THE UNIVERSITY OF MELBOURNE

Melbourne Institute Working Paper Series

Working Paper No. 16/05

A Micro-Macro Model for South Africa: Building and Linking a Microsimulation Model to a CGE Model

Nicolas Hérault

MELBOURNE INSTITUTE
of Applied Economic and Social Research

A Micro-Macro Model for South Africa: Building and Linking a Microsimulation Model to a CGE Model*

Nicolas Hérault
Centre d'Économie du Développement (IFReDE-GRES)
Université Montesquieu Bordeaux IV
and
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 16/05

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 0 7340 3195 5

November 2005

* This is a revised version of the paper originally prepared for the "8th Labour Econometrics Workshop", Center for Applied Economic Research, The University of New South Wales, Sydney, Australia, August 12-13, 2005. I would like to thank Jean-Pierre Lachaud, Daniela Casale, James Thurlow, Colette Muller, John Cockburn, Ramos Mabugu, Margaret Mabugu, Kalie Pauw and John Creedy for their suggestions and comments. I acknowledge support from the Melbourne Institute of Applied Economic and Social Research at the University of Melbourne, where I am visiting from 1 April to 31 December 2005. Specifically I am grateful to Guyonne Kalb for her assistance and invaluable comments on various aspects of this work. The support of the National Research Foundation in providing the two datasets used in this paper is also acknowledged.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 2100**
***Fax* (03) 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

This paper describes a newly-built micro-macro model for South Africa. A computable general equilibrium (CGE) model and a microsimulation (MS) model are combined in a sequential approach in order to build an effective tool to assess the effects of various macroeconomic policies and shocks on South African households. The CGE model is used to simulate the macro-changes in the structure of the economy after the policy change or the macro-shock. In a second step, these changes are passed on to the MS model. Micro-macro consistency equations, along with the direct transmission of prices, ensure that macro-changes are fully transmitted from the CGE to the MS model. Given any change in the macroeconomic structure of the economy predicted by the CGE model, the MS model predicts how individual agents modify their behaviours and how their incomes are affected, while accounting for individual heterogeneity.

1. Introduction

This paper aims to deal with the microeconomic effects of macroeconomic policies or shocks in South Africa. This is achieved by building a micro-macro model. In the South African context, where poverty and inequality both stand at a high level, it is important to identify the winners and losers of any policy change, so that the impact on poverty and inequality can be assessed in detail.

Generally speaking, the microeconomic effects of a policy are frequently ambiguous. Often the implementation of a macroeconomic policy benefits some households, while other households are made worse off or are unaffected For instance, the removal of tariffs on footwear products is expected to affect the workers of this sector negatively while at the same time it is also expected to have a positive impact on other households by reducing the cost of their expenditure. Models and empirical evaluations are important to assess the real impact of complex changes on poor households. However, the analyst is facing a problem since trade liberalisation is mainly a macroeconomic phenomenon while poverty and inequality are mainly microeconomic issues. In view of the fact that most of the available economic models have either a microeconomic or a macroeconomic focus,[1] the combination of a micro- and a macro-model appears to be the appropriate approach to address the question adequately, that is without missing the important micro-macro links.

Macro models in general and CGE models in particular are unable to address this issue because they have a necessarily limited number of representative households. It means that these models have to make strong assumptions regarding the intra-group income distribution, that is wether or not this distribution is assumed to be fixed. The high level of aggregation of CGE models means that they are generally not able to identify individual or groups of winners and losers. For this reeason it is not the appropriate tool to evaluate poverty and inequality impacts in a detailed way.

By contrast, MS models allow the modeller to assess the impacts directly at the microeconomic level since they account for the diversity of individual characteristics. These models generally use all the information available from household surveys in order to determine

[1] These models include macro models, microsimulation models, multiplier models and computable general equilibrium models.

precisely how a policy or a shock could affect each household. In our example of the removal of tariffs on footwear products, a MS model would be very useful to inform us on how the changes in prices and in employment levels would affect each household. However such a model is typically a partial equilibrium model and it is thus unable to determine how the new policy would affect employment or prices at the country level. A CGE model would be the appropriate tool to do this.

