Considering regional socio-economic outcomes in non-metropolitan Australia: A typology building approach*

Scott Baum[1], **Michele Haynes**[2], **Yolanda van Gellecum**[2], **Jung Hoon Han**[2]

[1] Urban Research Program, Griffith School of Environment, Griffith University, 170 Kessels Road, Nathan, QLD 4111, Australia (e-mail: s.baum@griffith.edu.au)
[2] Social Science Research Centre, School of Social Science, Michie Building (9) Level 8, University of Queensland, St Lucia Qld 4072, Australia (e-mail: m.haynes@uq.edu.au, y.vangellecum@uq.edu.au, jh.han@uq.edu.au)

Received: 15 September 2004 / Accepted: 30 March 2006

Abstract. Australia's large regional cities and towns display wide variation in how they are adjusting to the socio-economic transitions that have occurred over the past decade. One area of research interest has been in developing typologies of non-metropolitan performance. The current paper represents an analysis of Australian Bureau of Statistics 2001 Census data aimed at analysing non-metropolitan regions based on their performance across a range of selected socio-economic variables. Using model-based clustering methods, this paper places non-metropolitan regions into clusters depending on the degree to which they share similar socio-economic and demographic outcomes. These clusters form the basis of a typology representing the range of socio-economic and demographic outcomes at the regional level. Differences between the clusters are analysed using graphs of 95% confidence intervals on the individual means for each cluster. The typology provides a useful framework with which to develop a broad understanding of socio-economic processes and performance across different spatial scales.

JEL classification: R11, R14

Key words: Non-metropolitan Australia, cluster analysis, typology, socio-economic conditions

* This research was funded from an Australian Research Council Discovery project 'Spatially Integrated Socio-Economic Analysis: Australia at the New millennium' DP0208102. The authors would like to thank the anonymous reviewers for their comments.

1 Introduction

The economic and social performance of rural and regional Australia has received increased attention in recent years. Like large metropolitan cities, non-metropolitan localities have faced transitions and change associated with wider national and international economic restructuring, demographic shifts and the realignment of public policy at various levels. As the role of regions in national economies has changed, there have been commensurate shifts in the socio-economic characteristics of individual localities, together with shifts in the ways these changes are understood. One theme explored within the Australian context relates to a crisis in rural and regional areas as forces including falling commodity prices, metropolitan centred social and economic policies and population migration have combined to increase non-metropolitan disadvantage and social malaise (Banks 2000). Debate on these negative outcomes have centred on accounts dealing with non-metropolitan changes with particular attention focusing on the loss of services and infrastructure in non-metropolitan regions as population loss and reduced cash flow all combine to accentuate a cycle of decline (Lawrence et al. 1996, 1999; Scott et al. 2001). Discussing the problem of decline in rural and agricultural-based localities, Tonts (1996) says "The economic, social and environmental changes which have affected Australian agriculture since the mid-1970s have given rise to increasing concern not only for sustainable family farming, but also for the continuing viability of country towns".

However, while some see these outcomes in terms of an aggregate non-metropolitan response to changing social and economic trends, and reference to a city-country dichotomy, the more likely outcome is that while some places within non-metropolitan Australia have witnessed a decline in social and economic terms, others are doing well. Authors including Burnley and Murphy (2004) and Beer et al. (2003) have illustrated how some localities in non-metropolitan Australia have witnessed a turnaround, and have been better able to reposition themselves in the face of a different economic situation. This is highlighted by Macadam et al. (2004) who point out that "Australia is a picture of many dead and dying towns and villages, some large and thriving towns and cities, and a few smaller ones that appear to be surviving against the odds . . . Evidence for change lies in the surviving and thriving communities that have taken their future into their own hands and have identified and used their natural advantages. These are a sign that community decay can be avoided."

In short, non-metropolitan Australia is characterised by localities and regions that are socio-economic winners and localities and regions that are socio-economic losers. A significant level of empirical work has focused on understanding the socio-economic performance of regions, cities and towns. Within the empirical work one approach, which has its foundations in the early sociological research on the social ecology of cities, and more recently in understanding the structure of post-industrial cities and urban regions (Coulton et al. 1996; Massey and Eggers 1993; Berry 1996; Baum et al. 2002; Mikelbank 2004), has been the development of typologies of non-metropolitan Australian cities and towns. Emerging from a need to understand and simplify complex processes, the use of typologies quantitatively identifies similarities and differences between observations (in this case non-metropolitan regions), classifies observations according to these outcomes and provides substantive analysis and understanding of the groups. The typologies are not meant to be explanations of processes per se, but are "an attempt to systemise classification in aid of explanation" (Marcuse 1997) and they provide a "richer understanding of complex phenomena" (Mikelbank 2004). It is the ability to elucidate the overall structure of localities and regions that makes these typology building exercises useful.

In terms of non-metropolitan Australia, the early work by Beer and others (Beer et al. 1994; Beer and Maude 1995; Beer 1999) is an example of a typology building approach which takes a range of indicators and uses a multivariate analytical approach to assemble ideal types or typologies of localities and regions which represent the broad nature of socio-economic patterns emerging. In the case of the research by Beer and his colleagues, the focus was on considering the functional classification of regional cities by classifying urban centres with populations greater than 10,000 into several clusters. The basis of the classification scheme was industry employment, with the outcomes illustrating that economic development was occurring over a range of regional city types resulting in a diverse group of non-metropolitan regions "with disparate economies and social structures" (Beer et al. 1994). A more recent example of this typology building approach was the research by Stimson (Stimson et al. 2001, 2003) and Baum (Baum et al. 1999) who considered broad indicators of socio-economic performance across all levels of the settlement system including large non-metropolitan cites, towns and regions. This research has identified, that the socio-economic performance of places is mixed with many localities being places of opportunity and others being places of vulnerability.

The current paper follows the work of Stimson et al. (2001, 2003) and Baum et al. (1999) and develops a typology of socio-economic performance across Australia's non-metropolitan cities, towns and regions. Specifically, the paper focuses on large non-metropolitan urban regions (population greater than 10,000 with greater than 50% of population characterised as urban) using a range of data available at a spatial level (Statistical Local Areas/SLA's) to develop a typology of socio-economic outcomes. The paper uses a clustering approach (MCLUST) to group localities into meaningful subgroups and then uses plots of confidence intervals on the means for each variable to determine the difference between the clusters. This allows us to develop a typology of cities, towns and regions in non-metropolitan Australia, taking into account varying levels of socio-economic performance. In what follows, we first discuss the methods and data used in the analysis, prior to discussing the typology developed. The paper concludes by discussing the implications of both the methodology used and the patterns identified.

2 Typology building using model based clustering and confidence intervals

The context for developing typologies is to cluster observations into groups sharing similar features and then to provide some understanding of the ways in which the groups differ. Several methods are available to cluster or partition data into meaningful sub-groups. Clustering methods range from approaches that are largely heuristic to more formal modelling procedures that adopt statistical models to group data. While heuristic approaches have been widely used to cluster spatially based data (see for example Hill et al. 1998; Baum et al. 1999) they are limited in their ability to clearly identify the most appropriate clustering method to use, as well as the true number of clusters in the data and the presence of outliers. The current paper uses a strategy for implementing cluster analysis based on parameterised Gaussian (normal) mixture models (Fraley and Raftery 2002a) which provides a statistical approach to addressing these issues.

A mixture model refers to a statistical model that uses a mixture or weighted sum of standard probability distributions to describe the distribution of observed data. In cluster analysis the data usually consists of independent multivariate observations and so the most commonly used mixture model is the mixture of multivariate Gaussian or normal probability distributions defined by a mean and covariance matrix, corresponding to each component of the mixture distribution. Finite mixture models are well suited to cluster analysis because each component probability distribution corresponds to a cluster. This allows statistical inference to be made about the components of the mixture model and hence probability statements about the classification of observations to a cluster. That is, it provides a measure of uncertainty about how well each observation is classified to a cluster (component of the mixture model). The number of clusters in the data is determined through a statistical model selection procedure where mixture models with differing numbers of components are fit to the data and the model that provides the best fit to the data is selected. The number of identified clusters then corresponds to the number of components in the selected model.

