

**Centre for the Analysis of South African Social Policy  
University of Oxford**

# **Modelling Eligibility for the Child Support Grant**

**Helen Barnes and Michael Noble**



**Take Up of Social Grants Project  
June 2006**



**Take Up of Social Grants Project**

**Project 4.2.4 of the Department of Social  
Development's Social Policy Analysis Programme**

# **Modelling Eligibility for the Child Support Grant**

**By Helen Barnes and Michael Noble**

**June 2006**

## **Acknowledgements**

The broader project from which this report stems was undertaken by a research team within the Centre for the Analysis of South African Social Policy at the Department of Social Policy and Social Work, University of Oxford, UK. The project is funded by the UK Department for International Development Southern Africa.

The team comprised Michael Noble, Gemma Wright, Helen Barnes, Stefan Noble, Phakama Ntshongwana, Roxana Gutierrez-Romero, David McLennan and David Avenell. Ingrid Woolard from the University of Cape Town and Charles Simkins from the University of the Witwatersrand acted as consultants to the overall project.

The team would also like to extend thanks to Francie Lund, Roxana Gutierrez-Romero and Gemma Wright for their helpful comments.

## **Recommended citation**

Barnes, H. and Noble, M.(2006) *Modelling Eligibility for the Child Support Grant*, Project on Take Up of Social Grants, National Department of Social Development, South Africa.

## **Disclaimer**

The University of Oxford has taken care to ensure that the information in this report is 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. We do not guarantee that the information in this report is fit for any particular purpose. We do not accept responsibility for any alteration or manipulation of our data once it has been released. The information supplied is subject to the conditions of the contract between DSD and the University of Oxford.

# Contents

1	Introduction.....	4
2	Determining eligibility for CSG .....	5
	Preparation .....	5
	Identifying primary caregiver and spouse of primary caregiver using Census data .....	5
	Modified Woolard method to identify primary caregiver and spouse of primary caregiver .....	6
	Estimating eligibility.....	6
	Adjusting estimates for population change.....	6
	Imputing income .....	7
3	Modelling eligibility for CSG.....	8
	The logistic regression model .....	8
	Probability of eligibility.....	8
	Conclusion .....	12
	References.....	13

# 1 Introduction

This report is one of a series arising from a project entitled ‘Take Up of Social Grants’. The project is being carried out by the Centre for the Analysis of South African Social Policy (CASASP), which is based in the Department of Social Policy and Social Work at the University of Oxford in England. This project forms part of the South African National Department of Social Development’s (DSD) Social Policy Analysis Programme. This in turn is part of the broader ‘SACED’ Programme (Programme for Strengthening Analytical Capacity and Evidence Based Decision Making), and is funded by the UK Government Department for International Development Southern Africa.

In order to calculate take up rates, it is necessary to estimate the number of children throughout South Africa who are *eligible* to receive the Child Support Grant (CSG). This has been carried out by CASASP using the 2001 Census, and from this estimates of take up of CSG in January 2004 (Noble et al., 2005a) and January 2005 (Noble et al., 2005b) have been produced. This report details the methodology used to estimate eligibility for CSG in January 2005 using the 10% sample of the 2001 Census (**Section 2**) and then describes the development of a logistic regression model that examines the probability of caregivers with different characteristics being eligible for CSG on behalf of their child (**Section 3**). Children aged 0 to 10 inclusive are taken into account as this is the age group considered in Noble et al. (2005b).

## 2 Determining eligibility for CSG

This section describes the methodology developed by CASASP to estimate eligibility for CSG. The method involved the use of Statistics South Africa's 10% sample of the 2001 Census to identify the child's primary caregiver, the caregiver's income (and the income of their spouse if they had a spouse) and the type of housing (formal/informal) and area in which the children live (rural or urban). From this it was possible to establish whether the criteria of the means test were met. Three Census files relating to persons, households and housing, and geography were used in the estimation process.

### Preparation

Several variables from the household level file were merged into the person level file of the 10% sample of the 2001 Census. People in institutions and in other non-housing units were then deleted, as were people whose person number was greater than 30. The person file was restructured to become a household level file with a number of arrays relating to individual people within the household. The arrays included person number, age, sex, relationship to head of household, income, weight<sup>1</sup>, person number of spouse, person number of mother and person number of father. Surplus variables were dropped.

