

The South African Index of Multiple Deprivation 2007 at Datazone Level (modelled)

Michael Noble, Christopher Dibben and Gemma Wright

March 2010

Building a Caring Society. Together.



social development

Department:
Social Development
REPUBLIC OF SOUTH AFRICA

The South African Index of Multiple Deprivation 2007 at Datazone Level (modelled)

Michael Noble, Christopher Dibben and Gemma Wright

March 2010

Suggested citation:

Noble, M., Dibben, C. and Wright, G. (2010) *The South African Index of Multiple Deprivation 2007 at Datazone Level (modelled)*, Pretoria: Department of Social Development.

Disclaimer:

The facts presented and views expressed in this report are those of the authors and not necessarily those of the Department of Social Development or the UK Department for International Development. The University of Oxford has taken reasonable care to ensure that the information in this report and the accompanying data are correct. However, no warranty, express or implied, is given as to its accuracy and the University of Oxford does not accept any liability for error or omission. The University of Oxford is not responsible for how the information is used, how it is interpreted or what reliance is placed on it. The University of Oxford does not guarantee that the information in this report or in the accompanying file is fit for any particular purpose. The University of Oxford does not accept responsibility for any alteration or manipulation of the report or the data once it has been released.

The authors:

Prof Michael Noble is Professor of Social Policy at the Department of Social Policy and Social Work at the University of Oxford and Director of the Centre for the Analysis of South African Social Policy (CASASP) at the University of Oxford

Dr Christopher Dibben is a Senior Lecturer at the University of St Andrews, Scotland and Associate of the Centre for the Analysis of South African Social Policy (CASASP) at the University of Oxford

Dr Gemma Wright is a Senior Research Fellow at the Department of Social Policy and Social Work at the University of Oxford and Deputy Director of CASASP

Acknowledgements

The team would like to thank the Department of Social Development for sponsoring this project; the UK Department for International Development Southern Africa for funding it; and Statistics South Africa for their assistance in producing the data.

Contents

1	Background.....	5
2	Methodology	7
	Conceptualising multiple deprivation.....	7
	The data	7
	Domains of deprivation in the datazone-level SAIMD 2007	7
	Modelling the data	7
	Constructing the datazone-level SAIMD 2007	8
3	Validation: Comparison with municipality-level SAIMD 2007	10
4	Findings	12
5	Comparison with the datazone-level SAIMD 2001	25
	References.....	27
	Appendix.....	29

1 Background

This report provides details about a South African Index of Multiple Deprivation (SAIMD) that has been produced at small area level using modelled data relating to 2007. This datazone-level SAIMD 2007 is the most recent development in a series of Indices of Deprivation that have been produced at a small area level in South Africa.

The datazone-level SAIMD 2007 builds on several recent studies about deprivation in South Africa that were undertaken by a team from the University of Oxford's Centre for the Analysis of South African Social Policy (CASASP). Initially, a team comprising members of CASASP, the Human Sciences Research Council (HSRC) and Statistics South Africa (StatsSA) developed a ward-level measure of multiple deprivation for each of the nine provinces in the country, called the Provincial Indices of Multiple Deprivation (PIMD) (Noble *et al.*, 2006, 2009b). The PIMD refers to deprivation experienced by the total population (i.e. all ages including children) and was based on the 2001 Census.

Subsequently, a South African Index of Multiple Deprivation was produced at datazone level for 2001 for the whole of South Africa (Noble *et al.*, 2009a), again using data from the 2001 Census. Datazones are small area level statistical geographical units and enabled a much more fine-grained analysis of deprivation to be developed for 2001 (Avenell *et al.*, 2009). Most recently, a municipality-level SAIMD was produced for 2007 using the 2007 Community Survey (Wright and Noble, 2009).

However, as a Census has not been undertaken since 2001, and as the Community Survey 2007 is only robust down to municipality level, it has not so far been possible to update the datazone-level SAIMD 2001 using the techniques that had originally been used (Noble *et al.*, 2009a). In an attempt to produce a more fine-grained profile of deprivation in South African since 2001 than was achieved with the municipality-level SAIMD 2007, the research team considered using administrative data and using modelled survey data; this report is about the latter approach and involves modelling the 2007 Community Survey down to datazone level. The datazone-level SAIMD 2007 presented in this report is therefore an attempt to draw from the best features of its two most recent predecessors: like the SAIMD 2001 it is produced at datazone level and, like the municipality-level SAIMD 2007, it was developed using the most up to date data available in the form of the 2007 Community Survey (Statistics South Africa, 2007). The aim of this part of the project was to ascertain whether it was *possible* to produce plausible results using modelled data at datazone level.

In parallel to the work outlined above, a South African Index of Multiple Deprivation for Children 2001 (SAIMDC 2001) was produced at municipality level using data from the publicly available ten percent sample of the 2001 Census (Barnes *et al.*, 2007; Barnes *et al.*, 2009). This was further developed at datazone level for 2001 (Wright *et al.*, 2009a), and at municipality level for 2007 (Wright *et al.*, 2009b). However, a datazone-level version of the SAIMDC using modelled Community Survey data was not pursued due to concerns about modelling estimates of deprivation for such a small subset of the population (i.e. 0-17s).

Section 2 of this report presents the domains and indicators for the datazone-level SAIMD 2007 and summarises the methodological approach that was used. **Section 3** briefly compares the datazone-level SAIMD 2007 with the municipality-level SAIMD 2007, **Section**

4 presents the datazone-level SAIMD 2007 in more detail, and **Section 5** compares the datazone-level SAIMD 2007 with its 2001 counterpart.

2 Methodology

Conceptualising multiple deprivation

As with the municipality-level SAIMD 2007, the datazone-level SAIMD 2007 was conceptualised as a weighted combination of dimensions or ‘domains’ of deprivation (Wright and Noble, 2009).

The data

The source data for the datazone-level SAIMD 2007 was the 2007 Community Survey (CS). The CS was conducted in February 2007 by Statistics South Africa and covered 274,348 dwelling units across all of the provinces, and attained a response rate of 93.9% (Statistics South Africa, 2007: 10-11). It was a nationally representative large-scale household survey intended to provide information about the profile of the South African population between the 2001 and 2011 Censuses.

Domains of deprivation in the datazone-level SAIMD 2007

The domains within the datazone-level SAIMD 2007 are identical to those used in the municipality-level SAIMD 2007 (Wright and Noble, 2009)¹. Four domains of deprivation were produced:

- Income and Material Deprivation
- Employment Deprivation
- Education Deprivation
- Living Environment Deprivation

Modelling the data

There are many possible ways to model survey data to a small area level. These include synthetic estimation techniques, multiple imputation techniques and spatial microsimulation. A full review of approaches is outside the scope of this report but see for example Ballas *et al.* (2006) for a comprehensive review relating to the production of small area level income estimates, and Alderman *et al.* (2003) for a recent South African example of producing small area income estimates using a synthetic estimation technique.

The technique used here is a combination of direct estimation using data from the Community Survey at datazone level, and synthetic estimation using predictors from the 2001 Census. The technique uses multilevel modelling and is a modification of a technique employed in the development of the UK Health Poverty Index which was developed for the

¹ These dimensions of deprivation in the municipality-level SAIMD 2007 are the same as the datazone-level SAIMD 2001 except that there is no Health Domain (see Wright and Noble, 2009 for details about why it could not be derived using the CS data). Moreover, as a result of changes in wording of questions between the Census and the CS there are some very minor differences in the indicators comprising the domains in the municipality level SAIMD 2007 which are reflected in the modelled datazone-level SAIMD 2007.

UK Department of Health (Dibben *et al.*, 2001) This method is simply one way of calculating the commonly used ‘small area composite estimator’ (Rao 2003).

The method includes a direct estimation component. This is made possible by the fact that the Community Survey is a very large dataset and was supplied for this project by Statistics South Africa with a datazone code appended to the individual-level anonymised data. Using the same definitions of deprivation as for the municipality-level SAIMD 2007, deprived individuals were identified for each of the four domains (see Appendix 1 of Wright and Noble (2009) for a full definition of the variables used from the CS). A datazone-level score was created for each domain. For datazones where there are very few sampled cases or none, the estimate relies more on the synthetic estimate. The method therefore balances the advantages of using the synthetic estimate that may be biased for a particular datazone but will be well estimated against a direct estimate that will be unbiased but may have a large sampling error. It therefore produces a ‘best linear unbiased estimator’.

The method involves two main steps

Step 1 Using the Community Survey, with aggregated 2001 census variables linked to it, a multi-level, variable intercepts, logistic model was run, with level one being the individual i , level two datazone j , level three the municipality k and level four being the province h . Covariates, X , from the linked 2001 census were used to predict the probability P that individual i within a datazone jkh would be subject to the particular type of deprivation of interest as measured in the Community Survey. In each case the covariate X from the 2001 census was the equivalent measure from the Community Survey (ie with the same definition).

$$\text{Logit } (P_{ijkh}) = X_{ijkh} \mathbf{B} + \mathbf{U}_h + \mathbf{C}_{kh} + \mathbf{W}_{jkh} + \mathbf{E}_{ijkh}$$

Where P is a vector of probabilities associated with individual i in datazone j within municipality k and province h , B a vector of regression coefficients, X a matrix of covariates associated with the datazone and U , C , W are vectors of area effects associated with respectively province, municipality and datazone and E is a vector of independent random ‘noise’ elements (assumed to be binomially distributed).