The issue relating to the choice of clustering method used is addressed through the specification of the covariance matrix for the components of the multivariate normal mixture distribution. In a review of model-based clustering methods, Fraley and Raftery (2002a) note "... some of the most popular heuristic clustering methods are approximate estimation methods for certain probability models". One example is the standard *k*-means clustering method that is an approximation to the estimation of a multivariate normal mixture model for which the covariance matrix is proportional to the identity matrix and is the same for all components in the model. This means that the clustering method is altered by changing the specification of the component covariance matrix in a mixture model. Both the number of clusters and the most appropriate clustering method are then determined using statistical model selection procedures.

The three stages in the clustering process using mixture models are:

- initialisation via model-based hierarchical agglomerative clustering
- maximum likelihood estimation of the mixture model using the EM algorithm, and
- selection of the model and the number of clusters using the Bayesian Information Criterion (BIC).

The MCLUST procedure, developed by Fraley and Raftery (Banfield and Raftery, 1993; Fraley and Raftery, 1999, 2002a, 2002b, 2003), is a software package for implementing this model-based clustering strategy through the statistical software S-PLUS and R (R Development Core Team, 2003). It includes functions that combine hierarchical clustering, EM algorithm for estimation of mixture models and the Bayesian Information Criterion (BIC) for model selection. It also provides visual graphics for displaying the clustering and classification results.

The procedure estimates Gaussian mixture models for a range of component sizes as well as various parameterisations of the covariance matrix for each mixture component. The six different parameterisations of the covariance matrix available in MCLUST consider the volume, shape and orientation of the clusters and are denoted:

1. EII: spherical, equal volume
2. VII: spherical, unequal volume
3. EEI: diagonal, equal volume, equal shape
4. VVI: diagonal, varying volume, varying shape
5. EEE: ellipsoidal, equal volume, shape and orientation
6. VVV: ellipsoidal, varying volume, shape and orientation.

Given the maximum likelihood estimates for the chosen mixture model, MCLUST produces the conditional probabilities that each observation belongs to the different groups associated with the components (clusters) of the mixture model. The final classification of an observation is made to the group which corresponds to the greatest conditional probability for that observation.

A distinctive advantage of a model-based clustering approach is that it allows the researcher to use model selection techniques such as the BIC to compare outcomes (Schwarz 1978). This gives a systematic means of selecting both the parameterisation of the model and also the number of clusters. By computing the BIC for the single cluster model for each parameterisation and for the mixture likelihood with the optimal parameters from EM for 2 through to M clusters a matrix of BIC values is produced. This provides a value for each possible combination of parameterisation and number of clusters. Additionally, to aid in interpretation the BICs are plotted for each model, allowing the researcher to determine the optimal clusters and model parameterisation. The ‘ideal’ cluster is that in which the BIC is highest and shows significant gain.

Apart from clustering the SLAs that make up the group of large non-metropolitan cities, towns and regions, the aim of the paper is also to consider how the clusters of localities differ from one another. Methods such as discriminant analysis have been used in the past with the means of clusters, combined with discriminant functions to consider how clusters differ (see for example Hill et al. 1998; Baum et al. 1999). However, discriminant analysis assumes that the correct groupings of observations are known and this is not the case when groupings have been determined using cluster analysis as there is uncertainty associated with the allocation of observations to clusters. Also, the clusters are formed based on the variables selected for the analysis only and so groupings may change when different variables are included in the analysis. An alternative method and the one chosen in this paper, is to adopt a visual data interpretation method using confidence intervals (Masson and Loftus 2003). Basically, the method incorporates the use of confidence intervals (CI) in conjunction with visual presentation to allow the researchers to form inferences about the cluster outcomes that take account of both the cluster mean and also the wider spread of the data. The confidence intervals are used in two ways. Firstly, clusters whereby the CI is clearly different from others without overlap are considered to be singularly differentiated on that particular variable. Secondly, in some cases, groups of clusters may have CI that overlap but which are above or below the mean for the entire population and variables for which this occurs can also be considered to differentiate the clusters from others. The interpretation of the cluster outcomes then becomes an exercise in comparing outcomes on the interpretation of CIs. By considering the CIs for each variable across the clusters developed using the MCLUST procedure, the researcher can begin to identify the typology that has emerged.

3 A typology of socio-economic outcomes across non-metropolitan Australia

The objective of this paper is to apply the typology building process discussed above to understand the socio-economic outcomes that have emerged in non-metropolitan Australia.

In developing the typology, a range of data was used. These data were associated with the region's economic performance, as they were expressed in residents' and individuals' characteristics and with socio-economic and socio-cultural characteristics of households and residents more generally. The variables used correspond to those found in research on the economic and social transformations of communities and localities and have been widely used elsewhere (see for example Hill et al. 1998; Baum et al. 1999; Stimson et al. 2001, 2003). The variables are set out in Table 1 and were transformed where appropriate using a log transformation to account for floor and ceiling effects imposed by using percentage data.[1] Further discussion of the variables adopted can be found in Baum et al. (2005).

Because the data used in the analysis came from several sources it was necessary to select a level of aggregation which could be used across different data collection agencies. For this purpose Australian Bureau of Statistics Statistical Local Areas (SLAs) were used. Across the Australian settlement system over 1,300 SLAs are available for analysis. In this paper, SLAs were chosen that: firstly, were outside the extended metropolitan regions (Baum et al. 2005); secondly, had populations greater than 10,000 persons; and, thirdly, had an urban population, as defined by the Australian Bureau of Statistics, of over 50%.

[1] The transformation method used was a log transformation from p to $\log(p/(1-p))$ where $p = P/100$.

Table 1. Variables used in the analysis

Socio-economic change	change in population change in employment	% change between 1991 and 2001
Occupational characteristics	educated professionals (1) vulnerable occupations (2)	1: % of persons with degree qualifications or above classified as managers, professionals or para-professionals 2: % of persons classified as labourers, tradespersons and basic clerical with out post school qualifications
Industry characteristics	new economy (1) old economy (1) mass goods and services (1) mass recreation (1) agriculture (1) mining (1) specialisation index (2)	1: % of persons employed in a given industry sector. Characterisation following O'Connor and Healy (2001) 2: illustrates degree of specialisation in industry. It measures the degree of specialisation or diversification. A score approaching 0 indicates increasing diversification while a score approaching 100 indicates increasing specialisation
Human capital	low formal human capital	% of persons who left school at year 10 (generally a minimum level of education)
Income/wealth	wage and/or salary (1) ratio of high income to low income (2) tax imputation (3) interest earned (3)	1: average wage and salary earned (Australian Tax Office) 2: ratio of % high individual income to % low individual income 3: imputation credits and interest earned (Australian Tax Office)*
Unemployment and labour force participation	labour force participation (1) adult unemployment rate (2) youth unemployment rate (3) part time workers (4)	1: % of persons in the labour force 2: % of persons aged 25 to 64 unemployed 3: % of persons aged 15 to 24 unemployed 4: % of part time employees
Household/family measures	non-earner families	% of families with children (couples and single parents) where no parent is employed
Housing	owner occupiers (1) rental financial stress (2) mortgage financial stress (3)	1: % of households who are owner occupiers 2: % of low income renters paying more than 30% of income on rent 3: % of low income home purchasers paying more than 30% on mortgage repayments

Note: Imputation credits (or tax credits) are essentially a credit back on tax. Taxpayers are required to pay tax on the dividend income received through owning shares. But, if an Australian company has already paid tax on its income, and then distributed the dividends, making the taxpayer pay tax on these dividends would be taxing the same profits a second time

Table 2. Bayesian Information Criterion (BIC)

	EII	VII	EEI	VVI	EEE	VVV
1	−43,884.9	−43,884.9	−9,725.4	−9,725.4	−7,697.8	−7,697.8
2	−41,369.3	−40,492.8	−9,266.8	−9,230.1	−7,723.4	−8,149.8
3	−39,376.2	−38,399.8	−8,972.5	−8,851.0	−7,784.5	−9,031.3
4	−37,994.4	−37,172.3	−8,824.9	−8,692.3	−7,844.9	NA
5	−37,491.1	−36,059.5	−8,705.3	−8,765.6	−7,862.7	NA
6	−36,706.6	−35,386.8	**−8,565.8**	−9,005.4	−7,909.6	NA
7	−36,510.6	−34,740.7	−8,597.3	−9,005.8	−7,962.7	NA
8	−36,413.2	−34,292.9	−8,606.9	−9,097.1	−8,028.1	NA
9	−36,390.9	−33,945.3	−8,642.8	−9,060.6	−8,122.6	NA
10	−34,526.2	−33,726.5	−8,646.8	−9,215.7	−8,219.4	NA

In addition, in some cases several smaller SLAs were combined to make more meaningful regions. This occurred in situations where larger regional centres were represented by several SLAs. This decision resulted in the inclusion of 18 derived regions (highlighted in Table 3a and 3b below) and a total of 119 localities.

