



Pushing them to the edge:

**AN ASSESSMENT OF SPATIAL MICROSIMULATION
METHODS**

Paper Presented at the 2nd
International Microsimulation
Association Conference,
Ottawa, Canada, 8 – 10 June
2009

PREPARED BY

Robert Tanton and Yogi Vidyattama

ABOUT NATSEM

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

NATSEM aims to be a key contributor to social and economic policy debate and analysis by developing models of the highest quality, undertaking independent and impartial research, and supplying valued consultancy services.

Policy changes often have to be made without sufficient information about either the current environment or the consequences of change. NATSEM specialises in analysing data and producing models so that decision makers have the best possible quantitative information on which to base their decisions.

NATSEM has an international reputation as a centre of excellence for analysing microdata and constructing microsimulation models. Such data and models commence with the records of real (but unidentifiable) Australians. Analysis typically begins by looking at either the characteristics or the impact of a policy change on an individual household, building up to the bigger picture by looking at many individual cases through the use of large datasets.

It must be emphasised that NATSEM does not have views on policy. All opinions are the authors' own and are not necessarily shared by NATSEM.

Director: Ann Harding

© NATSEM, University of Canberra 2009

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.

National Centre for Social and Economic Modelling
University of Canberra ACT 2601 Australia
170 Haydon Drive Bruce ACT 2617

Phone + 61 2 6201 2780

Fax + 61 2 6201 2751

Email natsem@natsem.canberra.edu.au

Website www.natsem.canberra.edu.au

CONTENTS

About NATSEM	2
Author note	4
Author note	4
Acknowledgements	4
General caveat	4
Summary	5
1 Introduction	6
2 Data and Methods	7
2.1 Data	7
2.2 Methods	9
2.2.1 SpatialMSM/08C	9
2.3 Measures of accuracy	10
2.4 Methodological changes made	12
2.4.1 Adding additional benchmarks	12
2.4.2 Univariate Benchmarks	13
2.4.3 Limiting the source of households for the microsimulation	14
3 Results and Analysis	15
3.1 Adding additional benchmarks	15
3.2 Using univariate benchmarks	16
3.3 Limiting the source of households	17
4 Conclusions	20
Bibliography	21

AUTHOR NOTE

Robert Tanton is Acting Research Director of the Social Inclusion and Small Area Modelling team at the National Centre for Social and Economic Modelling (NATSEM) at the University of Canberra. Yogi Vidyattama is a Senior Research Officer at NATSEM.

ACKNOWLEDGEMENTS

This paper has been funded by a Linkage Grant from the Australian Research Council (LP775396), with our research partners on this grant being the NSW Department of Community Services; the Australian Bureau of Statistics; the ACT Chief Minister's Department; the Queensland Department of Premier and Cabinet; Queensland Treasury; and the Victorian Departments of Education and Early Childhood and Planning and Community Development. We would like to gratefully acknowledge the support provided by these agencies.

GENERAL CAVEAT

NATSEM research findings are generally based on estimated characteristics of the population. Such estimates are usually derived from the application of microsimulation modelling techniques to microdata based on sample surveys.

These estimates may be different from the actual characteristics of the population because of sampling and nonsampling errors in the microdata and because of the assumptions underlying the modelling techniques.

The microdata do not contain any information that enables identification of the individuals or families to which they refer.

SUMMARY

This paper will look at how well a spatial microsimulation model handles additional constraints, and how results from univariate constraint tables rather than multivariate constraint tables compare. This paper will also test how well non-Capital city households from a survey can estimate areas within capital cities. The argument for this approach is that the spatial microsimulation method has more households to choose from to represent the constraints in the area being estimated. In theory, this should improve the fit of the model. However, the alternative argument that could be made is that a household from another area may not be representative of households in the area being estimated – so, for example, in Australia, households in remote New South Wales will not be representative of households in Sydney. This paper aims to check the appropriateness of using households from anywhere in Australia to estimate SLAs in capital cities by running each model both ways (so using all households to estimate small areas in Sydney; and then using Sydney households to estimate small areas in Sydney), and comparing and contrasting the results.

The data used in the model is Australian Census and survey data from the Australian Bureau of Statistics (ABS).

Key words

Spatial microsimulation, Small area methodology

1 INTRODUCTION

In recent years, there has been increasing recognition of the importance of regional science in fields such as economics and human geography. This has meant an increased need for small area statistics. The need for small area statistics often cannot be addressed using direct estimates from survey data because most surveys use samples that are designed to provide reliable information for national or at most state level estimates, but nothing smaller. As a result it is usually impossible to derive estimates for small areas using sample surveys, and to derive a sample that would allow estimates for small areas would be inefficient (Heady et al. 2003).

This unmet demand for small area statistics has led to an increasing number of methods to derive these estimates. These methods are summarised in a number of papers (Ghosh and Rao 1994; Pfeffermann 2002; ABS 2006), and include simple ratio estimation, right through to random effects and Bayesian models.

Another technique that has emerged for small area estimation is spatial microsimulation. Microsimulation uses person, family or household level microdata to model real life individual conditions. Spatial microsimulation uses the microdata to estimate the condition of persons, families or households in a specific small area. Gonzales argued that the direct estimator from a survey can be a reliable estimator for a smaller area under the assumption that the small area has similar characteristics to the larger area for which the direct estimate is reliable (Gonzales 1973). Spatial microsimulation goes further in ensuring that the microdata used will represent the right characteristics of the small area by applying constraints or benchmarks, which are the characteristics of that small area, in the estimation process. This is done by populating a specific small area using persons, families or households from survey data based on census data that provides an accurate picture of the population in a small area.