### Identifying primary caregiver and spouse of primary caregiver using Census data

Datasets that have been used previously to estimate eligibility (2003 General Household Survey, 2000 Income and Expenditure Survey) do not contain information on the biological parents or the primary caregiver of each child. This has to be derived from the data using a series of assumptions. However, the 2001 Census does allow the biological mother and biological father of each child to be identified, if they are present in the household, and also the spouse<sup>2</sup> of each present biological parent, whether biologically related to the child or not. The primary caregiver was assumed to be the biological mother if present in the household. If the biological mother was not present in the household, the primary caregiver was assumed to be the biological father. Using the spouse variable it was then possible to identify the spouse of the primary caregiver.

---

<sup>1</sup> Both the household and person sample files contain a weight variable. This is the adjustment factor for undercount/overcount multiplied by ten to inflate the samples to the relevant population, and was derived from the Post Enumeration Survey carried out a month after the Census.

<sup>2</sup> The Census question asks 'Who in the household is (the person's) spouse or partner?' Although the Social Assistance Regulations state that the income of the spouse of the primary caregiver should be taken into account, for this purpose it was only possible to include the income of the partner of the primary caregiver, whether they are married to the primary caregiver or not. The term spouse is used henceforth to mean either spouse or partner.

## **Modified Woolard method to identify primary caregiver and spouse of primary caregiver**

The primary caregiver and/or spouse of approximately 20% of children under eleven could not be identified by the above criteria and so a modified version of the technique that Woolard applied to identify a child's primary caregiver and spouse in the Income and Expenditure Survey<sup>3</sup> was used. In these cases, the primary caregiver was identified as the oldest woman in the household aged 13-40 at the time of the birth of the child; and if none then the youngest woman in the household aged 41 or over at the time of the birth of the child; and if none then the oldest male in the household aged 13-40 at the time of the birth of the child; and if none then the youngest male in the household aged 41 or over at the time of the birth of the child.

## **Estimating eligibility**

Geographical variables were merged onto the file from the geography file of the 10% sample of the Census and a flag was created to identify households in urban areas in formal housing. This is necessary because eligibility for CSG is dependent partly on the type of housing unit and whether the house is in an urban or rural location. The income threshold for people living in formal housing in urban areas is R 9 600 pa, whereas the income threshold for people living elsewhere is R 13 200 pa.

Children under eleven years were flagged as eligible if the combined income of the primary caregiver and his/her spouse was below the relevant income threshold. Although the income variable in the Census is in bands and so cannot directly be used to determine eligibility as the means test thresholds fall in the middle of two income bands, there are various ways of assigning an actual income to the band. In this work the midpoint of each income band was calculated and the income band value was replaced with this midpoint value.

Separate variables were created for each year of age and sex (e.g. male 0 year olds, female 0 year olds, male 1 year olds, female 1 year olds). A household level count of children in each of the age/sex groups whose caregivers were eligible was created. This is the sum of the person weights for all relevant children (i.e. those flagged as eligible) in the separate age/sex groups.

## **Adjusting estimates for population change**

The municipality level counts of children whose caregivers were eligible for CSG were then adjusted to take into account population change between October 2001 and January 2005 using mid year population estimates from Stats SA. The following assumptions were made:

---

<sup>3</sup> A model of take up produced by Ingrid Woolard for Lydia Ntenga at DSD in 2004. A detailed comparison of the methods and estimates of eligibility for CSG produced by Woolard (2004), Budlender et al. (2005) and Noble et al. (2005a; 2005b) is presented in a report by Wright (2006).

- 1) The number of children in an institution has remained constant between October 2001 (Census date) and January 2005 at province level.
- 2) The rate of population change between mid 2004 and mid 2005 is constant over this time and the same for each age/sex group.
- 3) The institutions have remained in the same municipality between 2001 and 2005.
- 4) The rate of population change between October 2001 and January 2005 occurs evenly across all municipalities in a province.

A final count of children under eleven whose caregivers were estimated to be eligible to receive CSG in January 2005 was created at municipality and province level.

## Imputing income

A known problem with the 2001 Census is that there are large numbers of people reporting zero income or for whom no income data are recorded. Before releasing any Census products, Statistics South Africa used a single hot deck imputation to allocate values for unavailable, unknown, incorrect or inconsistent responses. Notwithstanding this, there is some concern with the large number of individuals claiming to have zero income, and in particular, the high proportion of individuals where the recorded income is not plausible, for example, individuals recorded as employed but claiming to have zero income.

In order to correct implausible and missing values in the income variable, a statistical procedure known as multiple imputation was employed by CASASP. This imputation was performed on the missing and implausible income values in the 10% sample of the 2001 Census and the results were then compared to the hot deck imputations produced by Statistics South Africa and the multiple imputations produced by Ardington and colleagues at the University of Cape Town (Ardington et al., 2005). The three different methods produce very similar results. Each method assigns roughly the same number of people to the first three income bands (84 to 85%).<sup>4</sup> This is important for work on take up as caregivers in the first three income bands are most likely to meet the requirements of the CSG means test.