The model-based clustering procedure and the BIC outcome (Table 2) lead to the selection of 6 clusters of localities (BIC = −8,565.8) using the EEI parameterisation (diagonal, equal volume, equal shape) for the component covariance matrix. Note that even though the BIC is highest for the EEE parameterisation with a single component, the BIC decreases consistently and does not achieve a maximum for greater numbers of components. In addition, the conditional probability for each locality's membership to a cluster are shown in Tables 3a and 3b and are generally low, suggesting that the choice of cluster is reasonable, with only ten out of the 119 places having a probability greater than 0.05.

The clustering exercise builds up the groups of regions that comprise the focus of our typology. To aid in interpreting the outcomes, plots of the 95% confidence interval on the mean of each variable are used to provide a visual indication of the significance of different indicators.

3.1 Income and wealth

The plots of the 95% confidence intervals for the means associated with the income and wealth variables are presented in Figure 1a to 1d. Cluster four is clearly differentiated from the other clusters by the average level of wages and salaries and together with cluster six is differentiated by the ratio of high income to low incomes. The remaining clusters either all have low aggregate levels of wage and salaries or lower ratios. In addition, none of the other clusters are clearly differentiated by these variables. The indicators of wealth (interest received and taxation imputation credits) show a different pattern of outcomes. Rather than being associated with high incomes they are more associated with regions that might be asset rich and income poor. Two groups, clusters three and five, are

Table 3a. Advantaged Regions: SLA name and uncertainty score

Income/workforce advantaged mining based regions		Income/labour market advantaged amenity based regions		Labour force advantaged, service based regions	
Banana (Qld)	0.00	Snowy River (NSW)	0.00	Albury (NSW)	0.00
Emerald (Qld)	0.00	Douglas (Qld)	0.00	Gladstone (Qld)	0.00
Mount Isa (Qld)	0.00	Broome (WA)	0.00	Mackay (Qld)	0.00
Kalgoorlie/Boulder (WA)	0.00	Alice Springs (NT)*	0.00	Wellington – Sale (Vic)	0.00
Port Hedland (WA)	0.00	Whitsunday (Qld)	0.00	Latrobe – Traralgon (Vic)	0.00
Roebourne (WA)	0.00	Wyndham-East Kimberley (WA)	0.00	West Tamar (Tas)	0.00
Singleton (NSW)	0.00	Katherine (NT)	0.00	Townsville (Qld)*	0.00
		Queanbeyan (NSW)	0.03	Thuringowa (Qld)*	0.00
				Bathurst (NSW)	0.00
				Wodonga (Vic)	0.00
				Greenough (WA)	0.00
				Orange (NSW)	0.00
				Bunbury (WA)	0.00
				Toowoomba (Qld)*	0.00
				Wagga Wagga (NSW)	0.00
				Calliope (Qld)	0.00
				Greater Bendigo (Vic)*	0.00
				Dubbo (NSW)	0.00
				Ballarat (Vic)*	0.00
				Maitland (NSW)	0.00
				Lake Macquarie (NSW)	0.00
				Warrnambool (Vic)	0.00
				Newcastle (NSW)*	0.00
				Tamworth (NSW)	0.00
				Muswellbrook (NSW)	0.00
				Cairns (Qld)*	0.00
				Port Lincoln (SA)	0.00
				Horsham – Central (Vic)	0.00
				Rockhampton (Qld)	0.02
				Goulburn (NSW)	0.02
				Greater Shepparton (Vic)	0.06
				Campaspe – Echuca (Vic)	0.07

* = combined SLAs

differentiated by the indicator of interest earned, while cluster three is also clearly differentiated according to the level of imputation credits received.

3.2 Labour force outcomes

The plots of the 95% confidence intervals for the means associated with the labour force outcome variables are presented in Figure 2a to 2d. All of the clusters differ to some degree across the four labour force variables. Cluster four is differentiated

Table 3b. Disadvantaged regions: SLA name and uncertainty score

Negative labour force/income disadvantaged regions		Agricultural based income disadvantaged/ labour market advantaged regions		Income poor/asset rich amenity based regions	
Cessnock (NSW)	0.00	Cowra (NSW)	0.00	Great Lakes (NSW)	0.00
Richmond Valley – Casino (NSW)	0.00	Young (NSW)	0.00	Tweed (NSW)	0.00
Grafton (NSW)	0.00	Tumut (NSW)	0.00	Ballina (NSW)	0.00
Broken Hill (NSW)	0.00	Griffith (NSW)	0.00	Byron (NSW)	0.00
Bundaberg (Qld)	0.00	Leeton (NSW)	0.00	Maclean (NSW)	0.00
Maryborough (Qld)	0.00	Bowen (Qld)	0.00	Bega Valley (NSW)	0.00
Latrobe – Moe (Vic)	0.00	Burdekin (Qld)	0.00	Eurobodalla (NSW)	0.00
Latrobe – Morwell (Vic)	0.00	Ararat (Vic)	0.00	Burnett (Qld)	0.00
Port Pirie (SA)	0.00	Campaspe – Kyabram (Vic)	0.00	Hervey Bay (Qld)	0.00
Burnie (Tas)	0.00	Moira – West (Vic)	0.00	Gold Coast (Qld)*	0.00
Central Coast (Tas)	0.00	Jondaryan (Qld)*	0.00	Maroochy (Qld)*	0.00
Devonport (Tas)	0.00	Swan Hill (Vic)*	0.00	Noosa (Qld)*	0.00
Waratah/Wynyard	0.00	Moree Plains (NSW)	0.00	Shoalhaven (NSW)*	0.00
Cooloola – Gympie (Qld)	0.00	Narrabri (NSW)	0.00	Hastings (NSW)*	0.00
Geraldton (WA)	0.00	Esperance (WA)	0.00	Coffs Harbour (NSW)*	0.00
Lismore (NSW)	0.00	Kingaroy (Qld)	0.00	East Gippsland – Bairnsdale (Vic)	0.00
Kempsey (NSW)	0.00	Gunnedah (NSW)	0.00	Port Stephens (NSW)	0.00
Whyalla (SA)	0.00	Mudgee (NSW)	0.00	Busselton (WA)	0.00
Greater Taree (NSW)	0.00	Atherton (Qld)	0.01	Albany – Central (WA)	0.00
Launceston (Tas)*	0.01	Baw Baw – Part B West (Vic)	0.02	Armidale Dumaresq (NSW)	0.02
Warwick – Central (Qld)	0.01	Inverell (NSW)*	0.03	Livingstone (Qld)	0.04
Murray Bridge (SA)	0.04	Mildura (Vic)	0.03		
Port Augusta (SA)	0.06	Albany- Bal (WA)	0.11		
Copper Coast (SA)	0.09	Parkes (NSW)	0.19		
Glenelg – Portland (Vic)	0.27	Wangaratta (Vic)	0.22		
		Mount Gambier (SA)	0.34		

* = combined SLAs

from the other clusters in terms of part time employment, with an above average mean. No other clusters stand out singularly, however, several of the clusters group together when we consider levels of unemployment and labour market participation. Most notable are clusters four, five and six which all have below average levels of total and youth unemployment and which together with cluster one (which has a below average adult unemployment rate) have above average levels of labour force participation.