Step 2 The fixed effects part of the model were then taken and applied to the matrix of small area covariates X from the 2001 census, the estimated area effects added, and the anti-logit applied. The probability was then summed and averaged over the datazone to produce a vector of synthetic datazone level estimates \hat{Y}_{jkh} :

$$\hat{Y}_{jkh} = \text{anti-Logit } (X_{ijkh} \hat{B} + \hat{U}_h + \hat{C}_{kh} + \hat{W}_{jkh})$$

Constructing the datazone-level SAIMD 2007

Once each of the four domain scores had been modelled, they were standardised by ranking, and transformed to an exponential distribution. Each transformed domain has a range of 0 to 100, with a score of 100 for the most deprived datazone. As before, equal

weights were assigned to the exponentially transformed domains in the absence of evidence suggesting differential weights should be used.

The SAIMD 2007 score is therefore the (equally) weighted sum of the exponentially transformed domain rank of the domain scores. The larger the SAIMD score, the more deprived the datazone.

It should be stressed that the data on which the datazone-level SAIMD 2007 is based are modelled estimates and should be regarded as experimental statistics.

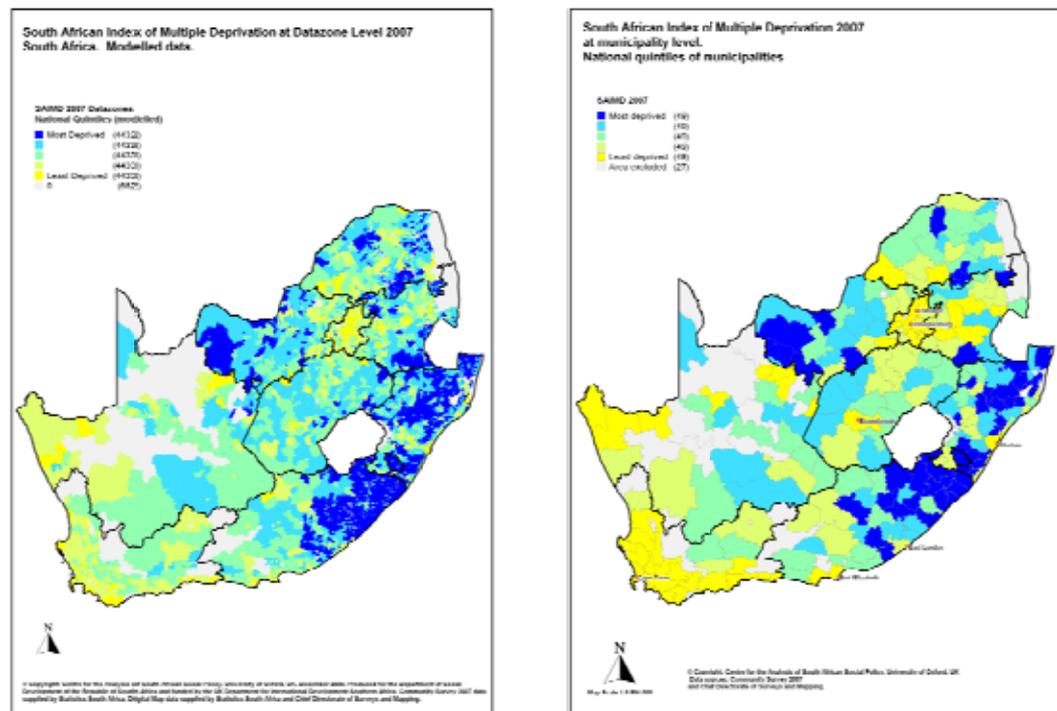
3 Validation: Comparison with municipality-level SAIMD 2007

One of the ways in which the modelled datazone SAIMD 2007 can be validated is to examine the extent to which the map of deprivation at datazone level accords with the already published municipality-level SAIMD 2007.

The map below shows the modelled datazone-level SAIMD 2007 (which is shown in more detailed form in Section 4 below) alongside the municipality-level SAIMD 2007 (taken from the municipality-level SAIMD 2007 report (Wright and Noble 2009)). As there are only 237² municipalities in the country, the municipalities have been quintised (divided into five equal groups) and mapped. In order to aid comparison, the datazone-level SAIMD 2007 has also been mapped here using quintiles of deprivation. The most deprived quintile in both maps is shown in a strong blue colour, and the least deprived quintile is bright yellow.

As one would anticipate, the overall picture of deprivation in both maps is the same, with most deprivation located in the former homeland areas, particularly the former Transkei, Ciskei, KwaZulu, Bophuthatswana and Lebowa. As with the datazone-level SAIMD 2001, pockets of deprivation are visible within otherwise less deprived municipalities (see for example the highly deprived datazones in the Free State, particularly around Thaba Nchu and Phuthaditjhaba).

Comparison of the Index of Multiple Deprivation 2007 at Datazone level (modelled) with Index of Multiple Deprivation 2007 at Municipality Level (on 2001 Municipality boundaries)



² 2007 Municipalities excluding District Management Areas and special Municipalities with very small populations.

Though these modelled estimates are only experimental statistics, this map demonstrates that the approach produces results in-line with what one might expect, given the municipality-level distribution of deprivation in 2007 that was produced using the Community Survey.

4 Findings

A profile of national, provincial and municipality-level deprivation is presented in the municipality-level SAIMD 2007 report for 2007 (Wright and Noble, 2009). As a reminder of the national levels of deprivation for each of the domains:

- 72% of people live in households that are income and/or materially deprived;
- 37.8% of the relevant working age population are unemployed or unable to work due to sickness/disability;
- 27.4% of the adult population aged 18-65 are education deprived (have no secondary schooling);
- 67.2% of the population experience living environment deprivation.

The rest of this section presents a profile of deprivation in 2007 at datazone level, based on the modelled CS data.

Table 1 Most Deprived Datazones by Province - Modelled SAIMD 2007

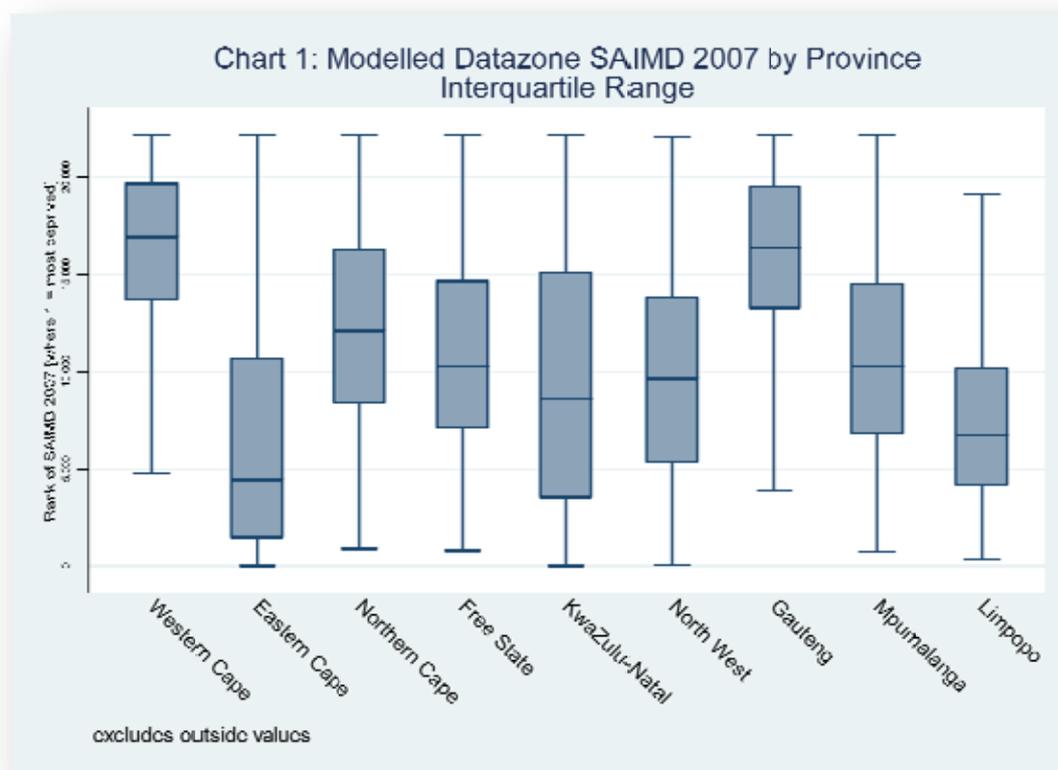
	No. of Datazones in Most Deprived Decile	No. of Datazones in Most Deprived Quintile	Total no. of Datazones in Province	% of Province's Datazones in Most Deprived Decile	% of Province's Datazones Most Deprived Quintile
Western Cape	0	3	2,184	0.0	0.1
Eastern Cape	1,061	1,591	3,181	33.4	50.0
Northern Cape	2	19	417	0.5	4.6
Free State	11	98	1,373	0.8	7.1
KwaZulu-Natal	731	1,394	4,663	15.7	29.9
North West	181	376	1,827	9.9	20.6
Gauteng	2	42	4,280	0.0	1.0
Mpumalanga	36	164	1,527	2.4	10.7
Limpopo	192	745	2,712	7.1	27.5

The 10% most deprived datazones in the country (the most deprived decile) and the 20% most deprived deciles in the country (the most deprived quintile) have been identified and Table 1 shows their distribution between the 9 provinces. Both the number of deprived datazones in the most deprived decile and quintile nationally per province as well as the percentage of such datazones in each province are shown. So, for example, in the Eastern Cape 1061 datazones of the province's 3181 datazones are in the 10% most deprived datazones nationally. This amounts to 33.4% of the province's datazones. If we look at

datazones in the Eastern Cape that are among the most deprived 20% of datazones in the country, 1591 datazones or 50% of the province's datazones are so placed.