There are a number of techniques that can be used for spatial microsimulation, but all use the basic idea of using benchmarks from a Census to populate the small area. The most common technique for spatial microsimulation is a reweighting technique, and there are a number of reweighting techniques available (Voas and Williamson 2000; Ballas et al. 2005; Anderson 2007; Tanton 2007).

This paper attempts to test one spatial microsimulation model (a model that reweights survey data using a generalised regression technique) to find out whether the model continues to give good estimates as the complexity of the model is increased. In this paper, the complexity of the model is increased by increasing the number of benchmarks.

This paper also tests how well a model with univariate benchmarks compares to a model with multivariate benchmarks. Using multivariate benchmarks allows a model to be constrained on marginal totals (so the total number of people with a certain income and rent), which should give better estimates in the final estimation process.

The third aspect of spatial microsimulation that this paper tests is whether survey sample observations which are very different to the characteristics of the small area will bias the

estimation for the small area. This is really testing whether someone from Sydney can be used to estimate an area outside Sydney (for instance, remote Western Australia).

The model tested in this paper, SpatialMSM, is a spatial microsimulation model that has been developed to estimate small area statistics in Australia. Besides estimating small area statistics, this model has also been linked to another microsimulation model to estimate the effect of changes in Government policy on small areas in Australia. The model has been used in Australia to fulfil the need for reliable small area data for research. The SpatialMSM model employs a generalised regression reweighting program from the Australian Bureau of Statistics' (ABS) called GREGWT. The GREGWT algorithm uses a generalised regression technique to create initial weights and iterates the estimation until the Microdata produce an overall characteristic that closely resembles the constraints for the small area. It is also used by the ABS to reweight their surveys to Australia wide and capital city benchmarks.

In this paper, Section 2 outlines the data and methods in detail, Section 3 provides results and analysis, and Section 4 provides conclusions.

2 DATA AND METHODS

2.1 DATA

This section describes the data that the model uses. The survey data used comes from two surveys – the 2002-03 and 2003-04 ABS Survey of Income and Housing (SIH) Confidentialised Unit Record Files (CURFs). These two survey files are combined to maximise the sample size available for the modelling.

The second source of data is the Australian Census of Population and Housing. The Australian Census is conducted every five years, and covers every resident in Australia. It therefore provides reliable estimates of socio-demographic variables for small areas. The latest Australian Census is for 2006. The Census data is used for the benchmark tables.

There are 11 Census benchmark tables used in this version of the spatial microsimulation model (SpatialMSM/08C), and these are shown in Table 1 below. The Census benchmark tables are derived from either standard output tables from the Census available through the ABS (Basic Community Profiles and Expanded Community Profiles) or special data requests from the ABS which we developed where the information was not available from ready made ABS tables.

Table 1 Benchmarks used in the procedures

Number	Benchmark
1	Age by sex by labour force status
2	Total number of households by dwelling type (Occupied private dwelling/Non private dwelling)
3	Tenure by weekly household rent
4	Tenure by household type
5	Dwelling structure by household family composition
6	Number of adults usually resident in household
7	Number of children usually resident in household
8	Monthly household mortgage by weekly household income
9	Persons in non-private dwellings
10	Tenure type by weekly household income
11	Weekly household rent by weekly household income

Source: ABS Census of Population and Housing, 2006

Given that the two surveys and the census were conducted at a different point in time, there are some adjustments needed so that the survey data was compatible with the Census data. First, the incomes from the surveys had to be uprated to 2006 dollar values, using changes in average weekly earnings. Second, the weekly household rent and mortgage were uprated to 2006 dollars using the housing component of the ABS Consumer Price Index (CPI).

Other adjustments to make the survey and Census compatible included removing non-classifiable households (for example, households which contain no persons over 15 or which contain visitors only) from several of the Census tables, as non-classifiable households were not on the survey dataset. We also added people in non-private dwellings to the survey dataset, as they were in the Census data, and we wanted to be able to keep them in for analysis of older people in non-private dwellings (in particular, nursing homes). The information on people in non-private dwellings came from the Census household sample file, which is a 1 per cent random sample from the Census. The household sample file used was the 2001 Census, as the 2006 file was not available when this work was done.

The Statistical Local Area (SLA) is the spatial unit used in this paper. The SLA is one type of standard spatial unit derived by the ABS and described in the Australian Standard Geographic Classification 2006 (ABS 2007a). There were two main reasons why the SLA was used as the unit of analysis in this study. First, the SLA is the smallest unit in the ASGC where there are no substantial issues with confidentiality. The ABS randomises any cells in tables where the number of people is less than 3, and as an area gets less populous, the chance of getting too many randomised cells increases. Second, SLAs cover the whole of Australia (as opposed to Local Government Areas which do not cover areas with no local government) and cover contiguous areas (unlike some postcodes) (McNamara et al. 2008).

2.2 METHODS

2.2.1 SpatialMSM/08C

The reweighting process in SpatialMSM uses an iterative constrained optimisation technique to calculate weights that will, when applied to the survey data, provide the best estimates of the Census Benchmarks. The technique uses a calibration estimator initially outlined by Singh and Mohl (Singh and Mohl 1996) and described and implemented by the ABS (Bell 2000) in a SAS macro called GREGWT. The SAS macro program is commonly used within the Australian Bureau of Statistics to benchmark survey datasets to known population targets, generally at the national or state level. In contrast, SpatialMSM uses this process to create a synthetic household microdata file for each Statistical Local Area (SLA) in Australia, containing a set of synthetic household weights which replicate, as closely as possible, the characteristics of the real households living within each small area in Australia (Chin and Harding 2007).