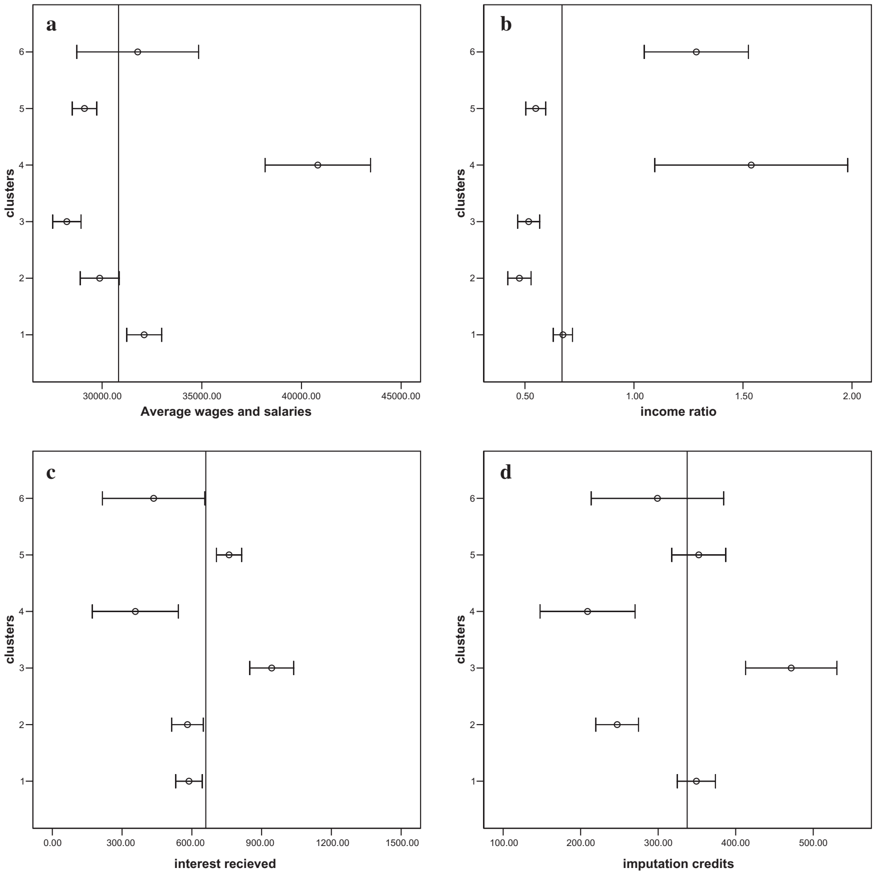

Fig. 1. a. Wage and salary, **b.** Ratio of high incomes to low incomes, **c.** Interest earned per tax payer, **d.** Imputation credits earned per taxpayer

3.3 Industry and occupation (including education)

The plots of the 95% confidence intervals for the means associated with the industry and occupation variables are presented in Figure 3a to 3j. Only a few differentiating variables are present when we consider industry variables. Most notable are cluster four which has an above average level of employment in the mining industry and cluster five which has an above average level of employment in agriculture. The only remaining differences are reflected in the graph for employment in mass recreation industries which are above average for cluster three and cluster six. There are no distinguishing variables in the two occupation variables or the variable measuring low levels of formal human capital.

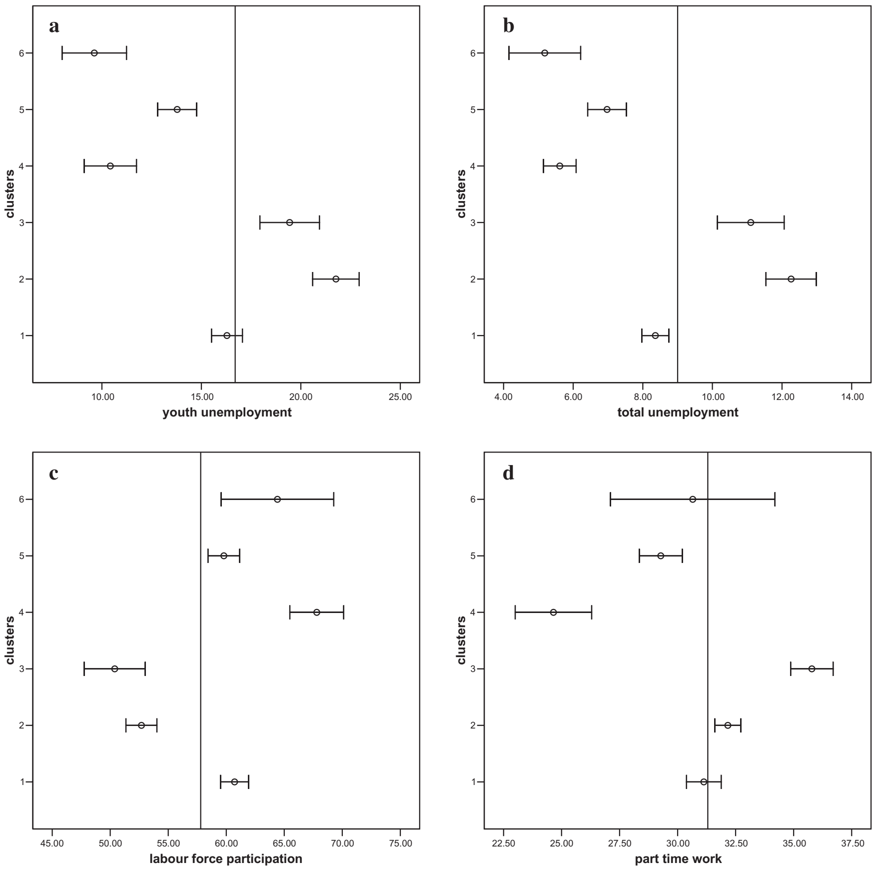

Fig. 2. a. Youth Unemployment, **b.** Total unemployment, **c.** Labour force participation, **d.** Part time workers

3.4 Household/family and housing

The plots of the 95% confidence intervals for the means associated with the household/family and housing variables are presented in Figure 4a to 4d. Cluster two is singularly differentiated from other groups in terms of a high proportion of families with no employed parent and above average levels of rental financial stress. The third cluster is singularly differentiated by high levels of households suffering rental financial stress and mortgage financial stress, while the fourth cluster is singularly differentiated by low levels of households suffering mortgage stress. Clusters four and six are together differentiated in terms of low levels of rental financial stress and clusters two, three, and five all have high proportions of home owners. In contrast, clusters one, four and six all have low levels of home ownership.