Obviously the combination of these two types of model is of great interest because it would allow for a combination of the advantages of both models, which in turn should offset their respective drawbacks. This paper presents a way of combining a macro-oriented Computable General Equilibrium (CGE) model and a MS model following the approach introduced by Robilliard et al. (2001). However, this approach is extended by also taking into account the changes in capital returns[2] and in transfers[3] at the household level..Combining these two models the microeconomic effects (on poverty and inequality) of a macroeconomic policy can be predicted.

The second section gives details about the micro-data and the MS model to be developed, focussing on the computation of labour market responses. The section describes the structure of employment and incomes in South Africa. Then the three components of the MS model are presented and discussed: firstly the selection model used to predict occupational choices, secondly the regression model used to predict earnings and thirdly the method applied to compute household real net income. The third section describes both the macro-data and the CGE model. The focus is on the main characteristics of the South African economy and the adjustment mechanisms in the CGE model. The fourth section provides details on the "top-down" approach used to link the MS model and the CGE model. The section presents and discusses the micro-macro consistency equations imposed on the MS model in order to achieve consistency with the CGE model. Finally, section 5 concludes and briefly explores the potential applications of the model. In this regard, Hérault (2005) applies the CGE-MS model described in this paper in order to assess the effects of trade liberalisation in South Africa.

[2] This is similar to what has been done in some more complex CGE-MS models. For instance, see the integrated model by Cororaton and Cockburn (2005).
[3] Transfers include transfers from other households, transfers form abroad and transfers from the government.

2. Microsimulation Modelling

The South African MS model is a static MS model allowing the analysis of social and economic policies' impacts, not only on the means of the variables of interest but also on their distributions. This particular feature derives from one of the main advantages of the MS models: the explicit accounting for individual heterogeneity. MS models can be subdivided into static or dynamic models, where static models are designed mainly for short-term analysis and dynamic models more for medium- to long-term analysis.[4] Given the short term focus of the CGE model, the South African MS model is implemented in a static framework. Therefore, the size and demographic characteristics of the population from the original cross-sectional dataset used in the simulation are fixed.

Moreover, the South African MS model allows for behavioural responses. There is a distinction between behavioural and non-behavioural models. The latter are designed for the analysis of the "morning after" effects of a policy change; that is, individual behaviour is assumed to remain the same as before the change. These models can be useful, for example, when looking at the impact on the government budget of some minor changes in the income tax rates or the eligibility rules for benefits and subsidies. However, once the changes at stake cannot be considered to be of minor importance, the assumption of unchanged behaviour becomes difficult to justify. A more complex behavioural MS model is used in this paper because this type of model allows individuals to adjust their behaviour in response to the simulated policy change.

2.1. Overview of the microeconomic data

The South African MS model is based on two household surveys: the Income and Expenditure Survey (IES) of 2000 and the Labour Force Survey (LFS) of September 2000. The two surveys can be linked since the IES is based on the same sample of households interviewed for the biannual LFS.

The LFS contains detailed employment data on occupation codes, activity codes and wages of workers. It is believed to be more reliable than the IES as far as labour incomes are concerned. The IES provides a detailed picture of income and spending of South African households but the

[4] For a detailed literature review on dynamic microsimulation modelling, see O'Donoghue (2001).

quality of the data is questionable (see PROVIDE project, 2005). Total income and total expenditure differ by more than 30% for almost one quarter of the records and this is only one of the inconsistencies pointed out by Van der Berg et al. (2003). Some adjustments were thus necessary due to inconsistencies within the IES, but also in order to reconcile both surveys. We followed the work done under the PROVIDE project (2005) by the National and Provincial Departments of Agriculture, whereby undeclared income and under-declared expenditure is added to total income and expenditure and all income and expenditure categories are scaled upwards, keeping the patterns of income and expenditure constant. In addition, labour incomes from the LFS were used to scale up labour incomes from the IES when the latter were obviously underreported.[5]