Because the reweighting process is an iterative process, there will be areas where the procedure will not find a solution. If there is no solution found after a number of iterations (which can be set by the user and for SpatialMSM is set at 30), then the process has not converged. Those SLAs where the process does not converge are usually SLAs where the population is quite different to the sample population – so for instance, industrial estates or inner city areas. For many areas, however, we found that the original GREGWT criteria for non-convergence was too strict: even after iterating 30 times and not converging, the estimate we got from the weights was still reasonable when compared with the benchmarks. In order to maximise the number of SLAs for which we could produce valid data, SpatialMSM uses the total absolute error (TAE) from all the benchmarks as a new criteria for reweighting accuracy. This measure was developed by Paul Williamson for a combinatorial optimisation reweighting method (Williamson et al. 1998), and is calculated as the sum of the absolute differences between the estimated population and the actual population in each category of each benchmark table for every SLA. The TAE will be 0 if all the benchmarks in the SLA are matched perfectly, and will increase as the estimation procedure fails to meet the benchmarks. A ‘failed’ TAE will depend on the population of the SLA – so for an SLA with a population of 100, a TAE of 50 is bad; but for an SLA with a population of 10,000, a TAE of 50 is good. So the criteria we use in this paper is that if the TAE divided by the population of the area is greater than 1 then the area has a failed accuracy, and is dropped from further analysis. In this paper, this is called a failed accuracy criteria (rather than non-convergence).

Using SpatialMSM/08C, we have been able to produce weights for 1214 SLAs. There were 138 SLAs where the method did not appear to work, and this was shown in the failed accuracy criteria. These SLAs have been dropped from further analysis. We found that most of the SLAs with failed accuracy criteria were industrial areas, office areas or military bases with very low population size. Therefore, the proportion of persons living in these SLAs is very small (Table 2). Only 0.7% of the total Australian population in 2006 were lost due to the failed accuracy criteria. Having said this, the process did not work for many areas in the Northern Territory, and 25 per cent of the Northern Territory population had

to be dropped due to failed accuracy. Therefore, small area estimates for the Northern Territory from SpatialMSM/08C should be treated cautiously.

Table 2 Number of SLAs dropped due to failed accuracy criteria

State/Territory	SLAs with failed accuracy	Total SLAs	Percent of SLAs with failed accuracy	Percent of population in SLAs with failed accuracy
NSW	2	200	1.0	0.4
VIC	4	210	1.9	0.0
QLD	43	479	9.0	0.8
SA	7	128	5.5	0.4
WA	17	156	10.9	0.9
TAS	1	44	2.3	0.1
NT	48	96	50.0	25.2
ACT	16	109	14.7	1.0
Australia	138	1422	9.7	0.7

Source: SpatialMSM/08C

2.3 MEASURES OF ACCURACY

While the TAE gives us some idea of whether the SLA converged, for those SLAs that have converged, we need to test whether the change in the model gave better or worse estimates, so we needed to have some measure of accuracy of the estimates. What we are interested in is some measure that is external to our model, and that we know is reliable for small areas.

Due to the nature of the benchmarking, we know that the model will estimate variables that are already benchmarked very well. These are called *constrained* variables. What the model needs to be able to do is estimate variables that have not been benchmarked, since the Census can provide reliable estimates of the benchmarked variables already.

The non-benchmarked (or *unconstrained*) variables that will be estimated reliably will need to be highly correlated with the benchmark variables, otherwise reasonable estimates cannot be provided. In some ways, the choice of output variable determines the choice of benchmarks. If poverty rates using equivalised disposable household income are required as the output variable, then variables like income, labour force status, housing tenure, and number of people in the household should be the benchmarks.

The unconstrained variable used for this testing was poverty rates. Because we are looking for a measure of how well our models are predicting small area poverty rates, we also need an accurate measure of poverty for small areas. We have therefore used Census data to calculate equivalised gross income for small areas in Australia. Because the Australian Census only has income available in groups, we have chosen a poverty line of \$400 per week. This was the closest group to the half median poverty line (\$281.50 per week in 2005/06 (ABS 2007b)) that we could get in the Census data we received from the ABS (which had already aggregated some income ranges for confidentiality).

A poverty indicator using equivalised gross household income and a poverty line of \$400 per week was then calculated using the same 2002-03 and 2003-04 income surveys on which the weights are based, with the incomes inflated to 2006 dollars using the change in average weekly earnings. We then applied the weights to this data to produce regional estimates of poverty rates. These spatially microsimulated poverty rates are therefore calculated in exactly the same way as the Census data poverty rates, and they are unconstrained (the benchmarks included income and number of adults/children resident, used for the equivalising process, in separate benchmark tables).