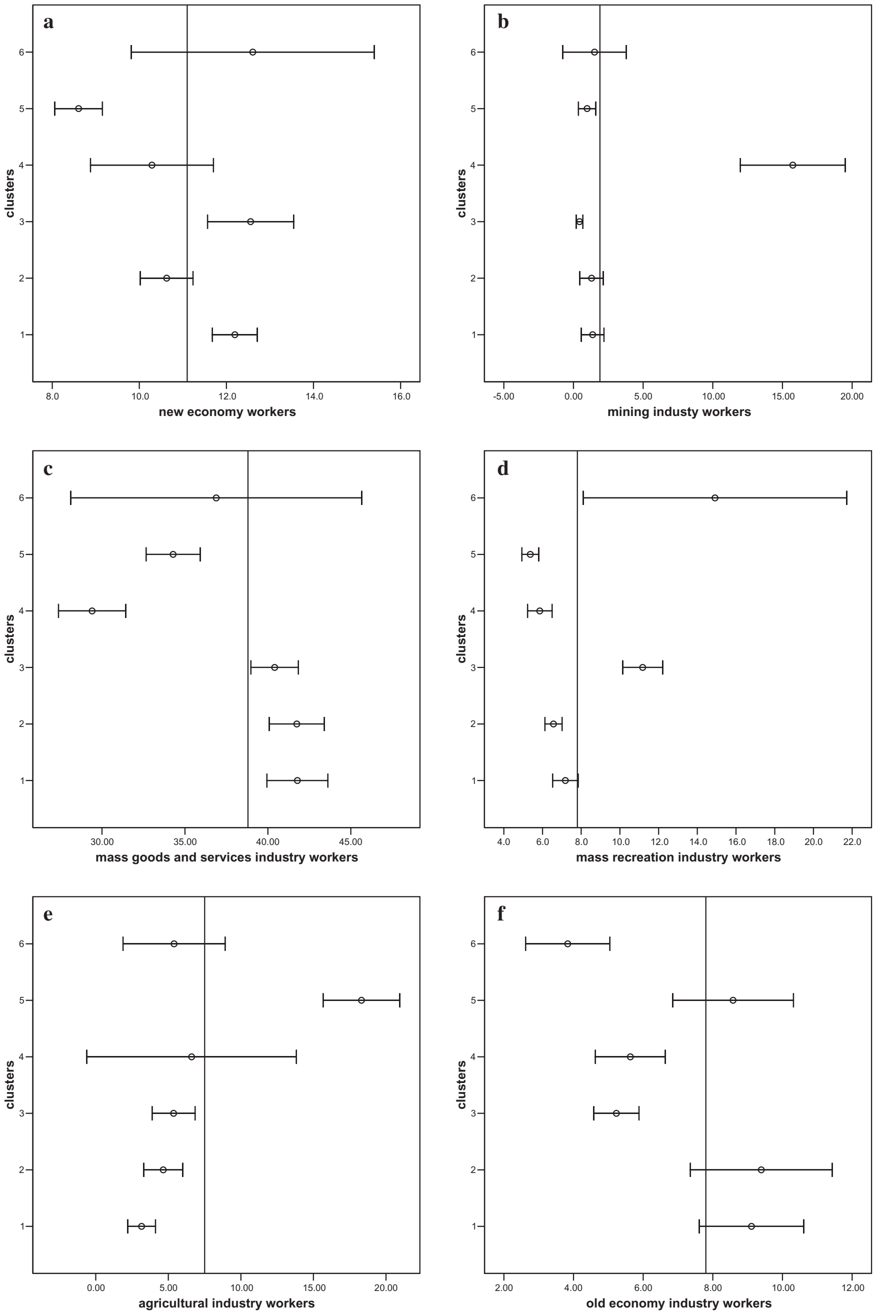

Fig. 3. a. Employment in new economy industries, **b.** Employment in mining indutry workers, **c.** Employment in mass goods and services industry, **d.** Employment in mass recreation industries, **e.** Employment in agriculture, **f.** Employment in old economy industries, **g.** Specialisation, **h.** Educated professionals, **i.** Vulnerable occupations, **j.** Low formal human captial

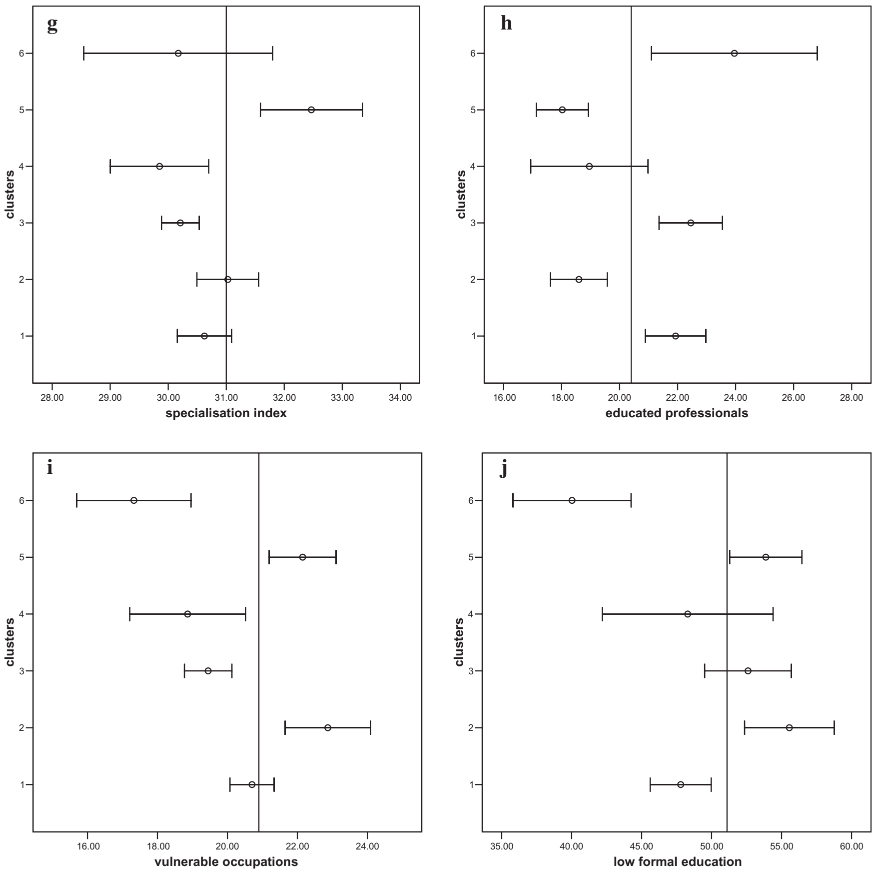

Fig. 3. Continued

3.5 Socio-economic change

The plots of the 95% confidence intervals for the means associated with the socio-economic change variables are presented in Figure 5a and 5b. None of the six clusters is differentiated on the measures of socio-economic change.

4 An overview of the non-metropolitan typology

The analysis of the plots of the confidence intervals allows us to develop our typology of non-metropolitan regions and localities. Readers interested in reviewing the clusters more closely can refer to the means presented in Table 4, with the list of individual SLAs in each cluster in Table 3a and 3b. Maps showing the location of each SLA by cluster across non-metropolitan Australia are presented in

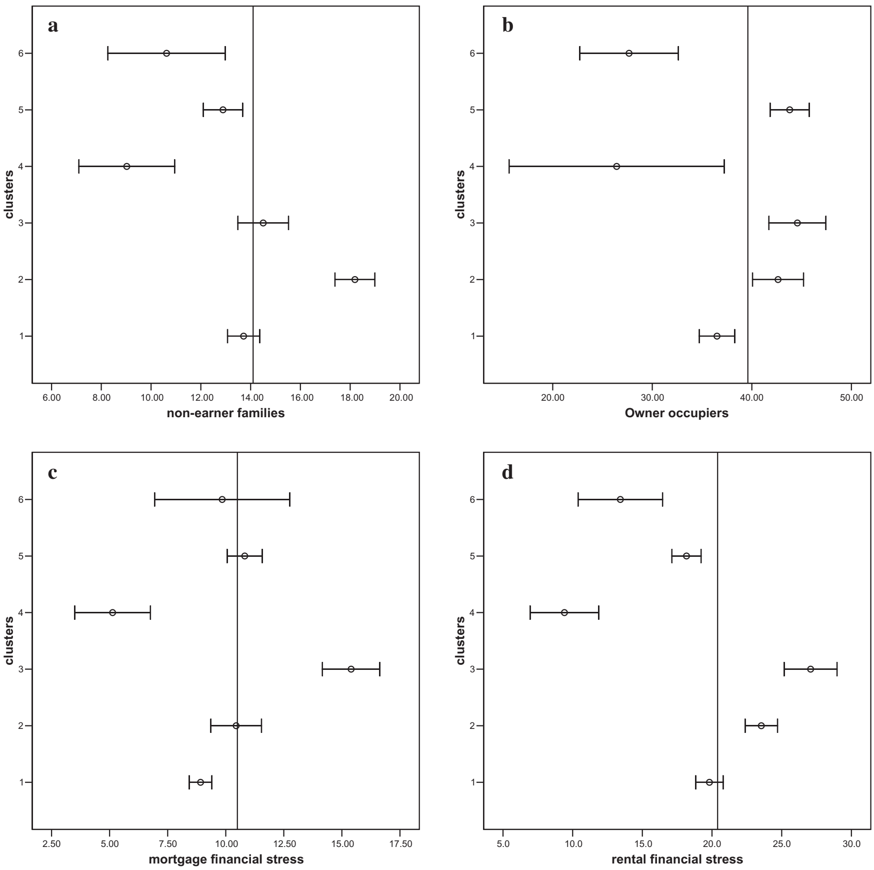

Fig. 4. a. Non-earner families, **b.** Owner occupiers, **c.** Mortage financial stress, **d.** Rental financial stress