CASASP's imputation process was carried out ten times and resulted in ten imputed datasets, each of which produced similar income distributions.<sup>5</sup> Although the method to model eligibility for CSG described in **Section 3** did not use the multiply imputed data, it is important to note how consistent the income variable is across the three different imputation methods. The original Stats SA hot deck imputations were used for the analysis in this report as it was felt that, on the basis of the multiple imputation work, they were sufficiently reliable for this purpose.

---

<sup>4</sup> This is for the whole Census sample (i.e. both cases where the income was imputed and cases where the income was neither implausible nor zero).

<sup>5</sup> A full account of CASASP's imputation procedure and a comparison of the three imputation methods is given in Barnes et al. (2006).

### 3 Modelling eligibility for CSG

#### The logistic regression model

Using the method described in **Section 2**, it has been estimated that in January 2005, of all children under eleven in South Africa, 67% have caregivers who are eligible for CSG (Noble et al., 2005b). These are children who live in families where the combined income of the primary caregiver and their spouse (if he/she exists) is less than R 9 600 for households that are formal dwellings in urban areas (urban-formal) and less than R 13 200 for households that are informal dwellings in urban areas (urban-informal) and all types of housing units in rural areas.

It is possible to determine a variety of characteristics of children whose caregivers are eligible for CSG, using the Census. This can be undertaken using simple bivariate analysis. However a more nuanced approach is to use a multivariate technique. A stepwise logistic regression was run to examine the probability of children with different characteristics being eligible for CSG.<sup>6</sup> Ten models were created with a new variable introduced and controlled for in each model. Each new variable improved the fit of the model and so was retained in subsequent models. The ten variables used are: population group of the child (it is likely that the caregiver is the same population group as the child for whom they are caring); four household variables (piped water, toilet facility, energy source for lighting, refrigerator); four variables relating to the child's caregiver (employment status, whether disabled, age, whether a lone parent); and province.

A number of variables were considered for inclusion in the model but were eventually dropped as they were either variables that were used as part of the means test calculation (e.g. urban/rural, formal/informal, caregiver's income) or variables that had a number of missing values (e.g. occupation) or variables that preliminary testing showed did not improve the fit of the model, often because a similar variable was already in the model (e.g. certain other household goods).

#### Probability of eligibility

The first model simply controlled for population group. It was found that compared to White children, children who are Black African are over 43 times more likely to have caregivers who are eligible for CSG, while for children who are Coloured or Indian/Asian, the odds ratios are lower (but still striking): Coloured children are almost 11 times more likely than White children to have caregivers who are eligible

---

<sup>6</sup> The analysis in this section focuses on the *children* rather than beneficiaries (i.e. children who have caregivers who are eligible for CSG). Therefore when it is claimed below that a child with a particular characteristic (whether a characteristic of his/herself, household or caregiver) is x times more likely to be eligible, this means that the caregiver of the child in question is x times more likely to be eligible than the caregiver of another child and not that the child his/herself is x times more likely to be eligible for CSG.

and Indian/Asian children are almost 3 times as likely to have caregivers who are eligible. The odds ratios from Model 1 and the other nine models are given in **Table 1**.

When the binary variable piped water (no piped water or piped water over 200m from dwelling compared to piped water less than 200m from dwelling, inside yard or inside dwelling) is controlled for (Model 2), the impact of population group is greatly reduced for Black Africans. Controlling for population group, children who have no piped water or piped water over 200m from their dwelling are almost 3 times more likely to have a caregiver who is eligible for CSG than those living in dwellings where there is piped water less than 200m away (including inside the yard or dwelling). As will be seen, the odds of children who have poor access to piped water having a caregiver who is eligible are reduced when other effects are entered into the model and thus controlled for. Models 3, 4 and 5 look at the impact of controlling for toilet facility, energy source for lighting, and whether the house has a refrigerator.

Model 6 shows that, controlling for population group and the four household characteristics mentioned above, children whose caregivers are disabled are 1.6 times more likely to be eligible for CSG. However, after controlling for the employment status and lone parenthood of the caregivers, this diminishes to 1.2 (Model 8).

A binary variable measuring whether or not the caregiver is a lone parent was added into the model (Model 8). Controlling for other factors including employment status, eligibility for CSG is 14 times more likely for a child whose caregiver is a lone parent<sup>7</sup> than for a child whose caregiver has a spouse. Employment status of the caregiver is another variable that has a large effect. When nine other variables are controlled for (the final model, Model 10), eligibility for CSG is 12 times more likely for a child whose caregiver is not in employment (unemployed or not economically active) than for a child whose caregiver is employed.