The data are broadly in accordance with Casale et al. (2004) and reveal a national unemployment rate of about 30% depending on the definition used. Table 1 includes the complete South African population and shows remarkable differences between racial groups. The 3.9 million whites account for less than 10% of the total population but for more than 20% of the formal workers. On the contrary, blacks represent almost 80% of the population but less than 60% of the formal workers. Indeed, blacks are overrepresented in all categories except formal workers, whereas coloureds, Asians and whites are underrepresented in all categories except formal workers. Table 1 also reveals that the level of hours worked is fairly high in South Africa averaging around 45 hours per week. It is somewhat lower in the informal sector than in the formal sector.[6] In this regard, it is interesting to note the distinct distribution of hours amongst racial groups in the formal and the informal sectors. Although whites work longer hours than blacks in the informal sector, the opposite is found in the formal sector. This can be explained by the different nature of jobs occupied by blacks and whites.

Table 2 focuses on average incomes, poverty and income distribution. South Africa is well known for being one of the most unequal countries along with Brazil and this is confirmed by a

[5] For this reason, the mean incomes calculated here are slightly higher than in the Provide project (2005); that is, 7.7% higher for total household income and 11.2% for total household labour income. Further adjustments were necessary, in particular to make incomes from labour consistent at the household and at the individual level (Stata do-files are available from the author).

[6] Following Statistics South Africa (questionnaire of the LFS of September 2000), "Formal sector employment is where the employer […] is registered to perform the activity. Informal sector employment is where the employer is not registered."

Gini coefficient as high as 0.67.[7] The recent trend is towards a more unequal income distribution even if inter-group inequality has been declining. The explanation is that the gap between rich and poor within each group has increased substantially despite very high initial levels. For instance, a Gini coefficient of 0.47 for whites is a very high figure for a population with an education and occupation profile similar to those observed in rich countries.

Table 1: Overview of employment (numbers in 1000's)

	Blacks	Coloureds	Asians	Whites	TOTAL
Inactive[(a)]	22,857	2,425	642	2,107	28,032
Unemployed	3,356	282	69	99	3,806
Subsistence agriculture	704	19	1	12	736
Informal workers[(b)]	2,935	268	32	122	3,357
(Average weekly hours of work)	*42.7*	*41.0*	*42.4*	*44.5*	*42.6*
Formal workers	4,327	1,019	359	1,602	7,307
(Average weekly hours of work)	*47.9*	*44.8*	*44.3*	*45.9*	*46.9*
TOTAL	34,180	4,013	1,104	3,941	43,238

Note: (a) Including 652,000 "unspecified" workers (b) Including 948,000 domestic workers. *Source: Author's calculations from IES 2000 and LFS 2000:2*

Table 2: Overview of incomes, poverty and inequality[(a]

	Blacks	Coloureds	Asians	Whites	TOTAL
Income per capita[(b)]	6,268	10,695	19,824	48,495	10,874
Headcount Index (P0)	35.6	10.3	1.1	0.2	29.2
Poverty Gap Index (P1)	14.0	3.2	0.6	0.07	11.4
Poverty Severity Index (P2)	7.3	1.4	0.4	0.03	5.9
Gini	0.59	0.51	0.47	0.47	0.67

Note: (a) The poverty line is the international $2/day poverty line (R174/month/capita in 2000 prices) (b) Average annual disposable income per capita in Rand. *Source: Author's calculations from IES 2000 and LFS 2000:2*

Despite the end of the Apartheid regime in 1994 and the first democratic elections, the white average income per capita was still, in 2000, 8 times higher than the black per capita income. As a consequence, it is not surprising to observe that poverty concerns almost only blacks; they account for more than 95% of all poor and more than 35% of all blacks are poor

[7] This Gini coefficient is computed on per capita disposable incomes. Using household incomes, the Gini coefficient is 0.59. Similar values are found in other studies. Simkins (2004) calculates a Gini coefficient on household incomes of 0.67 in 2000 and 0.69 in 2001. The Gini estimate by HRSC (2004) is 0.77 (using household incomes) for 2000. The World Bank (see http://iresearch.worldbank.org/PovcalNet), calculates a Gini coefficient, based on per capita incomes, of 0.58 in 2000 and the Gini index computed by Hoogeveen and Özler (2004) is also 0.58.

using the international $2/day poverty line. In contrast, poverty is virtually non-existent for Asians and whites despite the relatively high incidence of poverty at the national level of 29.2%.[8]