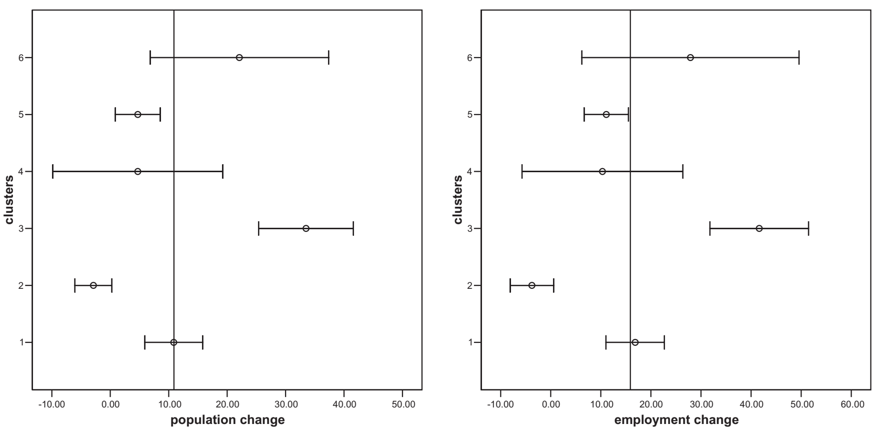

Fig. 5. a. Population change, **b.** Employment change

Table 4. Cluster means, selected variables

	Cluster 1: labour force advantaged, service based regions	Cluster 2: negative labour force/ income disadvantaged regions	Cluster 3: income poor/asset rich amenity based regions	Cluster 4: income/ workforce advantaged mining based regions	Cluster 5: agricultural based income disadvantaged/ labour market advantaged regions	Cluster 6: income/labour market advantaged amenity based regions	Total
Wage and salary ($)	30,833.72	28,805.48	27,349.80	38,975.43	27,796.88	30,386.62	29,578.17
Income ratio	0.67	0.47	0.51	1.5	0.54	1.28	0.67
Imputation credit per taxpayer ($)	346.73	247.11	469.64	208.87	253.36	299.00	337.40
Interest received per taxpayer ($)	577.43	581.20	942.14	357.10	760.02	435.78	659.99
Youth unemployment rate	16.12	21.8	19.5	10.4	13.8	9.6	16.7
Adult unemployment rate	8.3	12.3	11.05	5.6	7.0	5.2	9.0
Labour force participation rate	61.0	52.7	50.6	67.8	59.7	64.4	57.8
Part time work	31.0	32.2	35.7	24.6	29.3	30.7	31.3
New economy workers	12.2	10.62	12.5	10.3	8.6	12.6	11.1
Mining	1.4	1.3	0.4	15.7	0.9	1.5	1.9
Mass goods and services	41.4	41.7	41.0	29.4	34.3	36.9	38.8
Mass recreation	7.1	6.6	11.1	5.9	5.4	14.9	7.8
Agriculture	3.1	4.6	5.3	6.6	18.3	5.4	7.5
Old economy	9.3	9.4	5.01	5.6	8.6	3.8	7.8
Specialisation index	30.5	31.0	30.3	29.9	32.5	30.2	31.0
Educated professionals	21.6	18.6	22.8	18.9	18.0	23.9	20.4
Vulnerable occupations	20.8	22.9	19.4	18.9	22.1	17.3	20.9
Low education	48.2	55.6	51.8	48.3	53.9	40.0	51.1
No earner families	13.69	18.2	14.5	9.0	12.8	10.6	14.1
% owner occupied housing	36.5	42.6	44.1	26.4	43.8	27.7	39.6
Mortgage financial stress	8.9	10.4	15.1	5.1	10.8	9.8	10.5
Rent financial stress	19.6	23.53	27.0	9.4	18.2	13.4	20.4
Population change	11.4	–2.9	31.6	4.7	5.0	22.1	10.9
Employment change	17.6	–3.7	39.8	10.3	11.2	27.9	15.9

Fig. 6. Cluster membership, non-metropolitan New South Wales

Figures 6 to 12. Considering the outcomes from the analysis, the six clusters can be meaningfully divided into three groups of advantaged localities and three groups of disadvantaged places.

4.1 Advantaged localities

The first category of advantaged regions and localities was a set of income/workforce advantaged mining based regions (cluster 4), found in seven places in regional Western Australia, New South Wales and Queensland. They are characterised by employment in the mining industry and have high levels of income and commensurately low levels of households suffering from housing related stress. The cluster has good employment outcomes with low unemployment and high labour force participation. The cluster also has a low level of employees working on a part-time basis and has a low level of home ownership.

A second group of eight regions form a cluster of income/labour market advantaged amenity based regions (cluster 6). The cluster consists of eight localities found in New South Wales, Queensland, Western Australia and Northern Territory. They have significant employment in mass recreation industries and generally record more high-income individuals than low-income individuals. Like the mining based group, this cluster is characterised by good employment outcomes with low unemployment and above average labour force participation and below average home ownership.

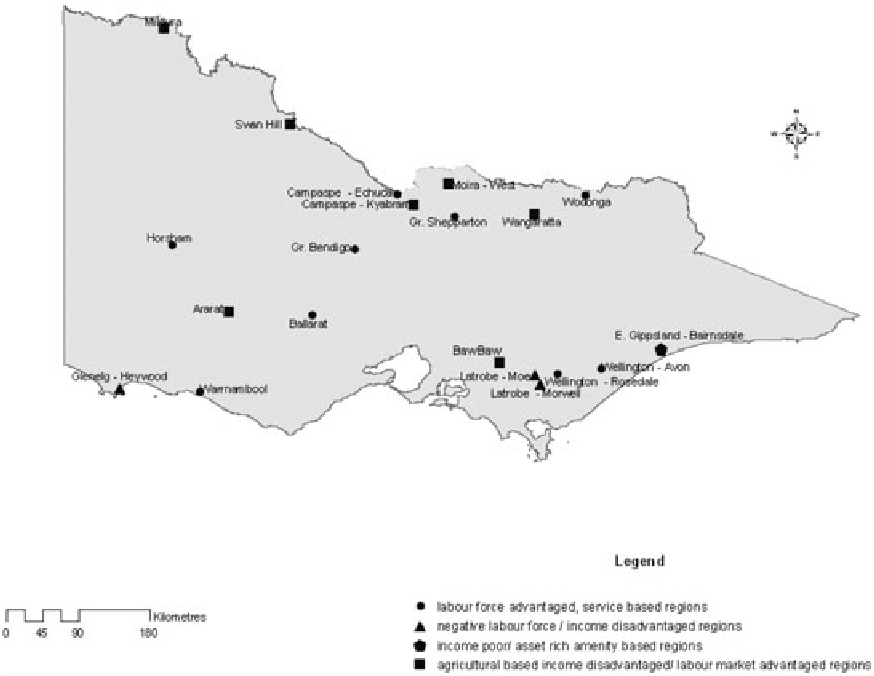

Fig. 7. Cluster membership, non-metropolitan Victoria

A third advantaged cluster comprises a large group of cities, towns and regions, many of which have important regional and rural service functions, which are defined as a labour force advantaged, service based regions (cluster 1). These 32 regions account for about one in three of the large regional cities and towns in Australia and are found in all states except Northern Territory. The regions in this cluster are differentiated by only three variables. This group's labour force advantage is reflected in a below average adult unemployment rate and an above average labour force participation rate. They also have a below average level of home ownership.

4.2 *Disadvantaged localities*

In contrast to the places labelled as advantaged cities, towns and regions were three clusters of localities characterised by varying degrees of relative disadvantage. A cluster of 24 regional cities and towns located in New South Wales, South Australia, Western Australia, Tasmania, Victoria and Queensland were identified as a negative labour force/income disadvantaged regions (cluster 2). Many of these localities developed during early periods of industrial growth in an era of protectionism and have since seen a reduction in their manufacturing fortunes. The cluster is distinguished in terms of poor labour market outcomes (above average

Fig. 8. Cluster membership, non-metropolitan Queensland

Fig. 9. Cluster membership, non-metropolitan Western Australia

unemployment and below average unemployment) and has above average levels of households suffering rental financial stress and families with no employed parent. This cluster is also disadvantaged in terms of income with a below average ratio of high-income earners to low income earners.