Focussing on the most deprived decile, the Eastern Cape (with 33.4%) has the highest proportion of its datazones in the most deprived decile nationally. This is followed by KwaZulu-Natal with 15.7%, North West Province with 9.9% and Limpopo Province with 7.1%. The Western Cape and Gauteng have the lowest percentages of their datazones in the most deprived decile nationally. If we look at the most deprived quintile a similar picture emerges with the Eastern Cape, KwaZulu-Natal and Limpopo having the highest proportions in the most deprived quintile and the Western Cape and Gauteng having the lowest.

Another way of looking at the distribution of datazones between provinces is to look at the interquartile range of the deprivation rank of each datazone by province. The following Chart (Chart 1) illustrates this by way of box plots for each province.



The interquartile range – represented by the boxes in Chart 2 – indicates the range of ranks in which the middle 50% of datazones in a province fall. A short box, as seen for example in the Western Cape or Limpopo indicates that the middle 50% of datazones have ranks quite close together and that deprivation (or lack of it) is quite similar for half the datazones. A longer box as seen for example in the Eastern Cape or KwaZulu-Natal indicates a wider range of deprivation for the middle 50% of datazones. If the box sits towards the bottom of the chart – as for the Eastern Cape, KwaZulu-Natal and Limpopo – this indicates that the middle group of datazones are more deprived as compared to provinces such as the Western Cape and Gauteng where the box is up towards the top of the chart or the less deprived part of the distribution. The lines and ‘whiskers’ indicate the entire range of deprivation within that province. The Eastern Cape and KwaZulu-Natal have deprivation

across the entire range whereas the Western Cape has no datazones in the most deprived part of the distribution and Limpopo, on the other hand, has no datazones in the least deprived part of the distribution.

If we now look at the results geographically the provincial level picture presented by the box plots can be unpacked and displayed at very small area level.

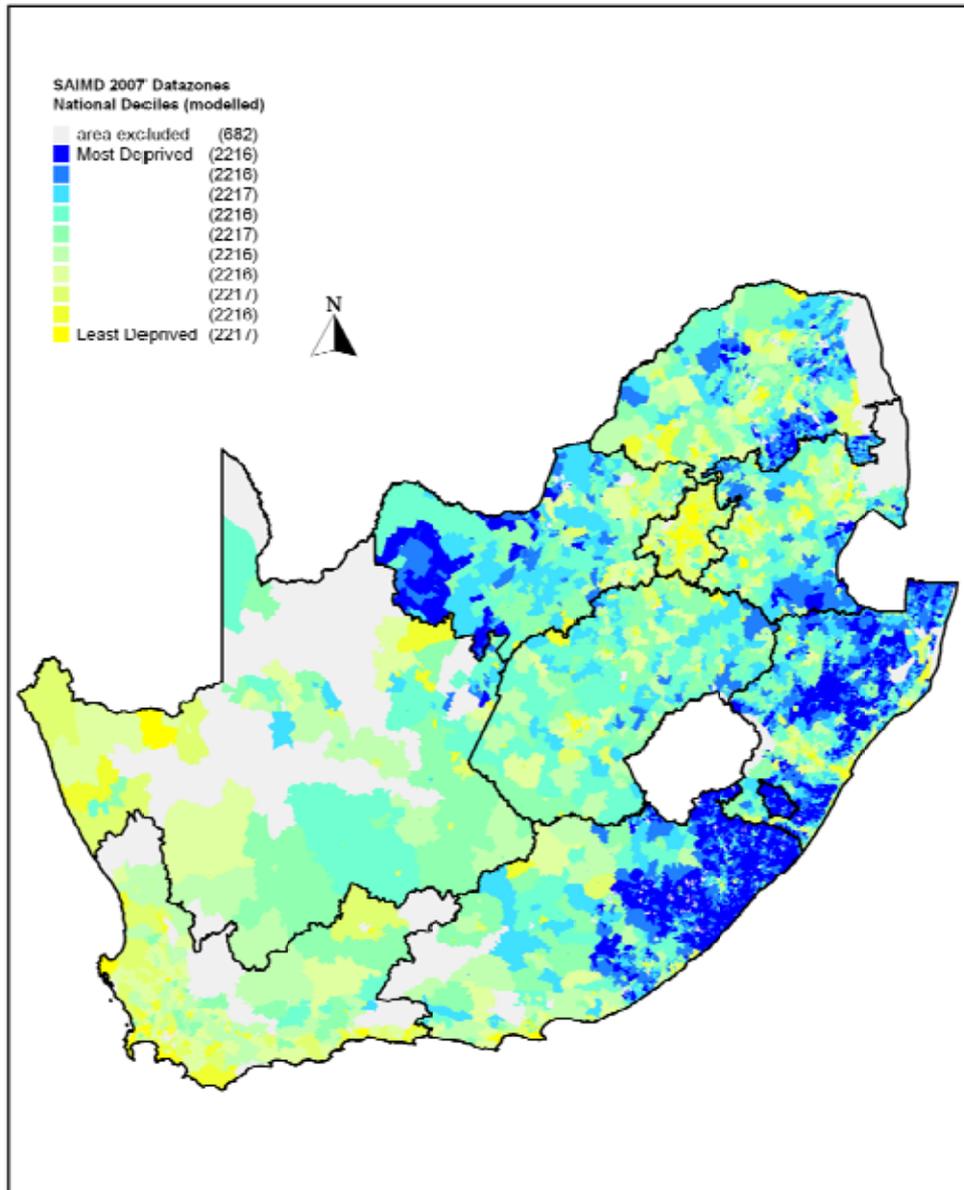
Map 1 shows the datazone-level SAIMD 2007 for the whole of South Africa. Maps 2-10 comprise zoom-ins of the same information for each province, again using national deciles of deprivation.

In Maps 1-10 all the datazones in the country have been divided into 10 equal groups (deciles). The most deprived decile is deep blue, while the least deprived decile is bright yellow. The map legends indicate the colour gradation between the two extremes.

As is the case with the datazone-level SAIMD 2001 (Noble *et al.*, 2009a), the striking finding emerging from these maps is the extent to which the most severe deprivation remains concentrated in the former homelands. This is apparent in Map 1 and can be seen particularly graphically in Map 3 (Eastern Cape) where deprivation is concentrated in the former Ciskei and Transkei homelands, in Map 6 where severe deprivation is apparent in the former KwaZulu homeland, Map 7 (North West) revealing deprivation in the former Bophuthatswana homeland, and Map 9 (Mpumalanga) where deprivation in the former KwaNdebele and KaNgwane can be seen. In Limpopo (Map 10) deprivation is particularly apparent in the former Lebowa and Gazankulu homelands.

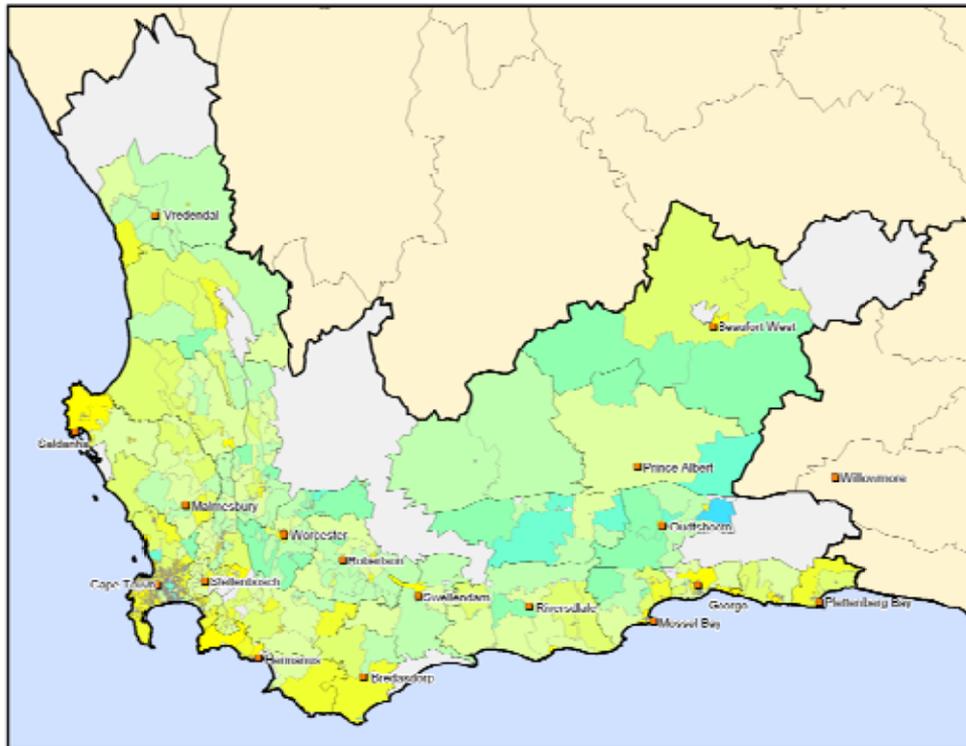
However, the strength of the datazone geography is that pockets of deprivation can be revealed in otherwise less deprived areas. Again as with the datazone SAIMD 2001, the pockets of deprivation can be seen in both urban and rural areas. So, for example in Map 3 (Eastern Cape) in the maps of Port Elizabeth and East London the deprivation in the townships is apparent.