As more variables are controlled for, the odds ratio decreases for all non-White population groups. So for example, although the caregivers of Black African children start out as over 40 times more likely to be eligible for CSG than the caregivers of White children, this can be explained by factors other than simply their population group, such as whether or not they have piped water, whether the caregiver is employed, whether the caregiver is a lone parent and so on. When nine of these factors are controlled for, Black African children end up as 6.6 times more likely than White children to have caregivers who are eligible for CSG. Coloured children end up as 5.2 times more likely than White children to have caregivers who are eligible for CSG, and Indian/Asian children are 1.8 times more likely than White children to have eligible caregivers.

---

<sup>7</sup> The term lone parent refers to any caregiver of a child, whether the child's biological parent or not, who does not have a spouse in the household.

**Table 1: Odds ratios from logistic regression modelling eligibility for CSG (models 1-5)**

<b>Effect</b>	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 4</b>	<b>Model 5</b>
Intercept	-0.8366	-0.4360	-0.2829	-0.1460	-0.1005
<i>Population group - White</i>	1.000	1.000	1.000	1.000	1.000
Population group - Black African	43.449	30.774	21.199	19.603	16.137
Population group - Coloured	10.687	10.559	9.501	9.192	7.861
Population group - Indian/Asian	2.739	2.754	2.728	2.716	2.680
<i>Piped water - &lt; 200m from dwelling</i>		1.000	1.000	1.000	1.000
Piped water - > 200m from dwelling/none		2.969	1.972	1.656	1.530
<i>Toilet facility - Flush toilet or VIP</i>			1.000	1.000	1.000
Toilet facility - No flush toilet or VIP			2.569	2.244	2.001
<i>Energy source for lighting - Electricity</i>				1.000	1.000
Energy source for lighting - No electricity				1.883	1.174
<i>Refrigerator in house - Yes</i>					1.000
Refrigerator in house - No					2.680
<i>Disability - No</i>					
Disability - Yes					
<i>Employment status - In employment</i>					
Employment status - Not in employment					
<i>Lone parent - No</i>					
Lone parent - Yes					
Age					
<i>Province - Limpopo</i>					
Province - Western Cape					
Province - Eastern Cape					
Province - Northern Cape					
Province - Free State					
Province - KwaZulu-Natal					
Province - North West					
Province - Gauteng					
Province - Mpumalanga					
N <sup>8</sup>	853430	853430	853430	853430	853430
Nagelkerke's R-Square	0.1976	0.2471	0.2802	0.2916	0.3215
-2 Log Likelihood	933318.6	897545.4	872787.5	864089.7	840841.6
Wald	79456.5	105366.8	125904.9	130586.7	146620.2
DF	3	4	5	6	7

<sup>8</sup> The logistic regression was run on a dataset which was not weighted to inflate the estimates to the relevant population, and cases were deleted where information was missing for any of the ten variables included in the models. For these reasons the number of children included in the logistic regression does not match the number of eligible children reported in Noble et al. (2005b).

**Table 1 (cont.): Odds ratios from logistic regression modelling eligibility for CSG  
(models 6-10)**

<b>Effect</b>	<b>Model 6</b>	<b>Model 7</b>	<b>Model 8</b>	<b>Model 9</b>	<b>Model 10</b>
Intercept	0.1136	-0.3821	-0.0437	0.4056	0.5172
<i>Population group - White</i>	1.000	1.000	1.000	1.000	1.000
Population group - Black African	16.013	13.196	6.811	6.927	6.648
Population group - Coloured	7.833	7.939	5.338	5.315	5.280
Population group - Indian/Asian	2.681	1.967	2.044	2.018	1.815
<i>Piped water - &lt; 200m from dwelling</i>	1.000	1.000	1.000	1.000	1.000
Piped water - > 200m from dwelling/none	1.526	1.347	1.199	1.225	1.187
<i>Toilet facility - Flush toilet or VIP</i>	1.000	1.000	1.000	1.000	1.000
Toilet facility - No flush toilet or VIP	1.996	1.839	1.872	1.890	1.690
<i>Energy source for lighting - Electricity</i>	1.000	1.000	1.000	1.000	1.000
Energy source for lighting - No electricity	1.172	1.114	1.228	1.228	1.254
<i>Refrigerator in house - Yes</i>	1.000	1.000	1.000	1.000	1.000
Refrigerator in house - No	2.672	2.491	2.791	2.777	2.756
<i>Disability - No</i>	1.000	1.000	1.000	1.000	1.000
Disability - Yes	1.590	1.298	1.192	1.301	1.271
<i>Employment status - In employment</i>		1.000	1.000	1.000	1.000
Employment status - Not in employment		6.951	12.431	12.230	12.346
<i>Lone parent - No</i>			1.000	1.000	1.000
Lone parent - Yes			14.032	13.603	13.493
Age				0.988	0.987
<i>Province - Limpopo</i>					1.000
Province - Western Cape					0.740
Province - Eastern Cape					*0.978
Province - Northern Cape					1.441
Province - Free State					1.567
Province - KwaZulu-Natal					**0.997
Province - North West					0.868
Province - Gauteng					0.618
Province - Mpumalanga					***0.994
N	853430	853430	853430	853430	853430
Nagelkerke's R-Square	0.3230	0.4414	0.6010	0.6023	0.6071
-2 Log Likelihood	839652.4	740711.7	586000.9	584567.3	579504.4
Wald	147105.9	196170.8	201862.0	202104.0	202419.1
DF	8	9	10	11	19