The next step is to calculate how far the microsimulated estimates are from the reliable Census estimates. To do this, we need to employ an accuracy criterion which compares our estimated per cent of people living in a household with equivalised gross income under \$400 per week to the reliable Census data from the ABS. To do this, we used a scatter plot of the rates from each data source. In theory, if the rates are exactly the same for each method, then all the data points will fall on the 45 degree line. The extent of dispersion from the 45 degree line can be calculated as:

$$MA = 1 - \frac{\sum (y_{est} - y_{ABS})^2}{\sum (y_{ABS} - \bar{y}_{ABS})^2}$$

Where

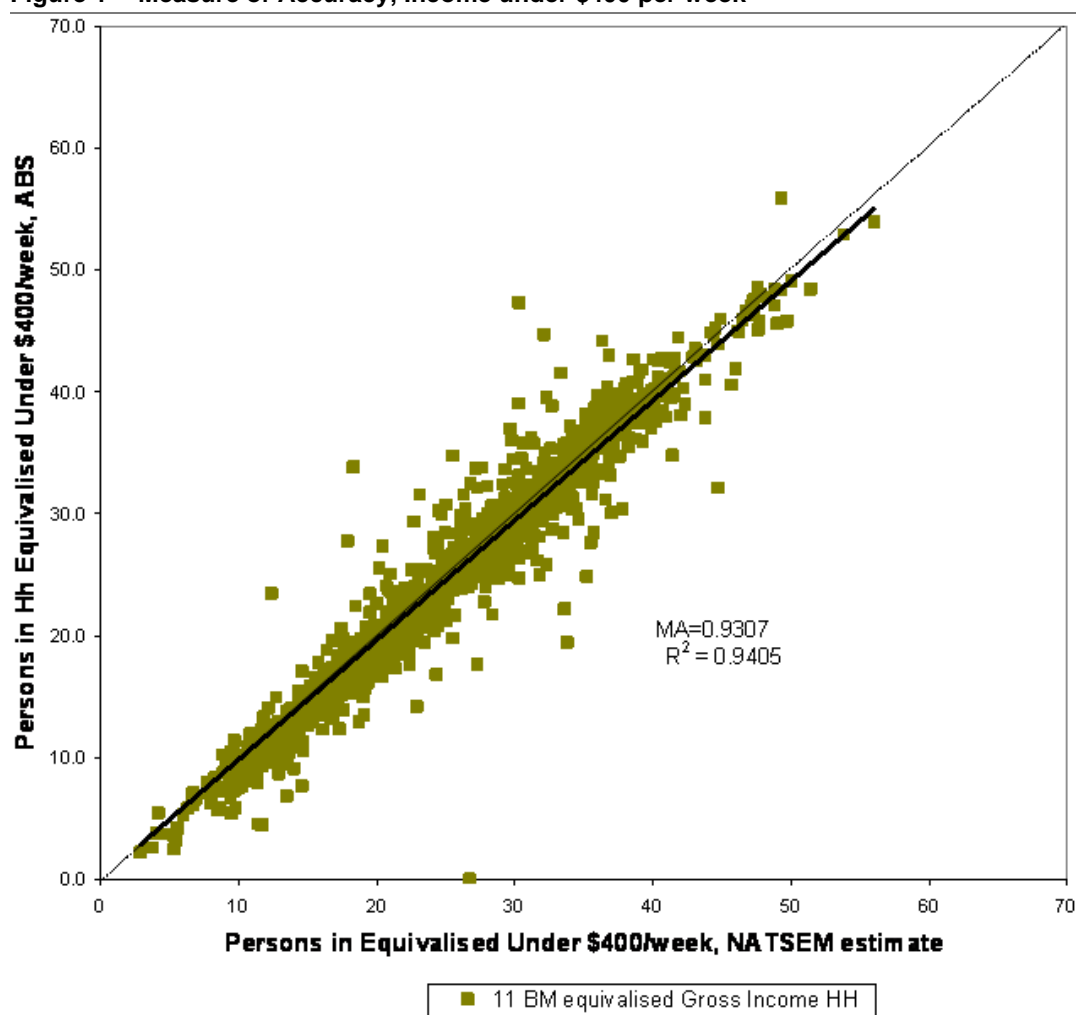
MA = *measure of accuracy*

y_{est} = *estimates of poverty rates from spatial microsimulation (gross income)*

y_{ABS} = *estimates of poverty rates from the ABS*

\bar{y}_{ABS} = *mean estimates of poverty rates from the ABS*

We have called this the measure of accuracy, as it is the dispersion of the rates around the 45 degree line. Note that this is different from the R squared used in a regression as this is the dispersion of the data points around a line of best fit. This is best seen in a graph, and Figure 1 shows this using SpatialMSM/08C. It can be seen that the points are all very close to the 45 degree line (shown in this graph as a thin line), so the MA is very high (0.9307).

Figure 1 Measure of Accuracy, Income under \$400 per week

Source: SpaitalMSM/08C applied to SIH 2002/03 and 2003/04

The interpretation of the MA is similar to the interpretation of the R squared. A higher MA is better and 100% or 1 is the highest accuracy that can be produced since it means that our estimate is exactly the same as the Census data.

One interesting point to note from Figure 1 is that the MA and R² are very close, and the 45 degree line (thin black line) and line of best fit (thick black line) are also very close. This means there is no systematic bias in the estimates. In fact, the SpatialMSM estimates have a very slight bias upwards.

2.4 METHODOLOGICAL CHANGES MADE

2.4.1 Adding additional benchmarks

In this test, two Benchmark tables will be added to the existing 11 and the impact of these additional tables will be analysed. The usual trade off in spatial microsimulation is that

when additional benchmarks are added, the procedure has greater trouble matching all the benchmarks, and fails to converge for a greater number of areas. However, with additional benchmarks, we can find that the accuracy of the final estimates increases, as there is more data being constrained to.

The aim of this exercise is to see whether it is possible to add additional benchmarks without losing too many areas due to a failed accuracy criteria, and then see how much the additional information affects the accuracy of the final estimates.