South African Index of Multiple Deprivation at Datazone Level 2007
Map 1 - South Africa. Modelled data.



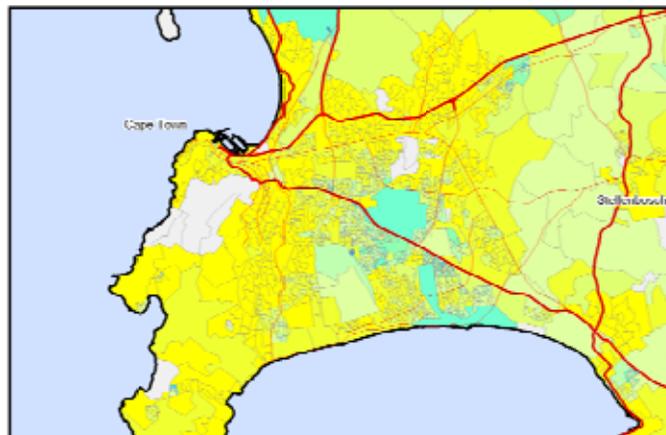
© Copyright: Centre for the Analysis of South African Social Policy, University of Oxford, UK, December 2008. Produced for the Department of Social Development of the Republic of South Africa and funded by the UK Department for International Development Southern Africa. Community Survey 2007 data supplied by Statistics South Africa. Digital Map data supplied by Statistics South Africa and Chief Directorate of Surveys and Mapping.

**South African Index of Multiple Deprivation at Datzone Level 2007
Map 2 - Western Cape. Modelled data.**



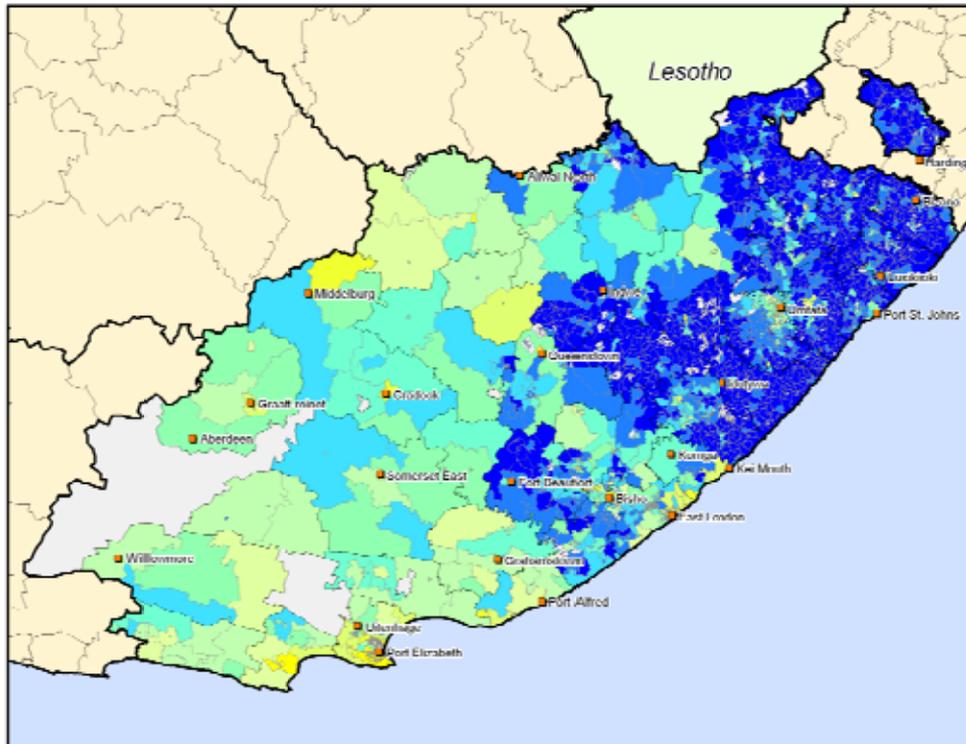
**SAIMD 2007 Datzones
National Deciles (modelled)**

area excluded	(60)
Most Deprived	(0)
	(3)
	(22)
	(44)
	(148)
	(279)
	(339)
	(429)
	(418)
Least Deprived	(502)



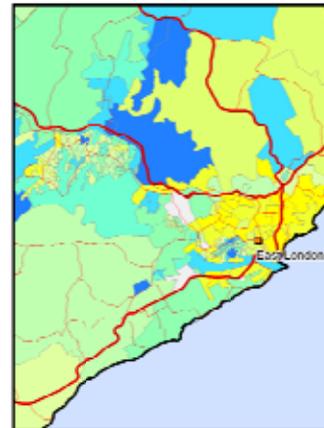
© Copyright: Centre for the Analysis of South African Social Policy, University of Oxford, UK, December 2009. Produced for the Department of Social Development of the Republic of South Africa and funded by the UK Department for International Development Southern Africa. Community Survey 2007 data supplied by Statistics South Africa. Digital Map data supplied by Statistics South Africa and Chief Directorate of Surveys and Mapping.

**South African Index of Multiple Deprivation at Datazone Level 2007
Map 3 - Eastern Cape. Modelled data.**



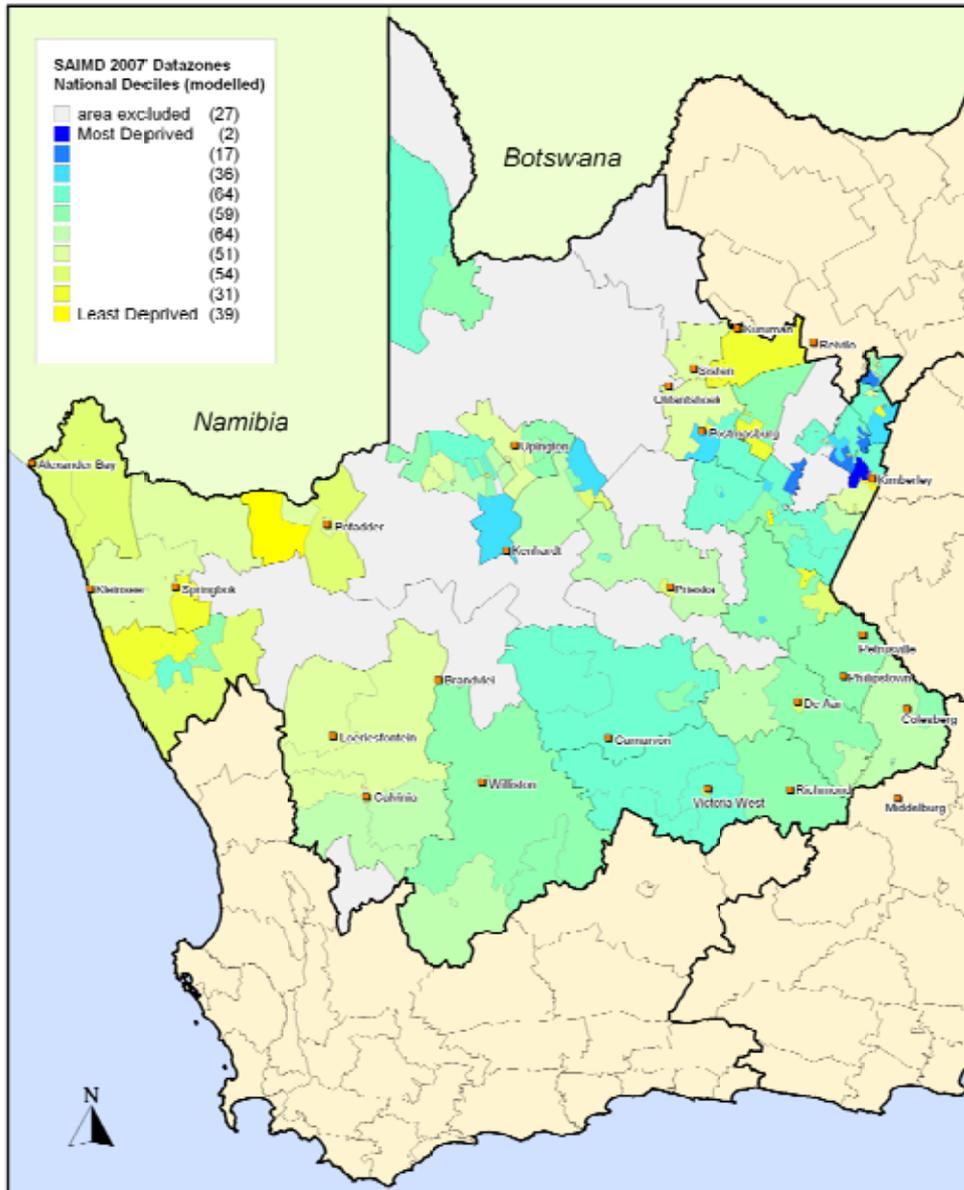
**SAIMD 2007 Datazones
National Deciles (modelled)**

area excluded	(286)
Most Deprived	(1061)
	(530)
	(320)
	(289)
	(220)
	(193)
	(151)
	(142)
	(130)
Least Deprived	(145)