p < 0.001 except where indicated

\* p = 0.4967

\*\* p = 0.1317

\*\*\* p = 0.3962

## Conclusion

Ten variables were entered into a logistic regression model to estimate the odds of eligibility for CSG for the caregiver of a child with particular characteristics. Of these ten variables, the binary variables employment status of the caregiver and whether or not the caregiver is a lone parent have the largest effect (see Model 10). This is unsurprising given that the means test to determine eligibility for CSG takes into account the income of the caregiver and spouse of the caregiver. What is perhaps surprising is the magnitude of the effect. A child whose caregiver is not in employment is over 12 times more likely to be eligible for CSG than a child whose caregiver is in employment, while a child whose caregiver is a lone parent is almost 13.5 times more likely to be eligible than a caregiver who has a spouse in the household.

The variable with the next largest effect is population group. Compared to a White child, a Black African child is over 6.5 times more likely to have a caregiver who is eligible for CSG. A Coloured child is 5 times more likely than a White child to have a caregiver who is eligible for CSG. When population group was the only variable in the model (Model 1), the respective odds ratios for Black African and Coloured children were 43.5 and 10.6 (times more likely than White children to have caregivers who are eligible for CSG). This probability was reduced by controlling for other factors, and it is likely that it would be further reduced if other variables had been included that tap into, and therefore control for, the effects of the legacy of apartheid, many of which are not found in the Census.

All other variables introduced into the models did not have a particularly large effect when controlling for nine other variables (odds ratios of between 0.6 and 2.8). The highest of these odds ratios was for the binary variable that measured whether or not there is a refrigerator in the household. Children without a refrigerator in the household are 2.8 times more likely than children with a refrigerator in the household to have a caregiver who is eligible for CSG.

The logistic regression model gives an interesting account of some of the factors underlying eligibility for CSG. Furthermore, eligibility could be taken as a proxy for poverty, and with certain amendments to the model it would be possible to analyse the likelihood of children with certain characteristics (either personal, household or related to their caregivers or other adults in the household) being in poverty.

## References

Ardington, C., Lam, D., Leibbrandt, M. and Welch, M. (2005) 'The sensitivity of estimates of post-apartheid changes in South African poverty and inequality to key data imputations', CSSR Working Paper No. 106, Cape Town: University of Cape Town.

Barnes, H. Gutierrez-Romero, R. and Noble, M. (2006) 'Multiple imputation of missing data in the 2001 South African Census', University of Oxford, UK (available on <http://www.casasp.ox.ac.uk>).

Budlender, D., Rosa, S. and Hall, K. (2005) 'The cost of the means test for the child support grant', A joint working paper by the Children's Institute and the Centre for Actuarial Research, University of Cape Town.

Noble, M., Wright, G., Barnes, H., Noble, S., Ntshongwana, P., Gutierrez-Romero, R., McLennan, D. and Avenell, D. (2005) *The Child Support Grant: A Sub-Provincial Analysis of Eligibility and Take Up in January 2004*, Project on Take Up of Social Grants, National Department of Social Development, South Africa.

Noble, M., Wright, G., Barnes, H., Noble, S., Ntshongwana, P., Gutierrez-Romero, R. and Avenell, D. (2005) *The Child Support Grant: A Sub-Provincial Analysis of Eligibility and Take Up in January 2005*, Project on Take Up of Social Grants, National Department of Social Development, South Africa.

Woolard, I.D. (2004) 'Constructing the DSD Model (Version 1.2)'.

Wright, G. (2006) *A Comparison of Province-level Estimates of Eligibility for the Child Support Grant*, Project on Take Up of Social Grants, National Department of Social Development, South Africa.