2.2. Key characteristics of the model

The model simulates the new labour market choices after changes in individual characteristics, such as earnings (due to macroeconomic changes as estimated in a CGE model), or in the coefficients of the model. The simulation is carried out for all individuals aged between 15 and 65 years. Incomes are simulated for each of the 26,000 households surveyed in the 2000 IES and LFS.[9] The underlying selection model, which drives the behavioural responses, assigns each individual from the working-age population to one of the five labour market categories distinguished in the model: inactive, unemployed, subsistence agricultural worker, informal worker and formal worker. This model takes the potential earnings in these categories into account. A regression model is estimated to predict earnings in each category. Finally, the results of both the selection and the regression model are used to compute household real net incomes.

2.2.1. The selection model for labour market choices

A multinomial logit specification is used for the selection model (see Maddala, 1983). The model assigns each individual to the sector with the highest associated probability. The probabilities are derived from the estimation of an implicit utility function. Therefore, the underlying assumption is that each individual chooses the sector with the highest associated utility. Equation (1) expresses the utility associated with each of the five labour market choices *j* for an individual *i*:

[8] Using the same poverty line, the World Bank (http://iresearch.worldbank.org/PovcalNet) produces an estimate of 34.1% while Hoogeveen and Özler (2004) estimated a poverty incidence of 34%. Using a poverty line of R250 per month per capita, our estimated poverty incidence of 41.7% is close to the 38.6% found by Van der Berg and Louw (2003).

[9] A few workers aged below 15 years or over 65 years are fixed at the observed labour market category. Likewise, 652,000 "unspecified" workers, for whom there is no data on occupation and on the sector (formal or informal), are also excluded from the model. That these occupation and sector variables are missing is partly due to the fact that, for various reasons, some people were observed as workers in the IES but not in the LFS. These workers are also excluded from the simulation, and their labour market choices are fixed at the observed levels. Moreover, all recipients of the Old Age and War pension, the Child Support Grant and the Disability Grant, as well as full-time students aged between 15 and 18 years, are excluded from the simulation.

$$U_{ij} = a_{h(i),j} + Z_{ij}.B_{h(i),j} + u_{ij} \quad (1)$$

This utility function is defined separately for the four demographic groups considered in the model: single women, married women, single men and married men.[10] The demographic group to which individual *i* belongs is indicated by an index function *h(i)*. The utility associated with each choice is a linear function of a set of individual characteristics Z_{ij}, which includes predicted net earnings,[11] skill level,[12] age, education, province of residence, race as well as the number and age of children if applicable. The utility associated with the first alternative (inactive) is set to zero to identify the model. The residual term u_{ij} describes the effects of unobserved determinants on labour market choices. The intercept $a_{h(i),j}$ is common to all individuals belonging to the same demographic group *h(i)*.

Taking the example of the single women (*sw*), equation (2) provides an expanded form of $Z_{ij}.B_{h(i),j}$ showing some variables of particular interest:

$$Z_{ij}.B_{sw,j} = SK_i.B1_{sw,j} + HS_i.B2_{sw,j} + PNE_{ij}.B3_{sw} + \ldots \quad (2)$$

The dummy variables SK_i and HS_i stand for skilled and high-skilled respectively, and indicate the skill level of individual *i*. The reference category is low-skilled workers. The coefficients $B_{h(i),j}$ of individual characteristics Z_{ij} are specific to each demographic group *h(i)* and each labour market choice *j*. The only exception is the coefficient $B3_{h(i)}$ associated with the category-specific predicted net earnings PNE_{ij}. That is, the coefficient differs between demographic groups but it is the same for all labour market choices (i.e. the same coefficient $B3_{sw}$ is used for $PNE_{i,0}$, $PNE_{i,1}$…, $PNE_{i,4}$). The variable PNE_{ij} is different from the other variables, because it is not constant across the choices. Each choice is associated with a particular individual-specific level of earnings. The constant coefficient across labour market choices

[10] The classification by marital status and gender was chosen for two main reasons. First because labor market choices and earnings are significantly different between the four demographic groups. Second, because the four groups are of a reasonable and comparable size. A classification by racial groups was also considered. However, given the very important differences in the sizes of the racial groups (see Table 1), such a classification would have made the estimation of certain modules of the model difficult or impossible.