The first benchmark table that we have added is Non schooling qualification for people aged 15 years and above. Unfortunately, the classification used in the 2006 Census was different to the classification used in the Survey of Income and Housing, and the classification in the 2002/03 Survey of income and Housing is different to that used in the 2003/2004 survey. However, what we have been able to do is aggregate the classes up to as broad a group as possible, and this means they are defined in the same way for each dataset. This means we end up with only three education levels available for benchmarking, “Bachelor degree or higher, postgraduate”, “Other post school qualifications” that contains certificates and advanced diplomas and ‘No higher degree’.

The second benchmark table added was the Occupation of Employed person aged 15 and above. Similar to the new non-school education benchmark, there was a different classification used for occupation on the Census compared to the survey. In the 2006 Census, Occupation was coded to the 2006 Australian and New Zealand Standard Classification of Occupations (ANZSCO). Both the 2002/03 and 2003/04 survey of Income and Housing use the Australian Standard Classification of Occupations (ASCO) Second Edition to code occupation. There is an occupation classification mapping to allow the ANZSCO to be recoded to ASCO, and this was used to get all the occupation data into the same classification (ABS 2008).

2.4.2 Univariate Benchmarks

For all the benchmarks currently used, we specify cross tabulations, so we are constraining on a number of variables together. Another way to specify the benchmark tables would be as univariate tables, so there is only one variable in each table, rather than two or three. We would expect that because we are constraining to simpler tables, that there would be a greater level of convergence. However, we could also expect lower accuracy, as the multivariate benchmarks allow constraining to marginal totals (the total number of people in one category *given* another category). The question will be whether the greater convergence offsets the lower accuracy.

This second exercise will examine the impact of reconstructing the SpatialMSM/08C multivariate benchmark tables into several univariate tables. Of the 11 tables in SpatialMSM/08C, 7 are multivariate tables. These benchmark tables are Age by sex by labour force status (3 variables), Tenure by weekly household rent, Tenure by household type, Dwelling structure by household family composition, Monthly household mortgage

by weekly household income, Tenure type by weekly household income and Weekly household rent by weekly household income (all with two variables).

There will be 10 new univariate benchmarks tables constructed from those 7 multivariate benchmark tables. As a result we now have 14 univariate benchmark tables. Table 3 gives the list of the new benchmarks tables and their sequence in reweighting process.

Table 3 List of Univariate benchmarks

Number	Benchmark table
1	Labour force status
2	Age
3	Sex
4	Dwelling Type
5	Tenure type
6	Weekly household rent
7	Household type
8	Dwelling structure
9	household family composition
10	Number of adults usually resident in household
11	Number of children usually resident in household
12	Monthly household mortgage
13	Weekly household income
14	Persons in non-private dwelling

2.4.3 Limiting the source of households for the microsimulation

In the first and second exercise, we have pushed the SpatialMSM model to the edge by modifying the constraints or benchmark tables used in the process. The next exercise will push the ability of this model by using a limited set of the microdata in the reweighting process. In particular, this exercise will examine the effect of using households from a specific capital city to estimate small area statistics in that capital city.

This exercise is important to address the question that often comes up regarding reweighting methods for spatial microsimulation, which is whether it is acceptable to use households from all around Australia to represent households in a specific SLA in Sydney or Melbourne (or any other capital city). The question centres on whether using a household from rural Western Australia to estimate a household in Sydney introduces some systematic bias into the model.

Theoretically there are advantages as well as disadvantages in having the entire Australian dataset available for estimation. The main advantage is that there will be more households with different characteristics to weight to a specific SLA. On the other hand, we know that non capital-city households have different characteristics to households in capital cities, so they may not be appropriate to use for estimating SLAs in capital cities.

The way that the estimation process works in GREGWT is that benchmarks listed earlier are given a higher priority, so the estimation procedure may give a household with a certain income a high weight because it represents a high income household in a Sydney SLA, but then if the high income household was from a mining town in rural Western Australia, then they may be renting as a fly in-fly out miner, whereas everyone in the Sydney SLA owns their house. So the method gives a high weight because the income benchmark is met, but the other characteristics are very different, so the use of this household could bias the estimation.

In this exercise we will examine the result of using households from 5 specific capital cities: Sydney, Melbourne, Brisbane, Adelaide, and Perth. Using these 5 capital cities will provide us with some confidence in our results if they are consistent for each capital city. The cities and survey sample for these cities is quite different. Sydney as the most populated capital city had around 1.5 million households in 2006 while the microdata used in SpatialMSM/08c provided around 4000 households in the sample. In contrast, Adelaide is a less populated capital city and had around 450 thousand households, represented by around 2000 households in the sample.

3 RESULTS AND ANALYSIS

3.1 ADDING ADDITIONAL BENCHMARKS

As expected, the use of these additional benchmarks reduced the number of converging SLAs. Using only the non-school qualification table as an additional benchmark, the number of SLAs that passed our TAE test was down from 1284 to 1280, so there were only four less SLAs estimated with the additional benchmark table. Using the new occupation table as an additional benchmark reduced the number of SLAs passing our TAE test to 1262, so 22 fewer SLAs. Introducing both the non-schooling qualification and the occupation table as the twelfth and thirteenth benchmark tables provided only 1257 accepted SLAs to be analysed.

Although the results met our expectations in terms of reducing the number of SLAs that passed our TAE test, the impact of these additional benchmarks on the process is not as straightforward as we thought it might be. In some areas, an additional benchmark meant that an area that failed the TAE test using 11 benchmarks was now accepted, so the new benchmark improved the estimation of that SLA. Conversely, some areas that were accepted with 11 benchmarks failed in meeting our TAE criteria once we added another benchmark.