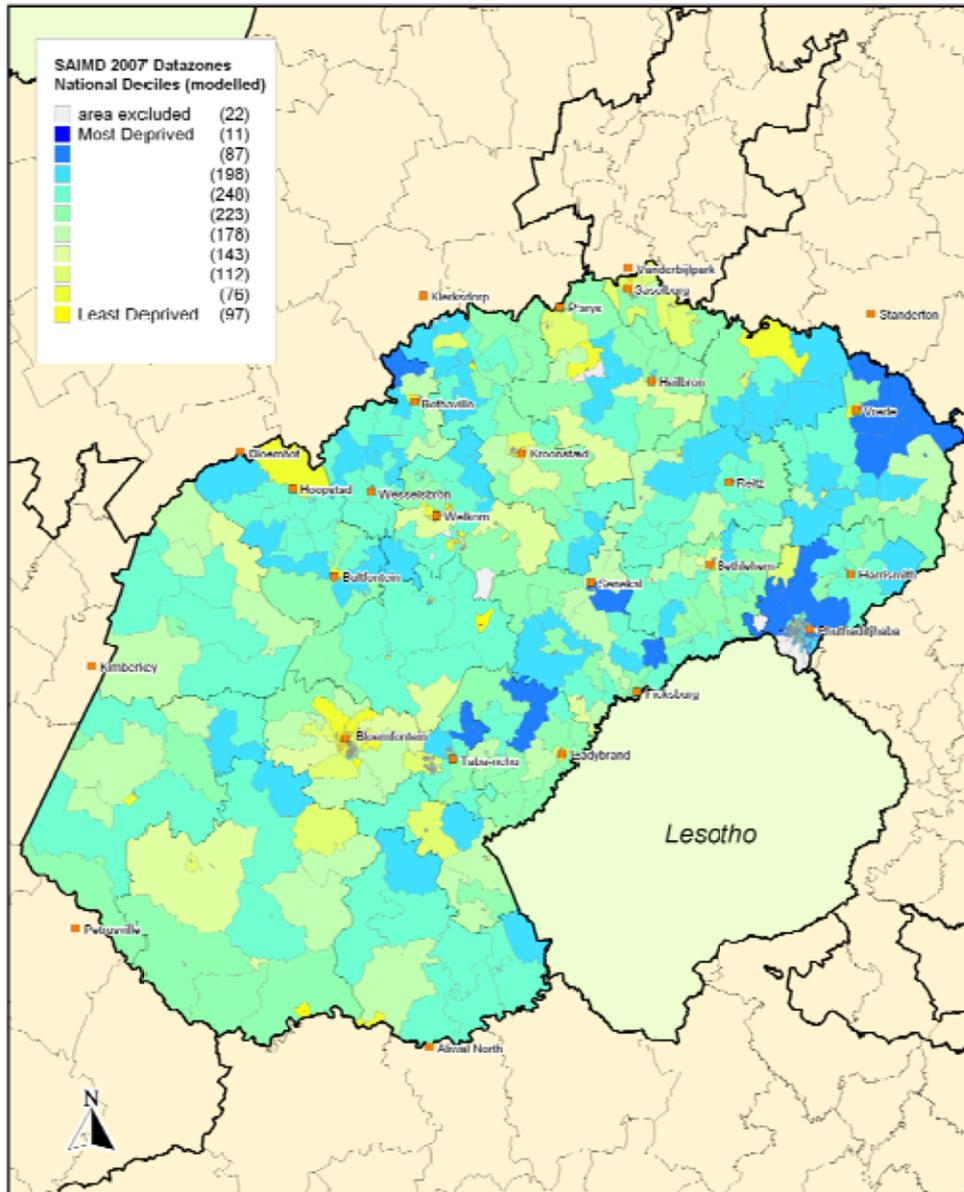
© Copyright: Centre for the Analysis of South African Social Policy, University of Oxford, UK, December 2009. Produced for the Department of Social Development of the Republic of South Africa and funded by the UK Department for International Development Southern Africa. Community Survey 2007 data supplied by Statistics South Africa. Digital Map data supplied by Statistics South Africa and Chief Directorate of Surveys and Mapping.

South African Index of Multiple Deprivation at Datazone Level 2007
Map 4 - Northern Cape. Modelled data.



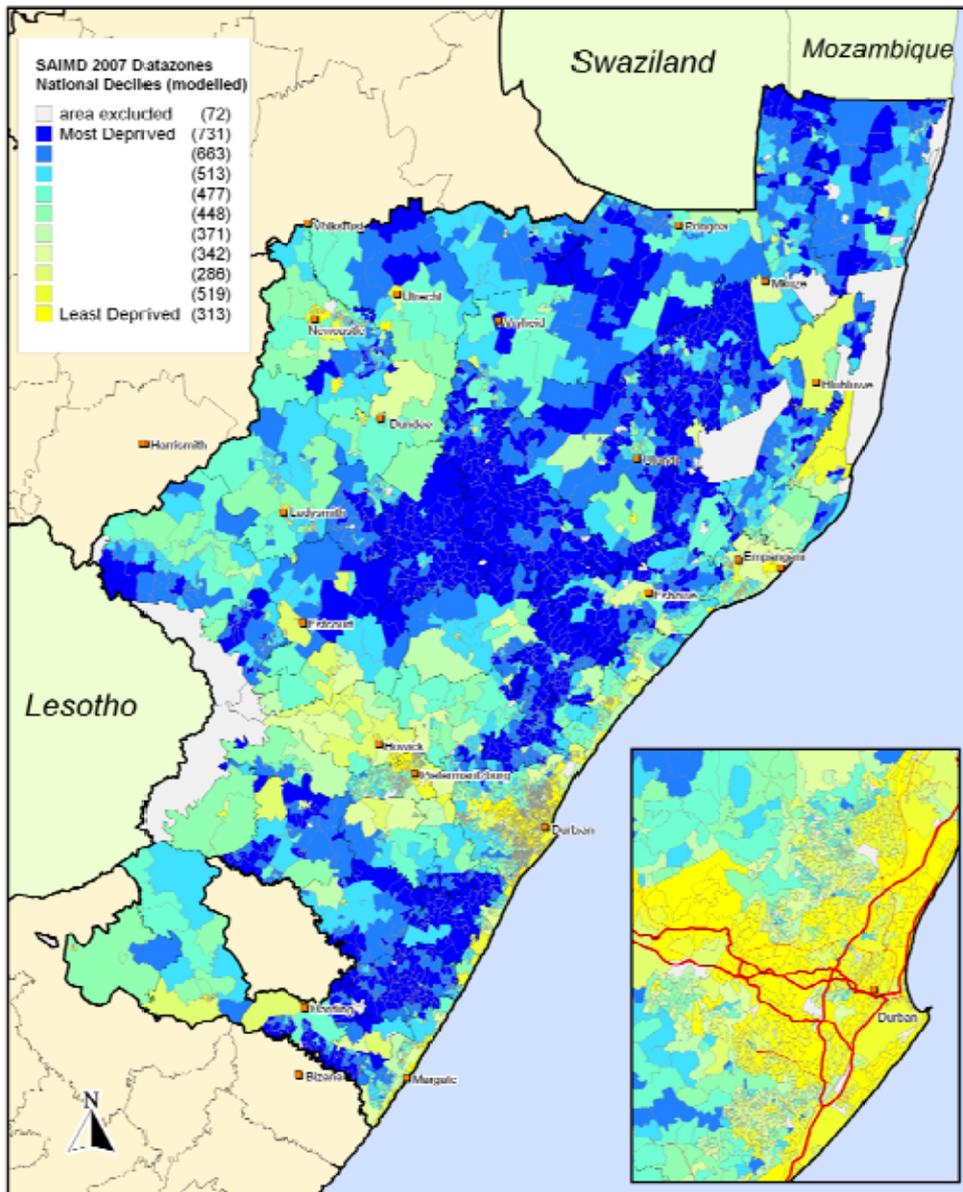
© Copyright, Centre for the Analysis of South African Social Policy, University of Oxford, UK, December 2005. Produced for the Department of Social Development of the Republic of South Africa and funded by the UK Department for International Development Southern Africa. Community Survey 2007 data supplied by Statistics South Africa. Digital Map data supplied by Statistics South Africa and Chief Directorate of Surveys and Mapping.

South African Index of Multiple Deprivation at Datazone Level 2007
Map 5 - Free State. Modelled data.



