significant policy impact and lead to social and economic change.

IGPA Director: Professor Laurie Brown

NATSEM Director: Professor Robert Tanton

© IGPA, University of Canberra 2019

All rights reserved. Apart from fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright Act 1968, no part of this publication may be reproduced, stored or transmitted in any form or by any means without the prior permission in writing of the publisher.

NATSEM at The Institute for Governance and Policy Analysis University of Canberra, ACT 2601, Australia  
Building 24, University Drive South, Canberra ACT 2617

Suggested Citation:

Vidyattama, Y. & Tanton, R., (2019), 'Mapping Significant Economic Disadvantage in New South Wales: Technical Appendix, NATSEM, University of Canberra. Report commissioned by NCOSS.

## SMALL AREA ESTIMATION METHOD

The small area estimation method used is a method published in peer reviewed journals<sup>1</sup>. It requires a Census (in this case the 2016 Census) for small area benchmarks; and a survey (in this case the 2015-16 Survey of Income and Housing) to provide individuals to fill areas. The benchmarked variables need to be available on both the population census and the survey, using the same definitions and the same categories. The benchmarks need to be related to the final variable that is required from the spatial microsimulation model – in this case, poverty rates. This means benchmarks like income and number of people in the household by age (so that the income can be equivalised to take into account the number of people in the household), and housing costs for after housing poverty, are required. The model used for this report uses 9 benchmarks from the 2016 Census as indicated in Table 1.

Table 1: Benchmarks

| Benchmark         | Description   |
|-------------------|---|
| 1 NPRD_2*HIND_2   | Number of Persons Usually Resident in Dwelling by Total Household Income (weekly) |
| 2 TENLLD_2*HIND_2 | Tenure and Landlord Type by Total Household Income (weekly)                       |
| 3 HCFMD_2*HIND_2  | Family Household Composition by Total Household Income (weekly)                   |
| 4 RNTRD_2*HIND_2  | Rent (weekly) by Total Household Income (weekly)                                  |
| 5 MRERD_2*HIND_2  | Mortgage repayments by Total Household Income (weekly)                            |
| 6 AGE_2*HIND_2    | Age of person (15+) by Total Household Income (weekly)                            |
| 7 HIED_2*HIND_2   | Equivalised Total Household Income (weekly) by Total Household Income (weekly)    |
| 8 LFSP_2*AGE_2    | Labour Force Status by Age of person (15+)  |
| 9 QALLP_2         | Non School Qualification  |

In addition, in this report we:

- Use only households from the Greater Capital City Statistical Area (GCCSA) to populate the SA2's in that GCCSA. This means we only used households from Sydney to populate SA2's in Sydney.

---

<sup>1</sup> Tanton, R., Vidyattama, Y., Nepal, B., & McNamara, J. (2011). Small area estimation using a reweighting algorithm. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 174(4), 931-951.

- Reduce the number of benchmarks if the model failed for an area. This is done according to the sequence in the table. The lower number of benchmarks means fewer constraints and a higher possibility of achieving an acceptable result. If the estimate is produced with less than 7 benchmarks, then the estimate is excluded from the overall database as unreliable.

The estimation is done at the ABS standard geography SA2 as the preferred small area unit. The estimate is then aligned to the state level results, which are calculated directly from the ABS survey. The process of the alignment is as follows:

- First, the aggregate value of the poverty rate is calculated for different groups of people and Greater Capital City Statistical Areas (GCCSA). The GCCSA represents the functional extent of Sydney compared to the rest of NSW.
- Then the number of poor in each area is adjusted until they add up to the GCCSA estimate.
- Finally, the number of poor in each group is readjusted so the total number of poor per small area is consistent.