Most of the SLAs that were affected by the additional benchmark were either in rural areas with a population less than one thousand people or inner city areas. Adding education as a benchmark meant that 13 more SLAs failed our TAE criteria, although it also meant 9 SLAs that previously failed now passed pass our TAE criteria (giving the net change of four SLAs). From the 13 SLAs that now failed the TAE criteria, only 5 come from capital city areas. These were Anstead in Queensland, Hobart-Inner in Tasmania, Ludmilla in the

Northern Territory and Acton and Harrison in the Australian Capital Territory. On the other hand, adding the benchmark meant that some SLAs now passed the TAE criteria, and these included City-Inner Brisbane and Duntroon and Pialligo in the ACT.

A similar pattern appears when occupation is used as the additional benchmark. The number of SLAs that failed our TAE test increased by 25, while the number that now pass was three, giving the net change of 22 SLAs. Only 5 of the 25 SLAs that now fail the TAE test were in capital cities. These five SLAs were Nathan in Queensland, Hobart-Inner in Tasmania, Ludmilla in the NT and Acton and Hall in the ACT. None of the three SLAs that now passed the TAE criteria were from capital cities.

In terms of the accuracy of the results when compared to Census data, there was an expectation that additional benchmarks would increase the accuracy of the estimates. When we compared the estimates to the number of people who lived in a household with equivalised gross income under \$400 per week, we found that the estimates were in fact no better. The addition of the occupation benchmark does increase the measure of accuracy from 93.1 per cent to 94.1 per cent, but the addition of the non-school qualification benchmark reduces the accuracy to 92.6 per cent. Using both tables as additional benchmarks resulted in an accuracy of 93.9 per cent (Table 4).

This slight reduction in accuracy may be because we are still validating against a variable that was benchmarked to when using 11 benchmarks. So poverty (or income) has always been benchmarked to. If we were validating against educational status (so the new benchmark variable), then we could expect to get much better results using a set of benchmarks with education. Essentially what this suggests is that with the 11 benchmark model, we have the best estimates of poverty; but if we also wanted to use these weights for educational status (so to look at how many people with a higher degree were in poverty), then we would need the education benchmark.

Table 4 Summary of the impact of additional benchmarks

Model	SLAs with TAE < 1	SLAs with TAE >= 1	Measure of Accuracy
SPATIALMSM08c (11BM)	1284	138	0.9307
11BM + non school Qualification (NSQ) BM	1280	142	0.9268
11BM + Occupation (OCC) BM	1262	160	0.9411
11BM + NSQ + OCC BM	1257	165	0.9388

Source: SpatialMSM/08c applied to SIH 2002/03 and SIH2003-04

3.2 USING UNIVARIATE BENCHMARKS

As mentioned in section 2, we expect that using univariate instead of multivariate benchmarks will increase the number of converging SLAs since it will allow benchmarking to single variables, rather than benchmarking to marginal totals (so the total in one classification given another classification). The results from this exercise reported here confirm that expectation.

Using the 14 univariate benchmarks shown in Table 3 has increased the number of SLAs that passed our TAE criteria to 1329 compared to 1284 converging SLAs using 11 Benchmark tables with 7 of them being multivariate (Table 5). All 45 additional accepted SLAs failed our TAE criteria when 11 benchmarks were used, so we have brought in an extra 45 SLAs. Only 16 of these 45 SLAs are in capital cities. The 16 SLAs in capital cities included Sydney-Inner, Brisbane City-Inner, Perth-Inner, Fremantle-Inner, Darwin City-Inner and Canberra City (Civic), so they were inner city areas, which are usually particularly difficult to estimate because of the diverse nature of the population in inner city areas.

Because the procedure is now only benchmarking to single variables (so we are not benchmarking to marginal totals), the accuracy when compared to reliable Census poverty rates has reduced. We now have a measure of accuracy of 87.8 per cent, down from 93.1 per cent (Table 5). However, this may be due to the fact that there are more SLAs accepted under our TAE criteria. Using the SLAs that were accepted when we were using 11 Benchmarks, the accuracy of the estimation with the univariate benchmarks is around 91.0 per cent. So the reason why we get more SLAs converging using univariate benchmarks may be that the procedure is now including SLAs that cannot be estimated well, even though they were acceptable against the Census benchmarks, and this reduced the correlation with reliable Census data.

Table 5 Summary of the impact of using univariate benchmarks

Model	Accepted SLAs with TAE<1	SLAs with TAE >= 1	Measure of Accuracy
SPATIALMSM/08c (11BM)	1284	138	0.9307
Univariate BM	1329	93	0.8781
Univariate BM and 1284 SLAs converged in SPATIALMSM/08c			0.9100

Source: SpatialMSM08/c applied to 2002-03 and 2003-04 SIH

3.3 LIMITING THE SOURCE OF HOUSEHOLDS

The next exercise is to analyse the effect of using all households in the survey dataset to derive estimates for small areas that may be very different from the area that the survey respondent is in. For instance, in Australia, using a survey respondent from remote New South Wales to derive an estimate for Central Sydney.