The second cluster of disadvantaged localities is an agricultural based income disadvantaged / labour market advantaged regions cluster (cluster 5) comprising 26 larger regional cities and towns found in all states except Tasmania and the Northern Territory. These places are mainly agricultural/pastoral-

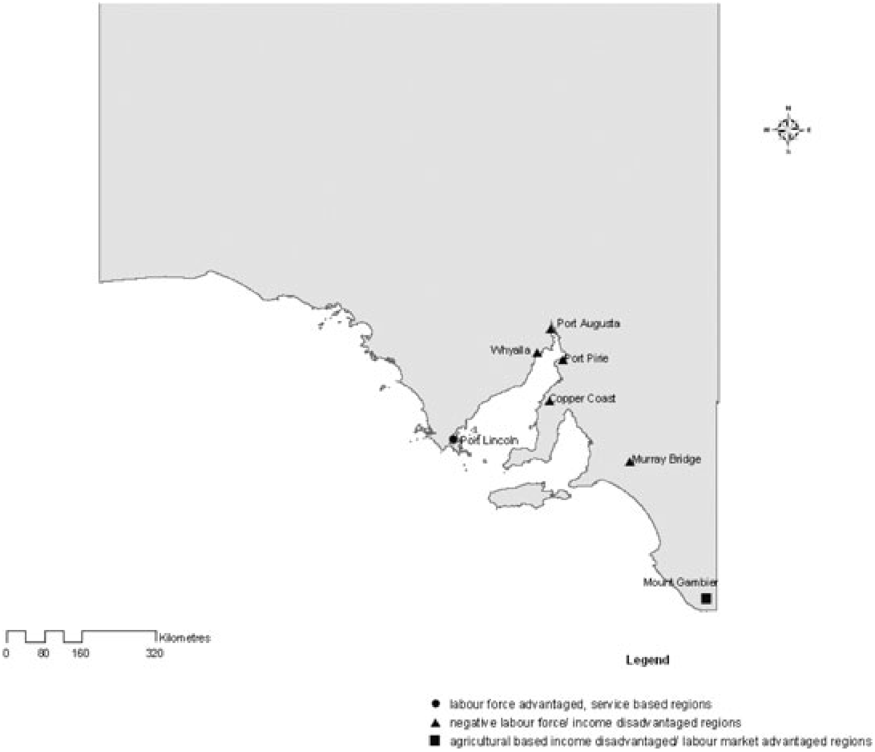

Fig. 10. Cluster membership, non-metropolitan South Australia

based towns that have become stagnant or are in decline. They are differentiated in terms of having high levels of employment in agriculture. The cluster has positive labour market outcomes with below average unemployment and above average labour force participation but has low incomes, with more low-income earners than high-income earners. Despite these low incomes however, the cluster does have an above average level of interest received suggesting that many places may be characterised as being income poor but asset rich.

A cluster of 21 regional cities and towns was identified as income poor / asset rich amenity based regions (cluster 3), located mainly along the coast of New South Wales and Queensland, but also in Western Australia. These types of places have been referred to in previous research as Australia's sunbelt migration regions or as welfare/retirement migration regions (Stimson et al. 2001). The cluster is differentiated in terms of low incomes and high levels of housing financial stress. In contrast to the low-income levels, the regions in the cluster do have high levels of assets with above average levels of interest received and imputation credits. The cluster has high levels of employment in mass recreation industries and an above average level of home ownership.

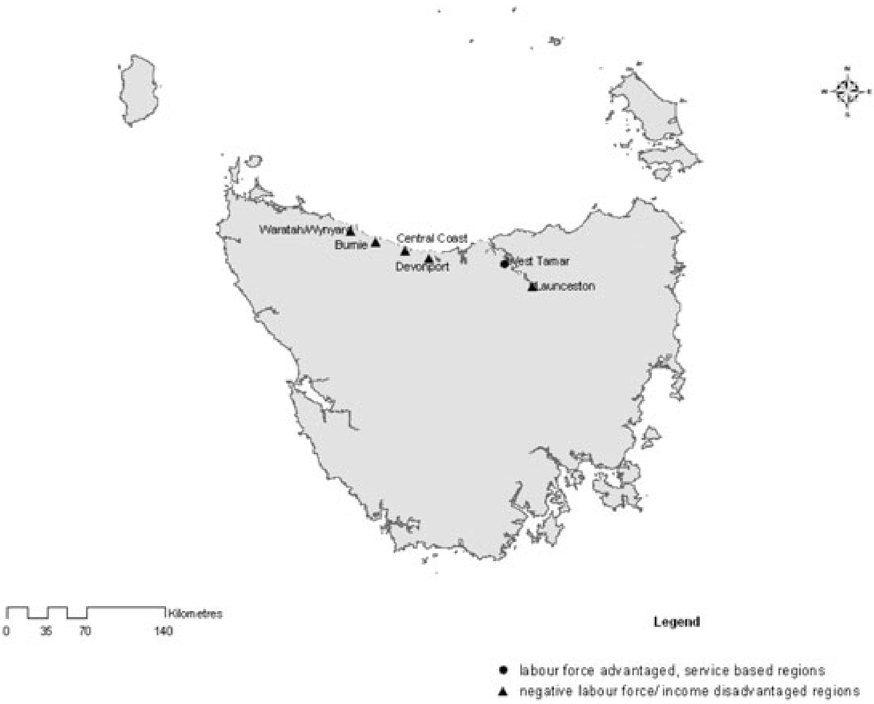

Fig. 11. Cluster membership, non-metropolitan Tasmania

5 Discussion

This paper has used a combination of model based clustering and visual analysis of confidence intervals to consider socio-economic outcomes in non-metropolitan Australia. Acknowledging that socio-economic outcomes in non-metropolitan Australia will be reflected in a range of non-metropolitan region types, we have followed the earlier research by Beer et al. (1994), Stimson et al. (2001, 2003) and Baum et al. (1999) and developed a typology of advantage and disadvantage based on a range of socio-economic indicators and adopting a multivariate classification technique. The research provided a useful way of categorising regions into groups based on their socio-economic performance.

Methodologically, the approach used in this paper represents an extension of earlier techniques used to cluster geographical areas. By adopting a model-based clustering approach, we have used a more rigorous statistical model selection technique to address the issues associated with determining an appropriate cluster method and the number of clusters present in the data. This approach has the advantage of identifying a probability mixture model that best describes the data of interest. The components of the selected mixture model correspond to clusters in the multivariate data and the model provides a measure of uncertainty about how well each observation has been classified. Typically, there is some level of uncer-