[11] A regression model is used to predict earnings for each individual for each category. More details are provided in the next section.

[12] The skill level is derived from the occupation for those observed working at the time of the survey and from the former occupation, when data is available, for the inactive and unemployed. An ordered probit model is estimated to predict the skill level when the relevant information is not available in the survey.

implies that for a given demographic group, the influence of predicted net earnings is the same over all labour market choices. The underlying assumption is that individuals, when evaluating the utility associated with each of the five choices, take into account the predicted net earnings in the same way. This model combines two types of specification: multinomial and conditional logit (see Maddala, 1983).

The estimated relative risk ratios (RRR) of the selection models for the four demographic groups are presented in tables 3 to 6. All RRRs are measured relative to the probability of the base category, which is the inactive group in our model. A RRR of less than one implies that the variable increases the relative probability of being inactive while a RRR greater than 1 implies that the variable reduces the relative probability of being inactive versus the labour market category considered. Some RRRs are not available in the subsistence agricultural equations because the corresponding coefficients had to be fixed to zero due to the lack of observations. For instance, there are no Asian single women in subsistence agriculture in the dataset.

The RRR of 1.27 associated with coloured in the unemployed equation for single women (see table 3) means that being coloured raises the relative probability of being unemployed versus being inactive by 27 per cent when all other independent variables are held constant. In all four tables, we note some large impacts for semi-skilled and agricultural activity variables on the RRRs for subsistence agriculture. This is caused by the fact that most of the subsistence agricultural workers are semi-skilled and have farming activities on top of their main occupation.

The results from the informal sector component of the multinomial-logit estimation show that as educational qualifications increase, the likelihood of being employed informally, rather than formally, declines. Having a young child has a significant impact on the labour market choices of singles by making them less likely to work in either the formal or the informal sector. The impacts on married people are less clear and often not significant. Unlike women, educated men are not significantly more likely to be in the formal sector than uneducated men but being educated reduces their odds of being informal worker. Moreover, having a high-skilled wife has a negative effect on their participation in both the formal and the formal sector.

Regardless of the demographic group, living in the North West province and in a former homeland makes people less likely to be a formal or informal worker, and more likely to be either inactive or in the subsistence agriculture.

Finally, RRRs greater than one show that as the predicted income increases, the likelihood of being in the corresponding labour market improves as well.