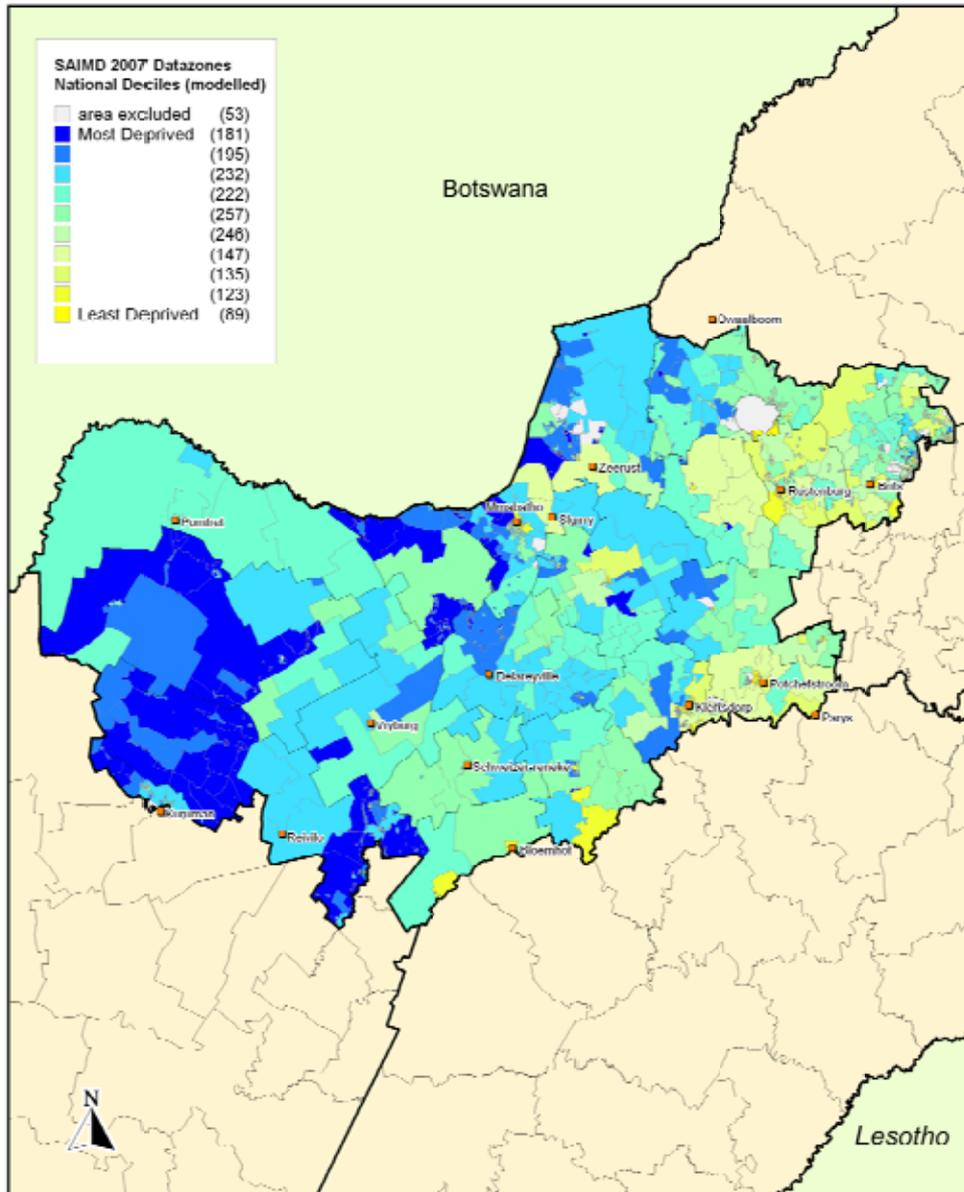
© Copyright: Centre for the Analysis of South African Social Policy, University of Oxford, UK, December 2008. Produced for the Department of Social Development of the Republic of South Africa and funded by the UK Department for International Development Southern Africa. Community Survey 2007 data supplied by Statistics South Africa. Digital Map data supplied by Statistics South Africa and Chief Directorate of Surveys and Mapping.

South African Index of Multiple Deprivation at Datazone Level 2007
Map 6 - KwaZulu-Natal. Modelled data.



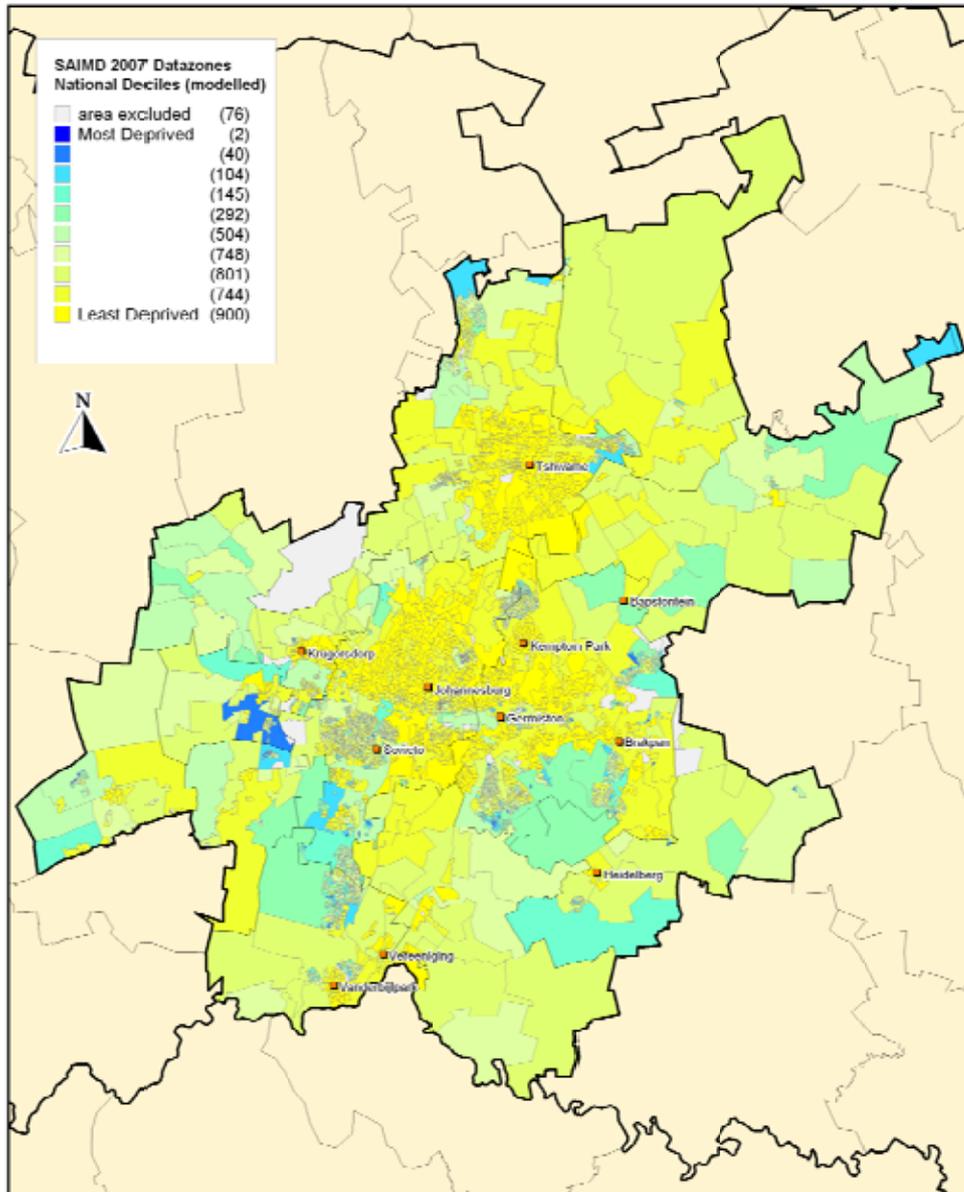
© Copyright, Centre for the Analysis of South African Social Policy, University of Oxford, UK, December 2005. Produced for the Department of Social Development of the Republic of South Africa and funded by the UK Department for International Development Southern Africa. Community Survey 2007 data supplied by Statistics South Africa. Digital Map data supplied by Statistics South Africa and Chief Directorate of Surveys and Mapping.

**South African Index of Multiple Deprivation at Datazone Level 2007
Map 7 - North West Province. Modelled data.**



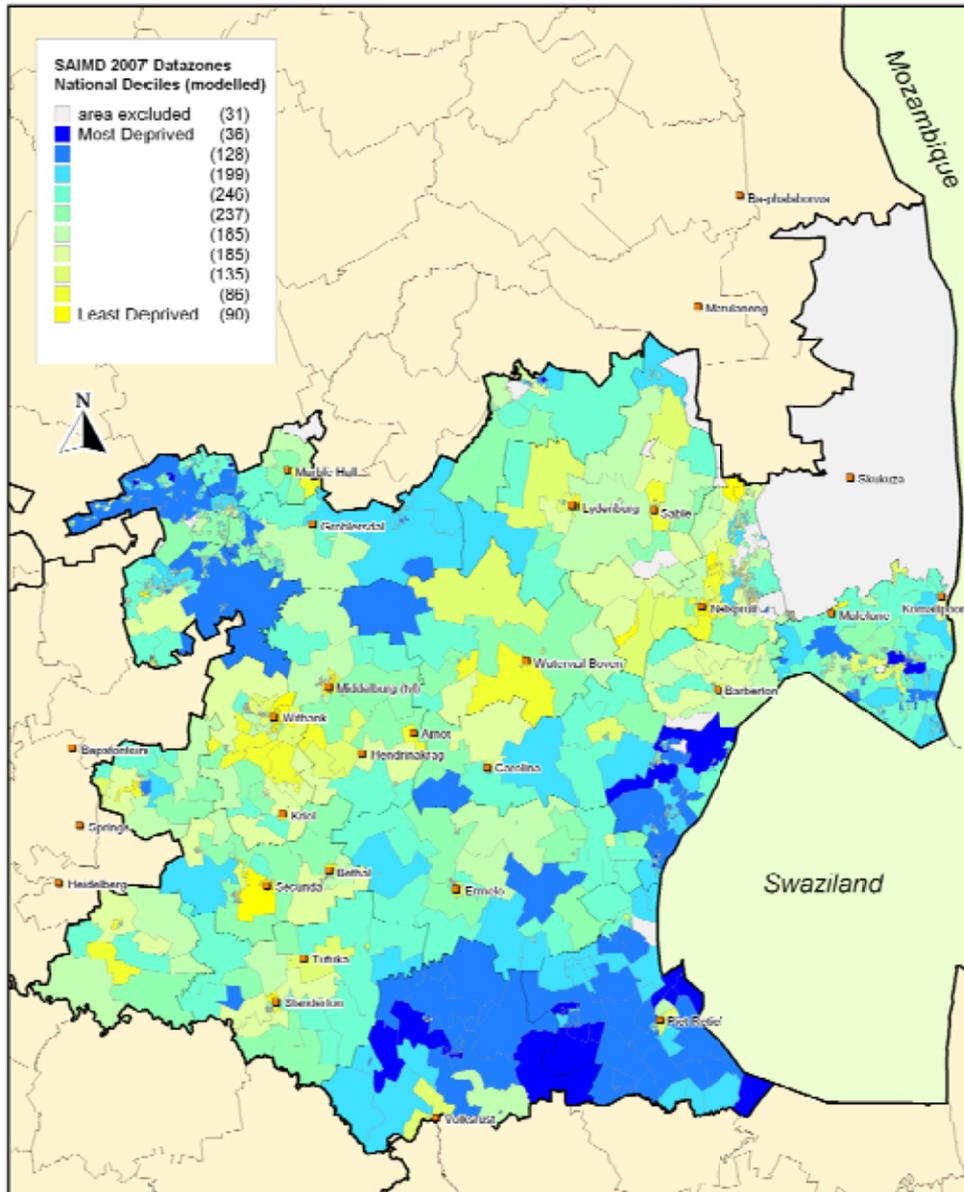
© Copyright: Centre for the Analysis of South African Social Policy, University of Oxford, UK, December 2008. Produced for the Department of Social Development of the Republic of South Africa and funded by the UK Department for International Development Southern Africa. Community Survey 2007 data supplied by Statistics South Africa. Digital Map data supplied by Statistics South Africa and Chief Directorate of Surveys and Mapping.