This will be tested by looking at whether different results, in terms of the number of SLAs passing our accuracy criteria and the measure of accuracy, are achieved when a sub-population from the survey is used. The survey data allows us to identify where the respondent came from (capital city and State). This allows us to form a subset of the sample that consists of only people in Sydney, Melbourne, Brisbane, Adelaide and Perth. We then use this subset of the sample to estimate all SLAs in these capital cities; as well as using the

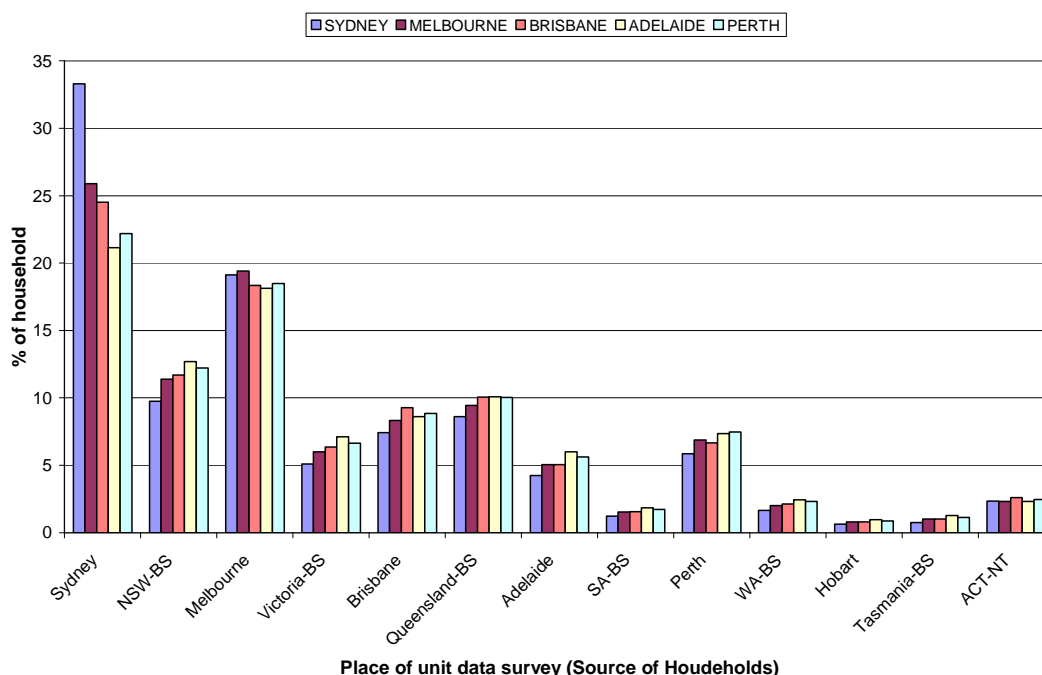
whole dataset to estimate SLAs in each of these cities. The results are then compared to see which sample gives better results.

Figure 2 shows how the sample for the 2002-03 and 2003-04 Surveys of Income and Housing are distributed. What this graph shows on the horizontal axis is the location that the respondent on the survey lived in (Sydney, NSW Balance of State, Melbourne, Victoria Balance of State, etc); and then the proportion of households used from this area to provide estimates for five capital cities in Australia (Sydney, Melbourne, Brisbane, Adelaide and Perth).

What the graph shows is that to estimate areas in Sydney (the blue bars), about 32 per cent of respondents came from Sydney; 10 per cent came from NSW – Balance of State; 19 per cent came from Melbourne; 5 per cent came from Victoria – Balance of State; about 7 per cent came from Brisbane; 8 per cent came from Queensland – Balance of State; 5 per cent came from Adelaide or Perth; and 2 – 3 per cent came from each of South Australia – Balance of State, Western Australia – Balance of State, Hobart, Tasmania – Balance of State and ACT-NT. All these add up to 100 percent of households.

Essentially, if we were using just ACT observations to estimate households in the ACT, we would be using far fewer households than if we use all households across Australia. So using all households gives a much more diverse set of households for the spatial microsimulation procedure to use for smaller States.

Figure 2 Source of Households to populate SLAs in SpatialMSM/08c in Five Capital Cities



Source: SpatialMSM/08c applied to 2003/04 and 2003/04 SIH

While this background information suggests that better results will be gained using all households, simply because it increases the number of households available to fill a small area, the exercise that confirms this would be to look at the measure of accuracy and the change in the number of SLAs with a total absolute error greater than 1 (so those with an unusable result) after using all households and households in the capital cities being estimated.

The results of the exercise using five capital cities in Australia are shown in Table 6. This table suggests that the results do not depend on the sample used. There is very little difference in the number of SLAs with a TAE less than one, and the Measure of Accuracy is only different for Perth.

For four out of five capital cities, using households from their own city has increased the accuracy slightly with little effect on the number of SLAs passing the TAE criteria. Melbourne is the only capital city where the results show a decrease in the accuracy when households from Melbourne are used to estimate Melbourne SLAs. Perth has the highest increase in accuracy with almost six percentage point increase. This is followed by Adelaide where the accuracy increased by two percentage points. The fact that the accuracy has increased more in the two most unpopulated capital cities used in this analysis shows that the Australian sample for the two smaller capital cities used (Adelaide and Perth) may give a slightly misleading result since the sample is dominated by households from the larger capital cities (see Figure 2).

Table 6 Effect of using households from each capital city to estimate areas in the capital city using spatial microsimulation

Source of data for estimation with SPATIALMSM/08c (11BM)	Accepted SLAs with TAE<1	SLAs with TAE >= 1	Measure of Accuracy
- Sydney for Sydney	63	1	0.9676
- Australia for Sydney	63	1	0.9618
- Melbourne for Melbourne	78	1	0.9263
- Australia for Melbourne	79	0	0.9511
- Brisbane for Brisbane	214	1	0.9263
- Australia for Brisbane	212	3	0.9224
- Adelaide for Adelaide	55	0	0.9735
- Australia for Adelaide	55	0	0.9534
- Perth for Perth	35	2	0.8478
- Australia for Perth	35	2	0.7856

Source: SpatialMSM/08c applied to 2003/04 and 2003/04 SIH

4 CONCLUSIONS

This paper has made a number of changes to a spatial microsimulation model to test the effect on the number of converging areas, and the accuracy of the estimates compared to reliable Census data. The aim is to test the reliability and stability of this model.