Poverty rates, and the proportion of people in poverty with certain characteristics, were calculated for these characteristics:

- Overall poverty rate
- Poverty rate by gender
- Poverty rate by Age
- Children aged under 15
- Adults aged 15 to 24
- Adults aged 25 to 64
- Adults aged 65 years and over
- Poverty rate by family type
- Couples without dependent children
- Couples with dependent children
- Lone parents with dependent children
- Lone individuals without dependent children
- Poverty rate by employment status
- Adults employed full-time
- Adults employed part-time

- Unemployed adults
- Adults not in the labour force
- Poverty rate by housing tenure
- Home owners with no mortgage
- Home owners with a mortgage
- Private renters
- Public renters (i.e. in public housing)
- Other renters
- Other types of housing tenure
- Poverty rates for disabled people and for the indigenous population were not available from the small area estimation modelling due to technical difficulties. Therefore, we used low income rates for these groups, as described in the report.

## VALIDATION

Validation of the small area estimates was done in two ways:

1. The proportion of areas for which we get convergence;
2. A comparison of small areas estimating an indicators that can be derived from the Census;

One of the factors analysed from our spatial microsimulation model was the percentage of areas that provide reasonable estimates given a number of benchmarks. Reducing the number of benchmarks then means that the model works (converges), but the estimates are not as accurate as we have used fewer benchmarks. At some point (in this report, fewer than 7 benchmarks), we decide that the estimate isn't accurate enough to be published. Areas where we cannot get reliable estimates are usually remote areas; or areas with very low population (e.g. industrial areas or national parks). The proportion of areas that have converged in this model are shown in Table 2. It can be seen that 9 benchmarks have been mostly used to get reliable estimates for small area in Sydney and the rest of NSW.