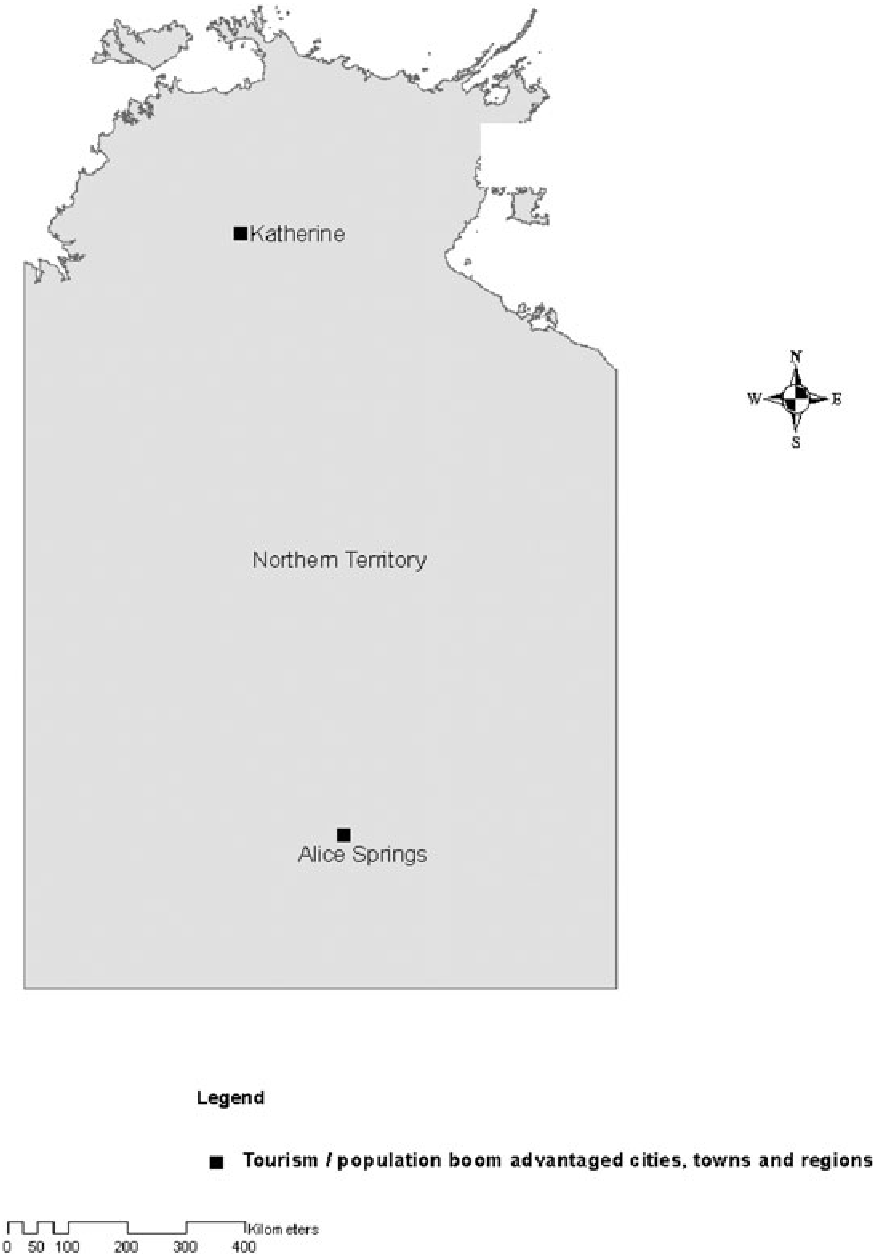

Fig. 12. Cluster membership, non-metropolitan Northern Territory

tainty associated with the classification of observations to a cluster. To allow for this uncertainty we generate 95 percent confidence intervals for the variable means within each cluster to identify variables that are strongly associated with the formation of a cluster.

Empirically, we have illustrated that socio-economic outcomes across Australia's large non-metropolitan cities, towns and regions are not even and reflecting earlier work illustrate the diverse nature of non-metropolitan regions. What this analysis of advantage and disadvantage across Australia's non-metropolitan regions demonstrates is a complex set of trends, in which:

(a) advantage is not evenly dispersed and is not always associated with the coastal sun-belt regions;
(b) growth can create disadvantage;
(c) disadvantage is spread across both coastal regions and inland agricultural service centres;
(d) coastal and inland centres that once prospered under protected manufacturing and utilities production have become vulnerable;
(e) opportunity in the new services sectors such as tourism and in mining has swept up a small number of often remote locations as localities that are thriving; and
(f) many inland cities and towns demonstrate both the continued strong performance and likely longer-term viability of some traditional regional service centres in both coastal and inland locations, where public funded functions in administration-education-health are important ingredients of that success.

Globalisation processes do seem to be creating advantaged regions in a relatively small number of places in non-metropolitan Australia, mainly through exporting of minerals, tourism and agricultural products processing. However, it is also evident that the processes of globalisation and the impacts of economic restructuring have created disadvantage across significant belts of inland Australia's agricultural communities and as well have adversely impacted its coastal and inland cities and towns that once prospered under protected manufacturing.

In conclusion, the analysis in this paper has given a descriptive overview of the socio-economic outcomes in Australia's large non-metropolitan cities, towns and regions. It is however important to note that the research reported in this paper is in some ways exploratory. Indeed, there could be considerable debate over the selection of variables used for the analysis and the inclusion of other variables might result in different outcomes. Additionally, it is likely that some regions might also be placed in more than one cluster and it could be that there may be subgroups contained within larger clusters that our methodology did not uncover. A finer empirical analysis of each cluster type would therefore pinpoint the presence of these sub-regions, and this could be done by perhaps using data at a finer level of aggregation (if these are available) or by utilising some other form of analysis such as ethnographic research. These issues aside, the typology developed does provide a useful background and provides a necessary basis for any detailed study of a particular region or group of regions because it provides the context within which more in depth study can be comprehended.

References

Banfield JD, Raftery AE (1993) Model-based Gaussian and non-Gaussian clustering. *Biometrics* 49: 803–821

Banks G (2000) Meeting the challenge of change in regional Australia. Address to Renaissance of the Regions Symposium, Melbourne, Victoria 9th November, http://www.pc.gov.au/speeches/cs20001109/cs20001109.pdf, date accessed 11th July 2005

Baum S, Stimson R, O'Connor K, Mullins P, Davis R (1999) *Community opportunity and vulnerability in Australia's cities and towns: Characteristics, patters and implications*. University of Queensland Press, for the Australian Housing and Urban Research Institute, Brisbane.

Baum S, Mullins P, Stimson R, O'Connor K (2002) Communities of the post-industrial city. *Urban Affairs Review* 37: 322–357

Baum S, O'Connor K, Stimson R (2005) *Fault lines exposed: Advantage and disadvantage across Australia's settlement system*. Monash University e-press, Melbourne

Beer A (1999) *Regional cities with Australia's evolving urban system, 1992–1996*. Paper presented at the 16th Pacific Regional Science Conference, Seoul

Beer A, Bolam A, Maude A (1994) *Beyond the capitals: Urban growth in regional Australia*. Australian Government Publishing Service, Canberra

Beer A, Maude A (1995) Regional cities in the Australian urban system, 1961 to 1991. *Urban Policy and Research* 13(3): 135–148

Beer A, Maude A, Pritchard B (2003) *Developing Australia's regions: Theory and practice*. University of New South Wales Press, Sydney

Berry B (1996) Technology-sensitive urban typology. *Urban Geography* 17: 674–689

Burnley I, Murphy P (2004) *Sea change: Movement from metropolitan to Arcadian Australia*. University of New South Wales Press, Sydney

Coulton C, Chow J, Wang E, Su M (1996) Geographical concentration of affluence and poverty in 100 metropolitan areas, 1990. *Urban Affairs Review* 32: 186–216

Fraley C, Raftery AE (1999) MCLUST: Software for model-based cluster analysis. *Journal of Classification* 16: 297–206

Fraley C, Raftery AE (2002a) Model based clustering, discriminant analysis and density estimation. *Journal of the American Statistical Association* 97: 611–631

Fraley C, Raftery AE (2002b) MCLUST: Software for model-based clustering, discriminant analysis, and density estimation, Technical Report No. 415, Department of Statistics, University of Washington

Fraley C, Raftery AE (2003) Enhanced software for model-based clustering, discriminant analysis, and density estimation: MCLUST. *Journal of Classification* 20: 263–286

Hill E, Brennan J, Wolman H (1998) What is a central city in the United States: Applying a statistical technique for developing taxonomies. *Urban Studies* 35: 1935–1969

Lawrence G, Lyons K, Momtaz S (eds) (1996) *Social change in rural Australia*. Central Queensland University, Rockhampton

Lawrence G, Gray I, Stehlik D (1999) Changing spaces: the effects of macro-social forces on regional Australia. In: Kasimis C, Papadopoulos A (eds) *Local responses to global integration*. Ashgate, Aldershot

Macadam R, Drinan J, Inall N, McKenzie B (2004) Growing the capital of rural Australia: The task of capacity building. A report for the Rural Industries Research and Development Corporation. RIRDC Publication no. 04/034

Marcuse P (1997) The enclave, the citadel and the ghetto: What has changed in the post-fordist U.S. city? *Urban Affairs Review* 33: 228–264

Massey D, Eggers M (1993) The spatial concentration of affluence and poverty during the 1970s. *Urban Affairs Quarterly* 29: 299–315
Masson M, Loftus G (2003) Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology* 57: 203–220
Mikelbank B (2004) A typology of U.S. suburban places. *Housing Policy Debate* 15:
O'Connor K, Healy E (2001) *The links between housing markets and labour markets in Melbourne*. Work in progress report prepared for the Australian Housing and Urban Research Institute. Australian Housing and Urban Research Institute, Melbourne
R Development Core Team (2003) *R: A language and environment for statistical computing*. R Foundation for statistical computing, Vienna, Austria
Schwarz G (1978) Estimating the dimension of a model. *The Annals of Statistics* 6: 461–464
Scott K, Park J, Cocklin C (2001) From sustainable rural community to social sustainability. *Journal of Rural Studies* 16: 433–446
Stimson R, Baum S, Mullins P, O'Connor K (2001) Australia's regional cities and towns: Modelling community opportunity and vulnerability. *Australasian Journal of Regional Studies* 7: 23–62
Stimson R, Baum S, O'Connor K (2003) The social and economic performance of Australia's large regional cities and towns: Implications for rural and regional policy. *Australian Geographical Studies* 41: 131–147
Tonts M (1996) Economic restructuring and small town adjustment: Evidence from the Western Australian Central Wheatbelt. *Rural Society* 6: 24–33