What we have found is that the spatial microsimulation model using GREGWT is very stable. We tend to get very similar results in terms of the measure of accuracy when we add benchmarks or limit the sample being used in the estimation. We have also found that using univariate benchmarks gives us more SLAs which pass our $TAE < 1$ criteria, but at a significant cost in terms of the accuracy of the model.

The two benchmarks that we have added in this paper did not have a huge effect on the number of SLAs with a $TAE < 1$, but did decrease the measure of accuracy slightly when using poverty rates. This may be different if we were using a variable that was correlated with the new benchmark being added to determine the measure of accuracy. So we might show that an estimate of the number of people receiving Austudy (a government allowance for students or the number of people with a Year 10 education in poverty) was much better using the new benchmark.

The advantage of adding benchmarks is that the weights become more general, so they can be used to estimate a wider range of variables. The model with education as a benchmark can be used to estimate poverty rates, housing stress, and Austudy recipients; whereas the model without the education benchmark would only provide reasonable estimates of poverty and housing stress, as there is no education benchmark.

We found that simplifying the benchmarks by creating a number of univariate tables gave many more useable SLAs (as shown by more SLAs with a $TAE < 1$), but the level of accuracy reduced. So there were advantages in benchmarking to more complicated bivariate tables.

In terms of the theory that using all households in a survey will give worse estimates for small areas, we find that the effect of using all households in Australia on the number of SLAs with a $TAE < 1$ and the measure of accuracy is very small. Using observations for the whole of Australia has a greater detrimental effect on the measure of accuracy for Adelaide and Perth, possibly because many of the observations in the survey come from the larger capital cities. One interesting result is that all the households gave a better estimate for Melbourne SLAs compared to only using Melbourne households for the estimation.

BIBLIOGRAPHY

- ABS 2006, A guide to small area estimation - Version 1.0,
[http://www.nss.gov.au/nss/home.NSF/533222ebfd5ac03aca25711000044c9e/3a60738d0abdf98cca2571ab00242664/\\$FILE/May%2006.pdf](http://www.nss.gov.au/nss/home.NSF/533222ebfd5ac03aca25711000044c9e/3a60738d0abdf98cca2571ab00242664/$FILE/May%2006.pdf), Accessed 30 July 2007.
- ABS 2007a, *Australian Standard Geographical Classification (ASGC)*, 1216.0, Australian Bureau of Statistics.
- ABS 2007b, *Household income and income distribution, Australia, 2005-06*, 6523.0, ABS.
- ABS 2008, *Census of Population and Housing: Link Between Australian Standard Classification of Occupations (ASCO) Second Edition and Australian and New Zealand Standard Classification of Occupations (ANZSCO)*, 2006, 1232.0, Australian Bureau of Statistics.
- Anderson, B. 2007, *Creating small area income deprivation estimates for Wales: Spatial microsimulation modelling* Chimera Working Paper 2007-11, Colchester, University of Essex.
- Ballas, D., Clarke, G.P. and Wiemers, E. 2005, Building a dynamic spatial microsimulation model for Ireland, *Population, Space and Place*, vol.11, no. 3, pp. 157-172.
- Bell, P. 2000, *GREGWT and TABLE macros - Users guide*, Unpublished, Australian Bureau of Statistics.
- Chin, S.F. and Harding, A. 2007, 'SpatialMSM - NATSEM's small area household model for Australia' in A. Gupta and A. Harding, (eds), *Modelling our future: Population ageing health and aged care*. Oxford, Elsevier, pp. 563 - 566.
- Ghosh, M. and Rao, J.N.K. 1994, Small Area Estimation: An Appraisal, *Statistical Science*, vol.9, no. 1, pp. 55.
- Gonzales, M.E. 1973, 'Use and Evaluation of synthetic Estimates', Proceedings of the Social Statistics Section, American Statistical Association, USA, pp. 33 - 36.
- Heady, P., PClarke, P., Brown, G., Ellis, K., Heasman, D., Hennell, S., Longhurst, J. and Mitchell, B. 2003, *Model-Based Small Area Estimation Series No. 2 - Small Area Estimation Project Report*, London, Office of National Statistics.
- McNamara, J., Harding, A., Daly, A. and Tanton, R. 2008, 'Child social exclusion: an updated index from the 2006 Census', paper presented at the Australian Institute of Family Studies Conference, Melbourne,
- Pfeffermann, D. 2002, Small area estimation - new developments and directions, *International Statistical Review*, vol.70, no. 1, pp. 125-143.
- Singh, A.C. and Mohl, C.A. 1996, Understanding calibration estimators in survey sampling, *Survey Methodology*, vol.22, no. 2, pp. 107 - 115.

- Tanton, R. 2007, 'The Australian Spatial Microsimulation Model', paper presented at the First General Conference of the International Microsimulation Association, Vienna, 20 - 22 August 2007.
- Voas, D. and Williamson, P. 2000, An evaluation of the combinatorial optimisation approach to the creation of synthetic microdata, *International Journal of Population Geography*, vol.6, pp. 349 - 366.
- Williamson, P., Birkin, M. and Rees, P. 1998, The estimation of population microdata by using data from small area statistics and samples of anonymised records, *Environment and Planning A*, vol.30, no. 5, pp. 785-816.