South African Index of Multiple Deprivation at Datazone Level 2007
Map 8 - Gauteng. Modelled data.



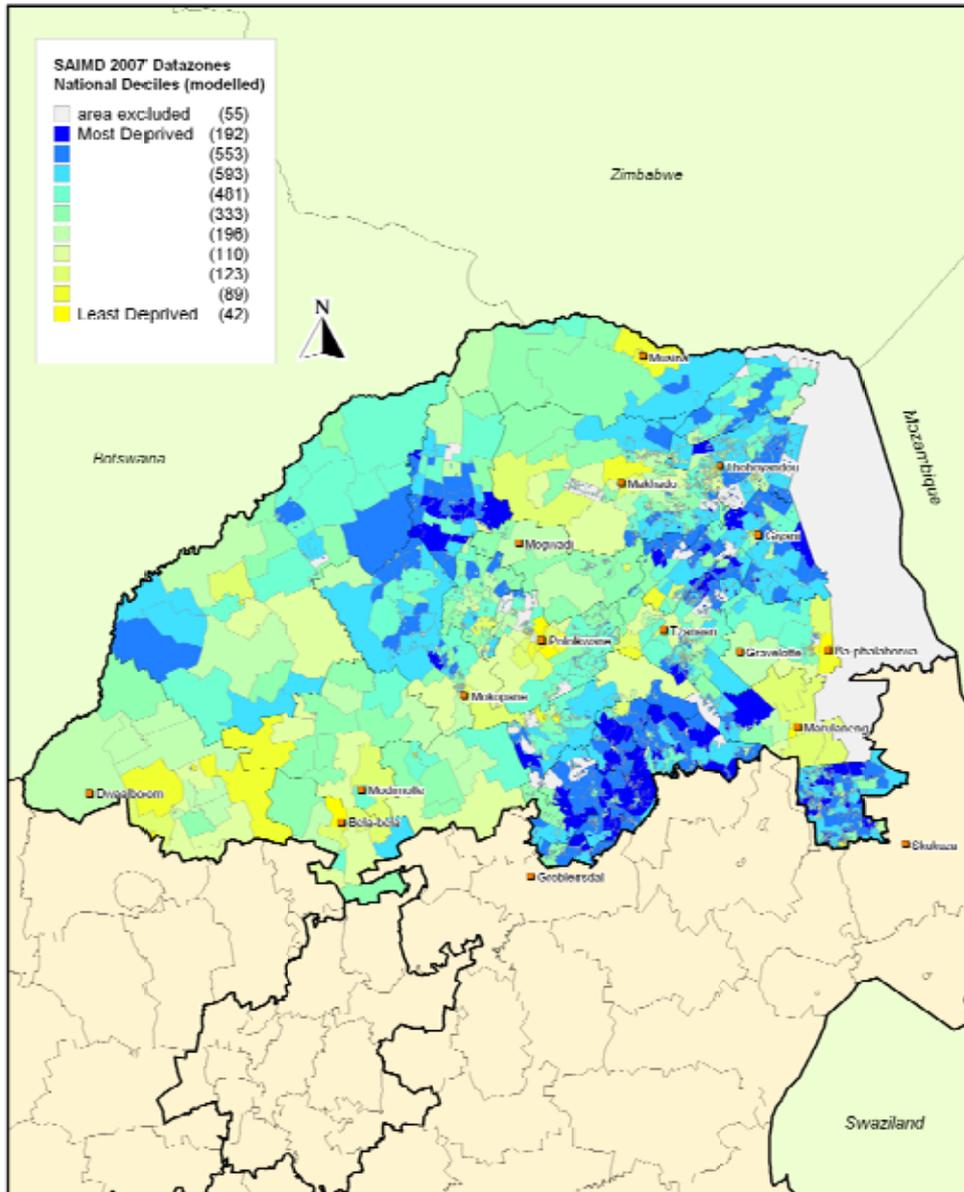
© Copyright: Centre for the Analysis of South African Social Policy, University of Oxford, UK, December 2008. Produced for the Department of Social Development of the Republic of South Africa and funded by the UK Department for International Development Southern Africa. Community Survey 2007 data supplied by Statistics South Africa. Digital Map data supplied by Statistics South Africa and Chief Directorate of Surveys and Mapping.

South African Index of Multiple Deprivation at Datazone Level 2007
Map 9 - Mpumalanga. Modelled data.



© Copyright: Centre for the Analysis of South African Social Policy, University of Oxford, UK, December 2008. Produced for the Department of Social Development of the Republic of South Africa and funded by the UK Department for International Development Southern Africa. Community Survey 2007 data supplied by Statistics South Africa. Digital Map data supplied by Statistics South Africa and Chief Directorate of Surveys and Mapping.

South African Index of Multiple Deprivation at Datazone Level 2007
Map 10 - Limpopo. Modelled data.



© Copyright, Centre for the Analysis of South African Social Policy, University of Oxford, UK, December 2005. Produced for the Department of Social Development of the Republic of South Africa and funded by the UK Department for International Development Southern Africa. Community Survey 2007 data supplied by Statistics South Africa. Digital Map data supplied by Statistics South Africa and Chief Directorate of Surveys and Mapping.

5 Comparison with the datazone-level SAIMD 2001

We know from the national analysis presented elsewhere that *absolute* levels of deprivation reduced slightly between 2001 and 2007 on each of the four domains (Wright and Noble, 2009).

The majority of datazones have also become less deprived in *absolute* terms since 2001. For the Income Domain 60.5% have become less deprived in absolute terms, while for the Employment Domain the figure is 76.5%, for the Education Domain it is 87.8% and for the Living Environment Domain it is nearly 70%. In many cases the improvement is very small.

The SAIMD is a composite relative measure and so it is not possible to measure absolute change for the overall SAIMD between 2001 and 2007. It is however possible to consider relative change (change in rank position) of datazones between these two time points.³ The following analyses look at this relative change.

Table 2 Most Deprived Datazones by Province - Modelled datazone-level SAIMD 2007 and datazone-level SAIMD 2001

	% of Province's Datazones in Most Deprived Decile 2007	% of Province's Datazones in Most Deprived Quintile 2007	% of Province's Datazones in Most Deprived Decile 2001 ⁴	% of Province's Datazones in Most Deprived Quintile 2001
Western Cape	0.0	0.1	0.1	0.6
Eastern Cape	33.4	50.0	29.9	47.4
Northern Cape	0.5	4.6	2.2	5.3
Free State	0.8	7.1	0.4	7.3
KwaZulu-Natal	15.7	29.9	18.4	32.4
North West	9.9	20.6	5.9	15.4
Gauteng	0.0	1.0	0.1	0.9
Mpumalanga	2.4	10.7	2.4	11.1
Limpopo	7.1	27.5	8.9	29.2

Table 2 is similar to Table 1 above. It presents the percentage of a province's datazones which are in the most deprived decile and quintile nationally. It repeats the last 2 columns of Table 1 which gives the picture for 2007 and also presents the same information for 2001.

³ The Appendix gives the overall correlation of the modelled datazone-level SAIMD 2007 with the datazone-level SAIMD 2001 (without the Health Domain).

⁴ For the purpose of comparing the modelled datazone-level SAIMD 2007 with 2001, a revised datazone-level SAIMD 2001 was created with no Health Domain as this could not be created for 2007.

Considering the percentages of their datazones in the most deprived decile, the Eastern Cape and North West are in a worse relative position while KwaZulu-Natal and Limpopo have improved somewhat.