Table 2: Number of Benchmarks Used

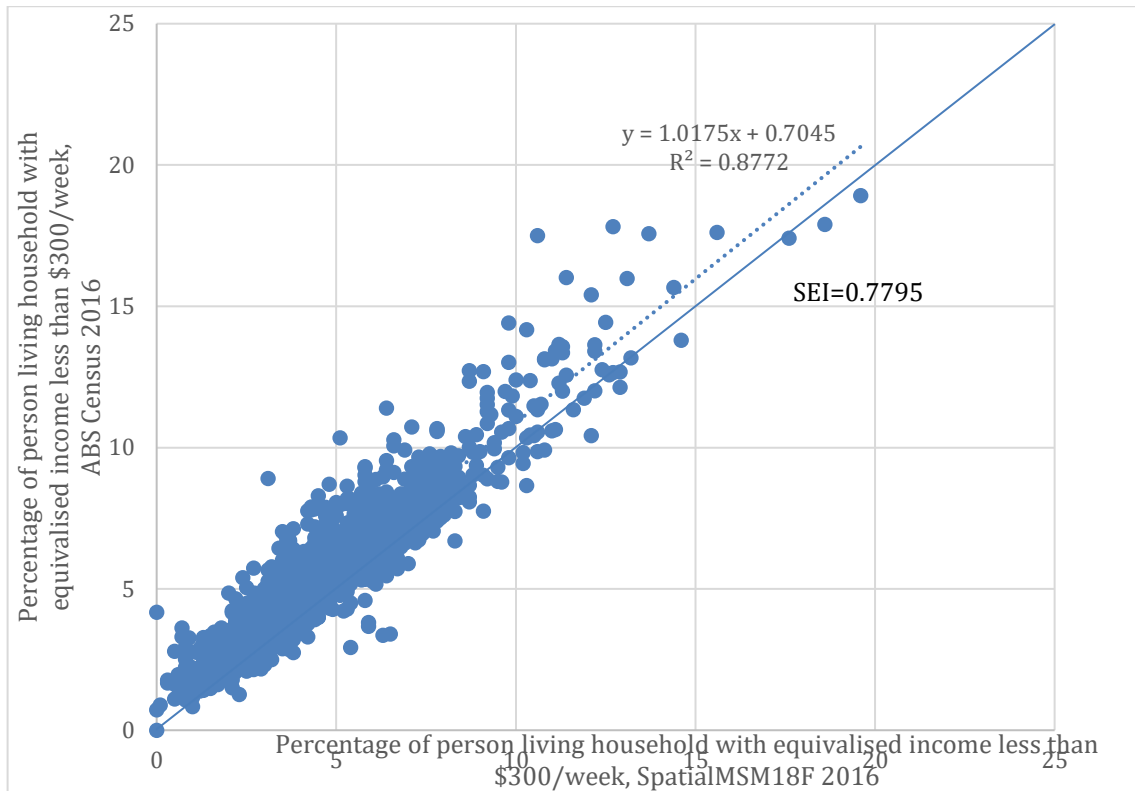
| GCCSA     | Number of Benchmarks used |        |       |        |        |        |           |
|-----------|---------------------------|--------|-------|--------|--------|--------|-----------|
|           | 3                         | 5      | 6     | 7      | 8      | 9      | 8 or more |
| 1GSYD     | 0.00%                     | 1.28%  | 1.28% | 1.92%  | 2.56%  | 92.95% | 95.51%    |
| 1RNSW     | 0.00%                     | 1.14%  | 0.38% | 1.89%  | 3.03%  | 93.56% | 96.59%    |
| 2GMEL     | 0.32%                     | 0.32%  | 0.65% | 1.62%  | 1.94%  | 95.15% | 97.09%    |
| 2RVIC     | 0.00%                     | 1.31%  | 0.65% | 0.00%  | 3.92%  | 94.12% | 98.04%    |
| 3GBRI     | 0.00%                     | 0.85%  | 0.85% | 1.27%  | 4.24%  | 92.80% | 97.03%    |
| 3RQLD     | 0.00%                     | 4.11%  | 1.71% | 3.42%  | 4.11%  | 86.64% | 90.75%    |
| 4GADE     | 0.00%                     | 0.00%  | 0.91% | 0.00%  | 1.82%  | 97.27% | 99.09%    |
| 4RSAU     | 0.00%                     | 3.23%  | 0.00% | 3.23%  | 1.61%  | 91.94% | 93.55%    |
| 5GPER     | 0.00%                     | 1.73%  | 1.16% | 0.00%  | 2.31%  | 94.80% | 97.11%    |
| 5RWAU     | 0.00%                     | 8.86%  | 7.59% | 3.80%  | 20.25% | 59.49% | 79.75%    |
| 6GHOB     | 0.00%                     | 0.00%  | 0.00% | 0.00%  | 5.71%  | 94.29% | 100.00%   |
| 6RTAS     | 0.00%                     | 0.00%  | 1.56% | 1.56%  | 0.00%  | 96.88% | 96.88%    |
| 7GDAR     | 0.00%                     | 4.55%  | 9.09% | 13.64% | 22.73% | 50.00% | 72.73%    |
| 7RNTE     | 0.00%                     | 62.50% | 8.33% | 8.33%  | 4.17%  | 16.67% | 20.83%    |
| 8ACTE     | 0.00%                     | 3.82%  | 4.58% | 3.05%  | 4.58%  | 83.97% | 88.55%    |
| Australia | 0.04%                     | 2.53%  | 1.62% | 2.05%  | 4.02%  | 89.73% | 93.75%    |

Note: G means Greater (Capital Cities Areas); R means the Remainder (of the State/Territory)

Based on this result we decided to use the estimate produced using 7, 8 or 9 benchmarks. Areas where results could not be derived using less than 7 benchmarks were removed. A list of removed areas is shown in Appendix 1.

Another method to validate estimates at the small area level was to use the standard error around identity (SEI) (Edwards and Tanton 2013). Note that another validation method based on aggregating estimated poverty rates from the model and comparing these to direct estimates from the survey cannot be applied since the estimates have been aligned, so there is no difference between the aggregated small area estimates and the survey estimate for the State. To validate the small area estimates, we have calculated the proportion of people living in a household with equivalised income less than \$300 a week from both the Census and from the model (SpatialMSM18F). Figure 1 indicates that we have achieved a reasonably close estimate (0.8722 R squared and 0.7795 SEI).