Table 3: Estimated relative risk ratios of the selection models for single women

Sample size: 15,137 obs.	**Unemployed**		**Subsistence agriculture**		**Informal sector**		**Formal sector**	
	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level
Coloured	1.270	0.360	0.602	0.530	0.627	0.071	2.726	0.000
Asian	0.801	0.565	***	.	0.183	0.001	1.644	0.144
White	0.719	0.324	***	.	0.070	0.000	1.747	0.056
Primary education	0.932	0.664	1.674	0.212	1.071	0.629	1.104	0.554
Standard 8	1.014	0.931	0.968	0.940	0.844	0.258	1.460	0.028
Standard 10	1.536	0.010	0.343	0.013	0.721	0.040	1.822	0.001
Tertiary education	1.205	0.519	***	.	0.002	0.000	1.500	0.131
Education attendance	0.034	0.000	0.263	0.000	0.123	0.000	0.107	0.000
Semi-skilled	1.179	0.012	32.333	0.000	0.949	0.494	1.824	0.000
High-skilled	0.453	0.061	***	.	0.547	0.174	3.206	0.000
Age	1.319	0.000	1.023	0.629	1.470	0.000	1.502	0.000
Age squared/100	0.656	0.000	0.939	0.326	0.630	0.000	0.603	0.000
Ability to write	1.551	0.006	0.691	0.381	1.415	0.013	1.194	0.276
Afrikaans language	0.536	0.017	1.040	0.962	1.239	0.404	0.752	0.274
English language	0.432	0.005	***	.	0.800	0.441	0.516	0.016
Agri. Activity	1.245	0.017	44.854	0.000	2.931	0.000	1.212	0.063
Western Cape	1.340	0.027	4.987	0.042	2.127	0.000	2.336	0.000
Eastern Cape	0.860	0.130	0.682	0.495	1.085	0.476	0.831	0.098
Northern Cape	0.842	0.256	2.502	0.324	1.348	0.080	0.813	0.218
Freestate	0.976	0.829	1.859	0.275	0.856	0.235	1.559	0.000
Kwazulu-Natal	1.035	0.697	1.827	0.263	1.284	0.015	1.419	0.000
North-West	0.823	0.051	0.171	0.047	0.854	0.178	0.777	0.026
Mpumalanga	1.272	0.026	1.215	0.733	1.432	0.004	1.596	0.000
Limpopo	0.788	0.043	0.243	0.016	0.770	0.050	1.064	0.631
Urban location	1.912	0.000	0.698	0.181	1.322	0.000	1.075	0.315
Homeland location	0.814	0.004	1.562	0.031	0.719	0.000	0.654	0.000
Children (number)	0.920	0.001	0.942	0.356	1.207	0.000	1.089	0.005
Household size	1.048	0.000	1.022	0.587	0.846	0.000	0.904	0.000
Children 0-1 yr old	0.743	0.005	1.282	0.454	0.453	0.000	0.479	0.000
Children 1-3 yrs old	1.094	0.312	1.212	0.504	0.588	0.000	0.633	0.000
Children 4-5 yrs old	0.795	0.023	0.974	0.932	0.574	0.000	0.650	0.000
Children 6-9 yrs old	0.887	0.195	1.450	0.197	0.597	0.000	0.675	0.000
Children 9-15 yrs old	0.991	0.927	0.667	0.230	0.694	0.000	0.767	0.008
Disability grant in the hh	1.194	0.076	0.496	0.036	0.577	0.001	0.764	0.043
OAP**** in the hh	0.920	0.157	1.223	0.218	0.485	0.000	0.611	0.000
Family allow. in the hh	1.335	0.006	0.541	0.084	1.094	0.499	0.927	0.567
Capital income**	0.998	0.578	0.930	0.126	1.003	0.243	0.991	0.074
Gov. transfers to the hh**	1.009	0.085	0.998	0.852	0.930	0.002	0.981	0.124
Predicted inactive inc.*	1.00005	0.000	1.00005	0.000	***	.	***	.
Predic. informal earnings*	***	.	***	.	1.00005	0.000	***	.
Predic. formal earnings*	***	.	***	.	***	.	1.00005	0.000

Note: *in Rand per year **in 1000's Rand per year ***coefficient fixed to zero **** Old Age and War Pension

Table 4: Estimated relative risk ratios of the selection models for married women