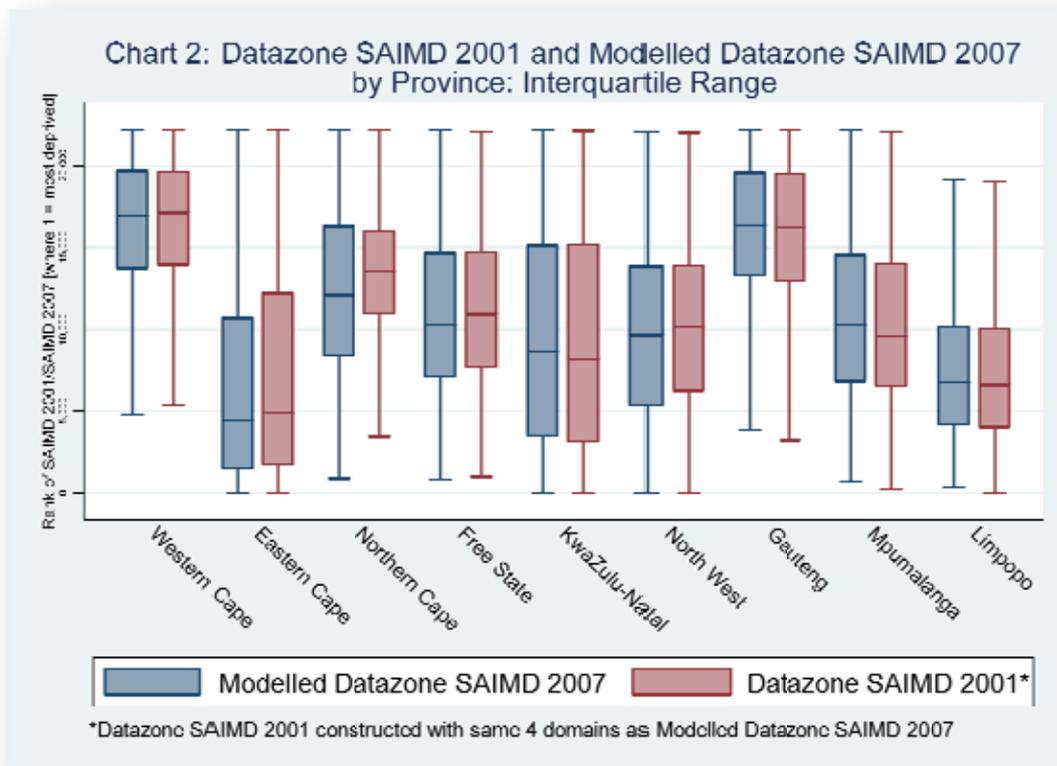


Chart 2 is a similar box plot to Chart 1 above. It reproduces the boxplots for the interquartile range for 2007 as shown in Chart 1 but places alongside them the equivalent box plots showing the interquartile range for 2001. The blue boxes (and their 'whiskers') indicate the position in 2007 while the red boxes (and their 'whiskers') present the 2001 picture.

As can be seen, the relative deprivation picture as presented by the interquartile range has changed little since 2001. The Eastern Cape, Northern Cape and North West have become slightly more deprived with the range narrowing in the Eastern Cape but expanding in the Northern Cape. Gauteng, Mpumalanga and Limpopo have all become very slightly less deprived relatively speaking.

In general, therefore, it seems that though deprivation has fallen nationally between 2001 and 2007 based on each of the four domains, the geographical distribution of deprivation has remained very similar between these two time points, and the former homelands in particular remain the most highly deprived areas in the country.

In conclusion, this report demonstrates that it is possible to produce plausible modelled estimates at datazone level for 2007. However, the statistics generated must remain experimental as, in the absence of other small area level data sources, validation beyond that presented here is hard to achieve.

References

- Alderman, H., Babita, M., Demombynes, G., Makhatha, N., & Ozler, B. (2003) 'How low can you go? Combining census and survey data for mapping poverty in South Africa', *Journal of African Economies*, 11(2):169–200.
- Avenell, D., Noble, M. and Wright, G. (2009) *South African datazones: A technical report about the development of a new statistical geography for the analysis of deprivation in South Africa at a small area level*, CASASP Working Paper No. 8, Oxford: Centre for the Analysis of South African Social Policy, University of Oxford.
- Ballas, D., Dorling, D., Anderson, B. and Stoneman, P. (2006) *Assessing the Feasibility of Producing Small Area Income Estimates*, Report to the Office of the Deputy Prime Minister, London: ODPM.
- Barnes, H., Noble, M., Wright, G and Dawes, A (2009) 'A Geographical Profile of Child Deprivation in South Africa', *Child Indicators Research*, doi:10.1007/s12187-008-9026-2.
- Barnes, H., Wright, G., Noble, M. and Dawes, A. (2007) *The South African Index of Multiple Deprivation for Children 2001*, Cape Town: Human Sciences Research Council Press.
- Dibben, C., Surender, R., Sims, A., Noble, M., Hill, A., Goldacre, M., Goodman, Lloyd, M., Wright, G. and Smith, G.A.N. (2001) *Health Poverty Index Scoping Project Report*, London: Department of Health.
- Noble, M., Babita, M., Barnes, H., Dibben, C., Magasela, W., Noble, S., Ntshongwana, P., Phillips, H., Rama, S., Roberts, B., Wright, G. and Zungu, S. (2006) *The Provincial Indices of Multiple Deprivation for South Africa 2001*, Oxford: University of Oxford, UK.
- Noble, M., Barnes, H., Wright, G., McLennan, D., Avenell, D., Whitworth, A., and Roberts, B. (2009a) *The South African Index of Multiple Deprivation 2001 at Datazone Level*, Pretoria: Department of Social Development.
- Noble, M., Barnes, H., Wright, G. and Roberts, B. (2009b) 'Small Area Indices of Multiple Deprivation in South Africa', *Social Indicators Research*.
- Statistics South Africa (2007) *Community Survey, 2007 (Revised Version)* Statistical Release P0301, Pretoria: Statistics South Africa.
- Statistics South Africa (2008) *Community Survey 2007: Unit Records Metadata*, Report No. 03-01-21 (2007), Pretoria: Statistics South Africa.
- Wright, G., Barnes, H., Noble, M., and Dawes, A. (2009a) *The South African Index of Multiple Deprivation for Children 2001 at Datazone Level*, Pretoria: Department of Social Development.
- Wright, G. and Noble, M. (2009) *The South African Index of Multiple Deprivation 2007 at Municipality Level*, Pretoria: Department of Social Development.

Wright, G., Noble, M., Barnes, H. and Noble, S. (2009b) *The South African Index of Multiple Deprivation for Children 2007 at Municipality Level*, Pretoria: Department of Social Development.

Appendix

In this Appendix the correlations of the modelled datazone-level SAIMD 2007 and its component domains with their 2001 counterparts are presented. As the datazone-level SAIMD 2001 included a Health Domain which could not be replicated in 2007, it was first necessary to re-create a datazone-level SAIMD 2001 which excludes a Health Domain. Comparisons are therefore made with this modified datazone-level SAIMD 2001.

The datazone-level SAIMD 2007 correlates 0.9451 ($p < .0001$) with the datazone-level SAIMD 2001. This can be seen graphically in the scatter plot below (Chart 3)

The Spearman correlations for the other domains are given in the table below – again demonstrating relatively high correlations.

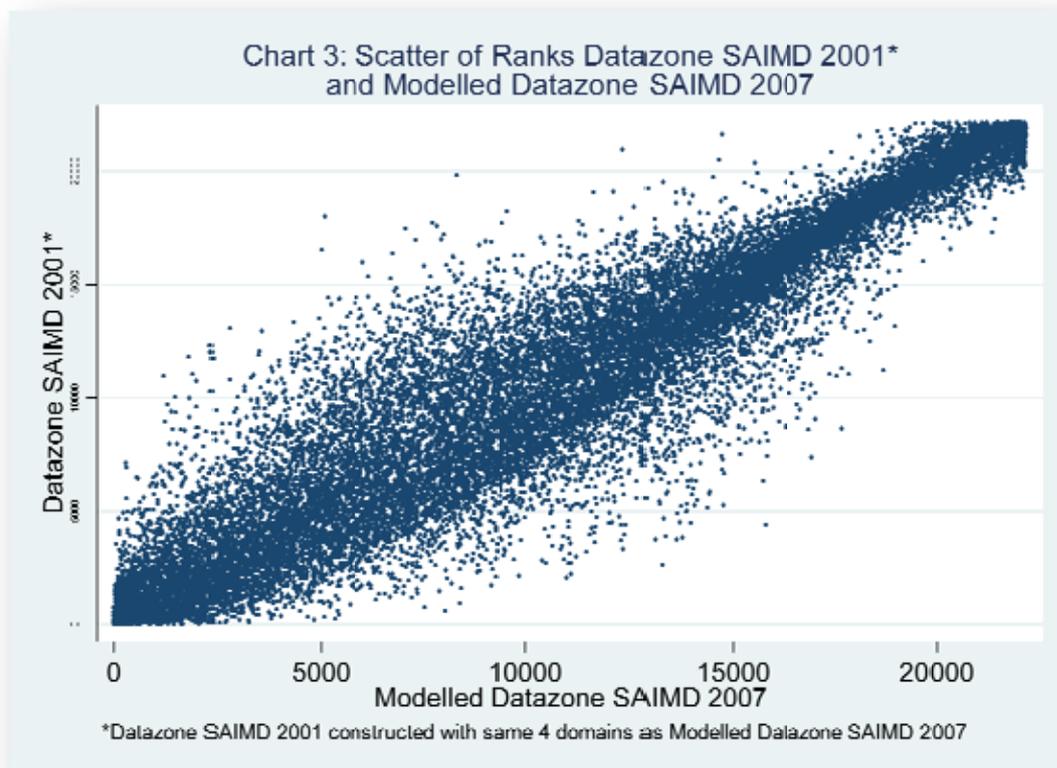


Table 3 Spearman's Correlations for the Individual Datazone Domain Scores 2001 and 2007

	Income 2007	Employment 2007	Education 2007	Living Environment 2007
Income 2001	0.8759*			
Employment 2001		0.8633*		
Education 2001			0.9573*	
Living Environment 2001				0.9044*

* $P < .0001$