Figure 1: Validation of proportion of persons living with equivalised income less than \$300/week  
(Spatial MSM and Census data)



## APPENDIX 1

### LIST OF EXCLUDED AREAS FOR THE POVERTY RATE CALCULATION AND REASONS FOR EXCLUSION

#### Due to low estimation accuracy (< 7 benchmarks)

Lord Howe Island

Greenfield Park - Prairiewood

Kingsford

Lurnea - Cartwright

Canley Vale - Canley Heights

Far West

Badgerys Creek

Royal National Park

Chullora

Ettrema - Sassafras - Budawang

Blue Mountains - North

Deua – Wadbilliga

#### Additional exclusion due to low population

Illawarra Catchment Reserve

Newcastle Port - Kooragang

Prospect Reservoir

Banksmeadow

Port Botany Industrial

Wetherill Park Industrial

Yennora Industrial

Wollangambe - Wollemi

Port Kembla Industrial

Sydney Airport

Centennial Park

Holsworthy Military Area

Blue Mountains - South

Rookwood Cemetery

Smithfield Industrial

#### Additional exclusion due to low number of unemployed people

Lilli Pilli - Port Hacking - Dolans Bay

Bombala

#### Additional exclusion due to low number of people in public housing

Terrey Hills - Duffys Forest

Lilli Pilli - Port Hacking - Dolans Bay

Double Bay - Bellevue Hill

Loftus - Yarrawarra

Pymble

Avalon - Palm Beach

Dover Heights

Turrumurra

West Pennant Hills

Lindfield - Roseville

St Ives

Putney

Acacia Gardens

Winston Hills

Bayview - Elanora Heights

Castle Hill - East

Cherrybrook

Mulgoa - Luddenham - Orchard Hills

Erskine Park

Castle Hill - Central

North Rocks

Gordon - Killara

Cecil Hills

Caringbah South

Wamberal - Forresters Beach



Heathcote - Waterfall  
Berowra - Brooklyn - Cowan  
Kiama Downs - Minnamurra  
Southern Highlands  
Warriewood - Mona Vale  
Glenwood  
Homebush Bay - Silverwater  
Cobbitty - Leppington  
Queanbeyan Region  
Kurrajong Heights - Ebenezer  
Avoca Beach - Copacabana  
Rouse Hill - Beaumont Hills  
Yarramundi - Londonderry  
Hill Top - Colo Vale  
The Oaks - Oakdale  
Woolaware - Burraneer  
West Wallsend - Barnsley - Killingworth  
Woronora Heights  
Galston - Laughtondale  
Pitt Town - McGraths Hill  
Bilpin - Colo - St Albans  
Castle Hill - South  
Bangalow  
Korora - Emerald Beach  
Newport - Bilgola  
Calga - Kulnura  
Jindabyne - Berridale  
Tuggerah - Kangy Angy  
Yass Region  
Broulee - Tomakin  
Valentine - Eleebana  
Branxton - Greta - Pokolbin  
Kiama Hinterland - Gerringong  
Saratoga - Davistown  
Dubbo Region  
Tea Gardens - Hawks Nest  
Redhead  
Warragamba - Silverdale  
Horsley - Kembla Grange  
Seaham - Woodville  
Sydenham - Tempe - St Peters  
Kellyville  
Summerland Point - Gwandalan  
Coramba - Nana Glen - Bucca  
Dural - Kenthurst - Wisemans Ferry  
Glenhaven  
Wagga Wagga - North  
Wangi Wangi - Rathmines  
Anna Bay  
Lennox Head - Skennars Head  
Bombala  
Jilliby - Yarramalong  
Horsley Park - Kemps Creek  
Orange Region  
Sussex Inlet - Berrara  
Terranora - North Tumblegum  
Bega-Eden Hinterland  
Cessnock Region  
Maitland - North  
Singleton Region  
Port Macquarie Region  
Moama  
Lemon Tree Passage - Tanilba Bay  
Austral - Greendale  
Tamworth Region  
Castle Hill - West  
Robertson - Fitzroy Falls  
Box Head - MacMasters Beach  
Lake Munmorah - Mannering Park  
Bonnells Bay - Silverwater

Tomerong - Wandandian - Woollamia

Bargo

Oatley - Hurstville Grove

Gunnedah Region

Bathurst Region

Old Bar - Manning Point - Red Head

Culburra Beach

Ulladulla Region

Callala Bay - Currarong

Bulahdelah - Stroud

Braidwood

Muswellbrook Region

Tumbarumba

South West Rocks

Wahroonga (East) - Warrawee

Kempsey Region

Murwillumbah Region

Grenfell

Grafton Region

Corowa Region

Deniliquin Region

Taree Region

Forster-Tuncurry Region

Inverell Region - West

Cooma Region

Eurobodalla Hinterland

Casino Region

Dorrigo

Mudgee Region - East

Tumut Region

Walcha

Woollahra

Rose Bay - Vaucluse - Watsons Bay

## LIST OF EXCLUSIONS FOR THE LOW INCOME HOUSEHOLD CALCULATION

### **Due to low population with valid household income**

Illawarra Catchment Reserve

Newcastle Port - Kooragang

Prospect Reservoir

Banksmeadow

Port Botany Industrial

Wetherill Park Industrial

Royal National Park

Blue Mountains - North

Deua - Wadbilliga

Yennora Industrial

Wollangambe - Wollemi

Port Kembla Industrial

Sydney Airport

Centennial Park

Holsworthy Military Area

Blue Mountains - South

Rookwood Cemetery

Smithfield Industrial

### **Additional exclusion due to low number of Aboriginal people**

Terrey Hills - Duffys Forest

Lilli Pilli - Port Hacking - Dolans Bay

Pymble

St Leonards - Naremburn

Turrumurra

West Pennant Hills

Willoughby - Castle Cove - Northbridge

Lindfield - Roseville

St Ives

Putney

Acacia Gardens

Castle Hill - East

Lord Howe Island

Chatswood (West) - Lane Cove North

Castle Hill - Central

North Rocks

North Ryde - East Ryde

Gordon - Killara

Cecil Hills

Sydney - Haymarket - The Rocks

Strathfield South

Chatswood (East) - Artarmon

Auburn - North

Woronora Heights

Galston - Laughtondale

Hunters Hill - Woolwich

Homebush

Waitara - Wahroonga (West)

Berala

Paddington - Moore Park

Glenhaven

North Sydney - Lavender Bay

Castle Hill - West

Robertson - Fitzroy Falls

Regents Park

Auburn - Central

Eastwood - Denistone

Auburn - South

Woollahra

Wiley Park

Rose Bay - Vacluse - Watsons Bay

Badgerys Creek

Chullora

Ettrema - Sassafras – Budawang

**Additional exclusion due to low number of people with disability**

Lord Howe Island

Badgerys Creek

Chullora

Ettrema - Sassafras – Budawang

**Additional exclusion due to low number of people from culturally and linguistically diverse community**

Deua - Wadbilliga

Grenfell

Wollangambe - Wollemi

Port Kembla Industrial

Illawarra Catchment Reserve

Lord Howe Island

Walcha

Newcastle Port - Kooragang

Ettrema - Sassafras - Budawang

Prospect Reservoir

Banksmeadow

Port Botany Industrial

Sydney Airport

Centennial Park

Holsworthy Military Area

Blue Mountains - North

Blue Mountains - South

Rookwood Cemetery

Smithfield Industrial

Yennora Industrial

Badgerys Creek

Wetherill Park Industrial

Royal National Park