Sample size: 11,378 obs.	Unemployed		Subsistence agriculture		Informal sector		Formal sector	
	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level
Coloured	0.539	0.022	2.982	0.365	0.478	0.001	1.151	0.514
Asian	0.237	0.000	***	.	0.099	0.000	0.488	0.004
White	0.282	0.000	1.705	0.669	0.074	0.000	0.400	0.000
Primary education	1.120	0.530	5.168	0.000	1.188	0.202	0.926	0.657
Standard 8	1.256	0.226	3.071	0.013	0.931	0.626	1.093	0.617
Standard 10	1.423	0.077	0.577	0.271	0.832	0.269	1.425	0.058
Tertiary education	1.225	0.587	1.174	0.892	0.186	0.000	1.687	0.043
Education attendance	0.246	0.000	0.647	0.359	0.491	0.001	0.766	0.065
Semi-skilled	1.113	0.210	114.405	0.000	1.173	0.041	1.708	0.000
High-skilled	1.619	0.113	***	.	2.015	0.013	2.391	0.000
Age	1.178	0.000	1.067	0.291	1.197	0.000	1.317	0.000
Age squared/100	0.767	0.000	0.909	0.208	0.813	0.000	0.712	0.000
Ability to write	1.001	0.996	0.189	0.000	0.903	0.445	1.204	0.292
Afrikaans language	0.669	0.131	0.075	0.035	1.247	0.320	1.192	0.411
English language	0.569	0.060	0.075	0.031	0.982	0.945	0.835	0.418
Agri. Activity	0.739	0.005	15.435	0.000	2.096	0.000	0.861	0.128
Western Cape	0.627	0.004	3.431	0.081	1.015	0.918	1.369	0.008
Eastern Cape	0.621	0.000	1.555	0.403	0.698	0.004	1.016	0.889
Northern Cape	0.659	0.025	1.095	0.923	1.027	0.870	0.796	0.117
Freestate	0.869	0.259	1.840	0.258	0.979	0.862	1.319	0.021
Kwazulu-Natal	0.700	0.003	1.852	0.227	0.813	0.072	1.187	0.107
North-West	0.670	0.001	0.823	0.741	0.617	0.000	0.824	0.097
Mpumalanga	1.107	0.418	1.985	0.191	0.964	0.772	1.069	0.599
Limpopo	0.718	0.016	0.472	0.168	0.531	0.000	1.046	0.726
Urban location	1.856	0.000	0.317	0.000	0.944	0.421	0.914	0.212
Homeland location	0.856	0.101	0.646	0.047	0.716	0.000	0.566	0.000
Children (number)	0.859	0.000	0.860	0.077	0.930	0.037	0.827	0.000
Household size	1.082	0.000	0.978	0.675	1.013	0.508	1.020	0.341
Children 0-1 yr old	0.666	0.003	1.219	0.574	0.566	0.000	0.926	0.548
Children 1-3 yrs old	0.962	0.730	0.975	0.933	0.712	0.001	1.011	0.916
Children 4-5 yrs old	0.862	0.239	0.891	0.721	0.717	0.004	0.959	0.716
Children 6-9 yrs old	1.057	0.631	1.470	0.197	0.888	0.244	1.275	0.018
Children 9-15 yrs old	1.055	0.662	0.851	0.605	0.855	0.137	1.025	0.807
OAP**** in the hh	1.085	0.477	0.746	0.229	0.628	0.000	0.876	0.230
Family allow. in the hh	1.272	0.218	1.756	0.131	1.250	0.190	0.917	0.688
Capital income**	0.986	0.029	0.845	0.001	0.998	0.024	0.995	0.002
Gov. transfers to the hh**	0.970	0.049	1.008	0.811	0.980	0.172	0.937	0.000
Partner live here	2.316	0.002	2.442	0.152	1.649	0.036	3.288	0.000
Primary education (partner)	1.215	0.118	0.985	0.951	0.986	0.875	1.028	0.815
Standard 8 (partner)	1.158	0.268	0.787	0.424	0.788	0.025	1.000	0.999
Standard 10 (partner)	1.204	0.205	1.190	0.647	0.560	0.000	0.938	0.630
Tertiary education (partner)	0.920	0.760	3.657	0.113	0.415	0.006	0.883	0.489
Age (partner)	0.981	0.000	0.987	0.254	0.986	0.002	0.974	0.000
Semi-skilled (partner)	0.798	0.220	0.464	0.242	0.969	0.865	1.184	0.134
High-skilled (partner)	1.049	0.553	1.290	0.185	1.359	0.000	1.064	0.382
Predicted inactive inc.*	1.00004	0.000	1.00004	0.000	***	.	***	.
Predic. informal earnings*	***	.	***	.	1.00004	0.000	***	.
Predic. formal earnings*	***	.	***	.	***	.	1.00004	0.000

Note: *in Rand per year **in 1000's Rand per year ***fixed to zero **** Old Age and War Pension

Table 5: Estimated relative risk ratios of the selection models for single men

Sample size: 13,252 obs.	Unemployed		Subsistence agriculture		Informal sector		Formal sector	
	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level
Coloured	0.678	0.144	0.376	0.127	0.768	0.421	1.134	0.615
Asian	0.838	0.644	0.639	0.667	0.351	0.095	0.944	0.867
White	0.327	0.000	0.429	0.247	0.102	0.000	0.454	0.005
Primary education	0.886	0.460	1.186	0.669	1.030	0.876	0.880	0.452
Standard 8	1.168	0.356	0.833	0.669	0.977	0.908	0.824	0.276
Standard 10	1.420	0.041	0.139	0.000	0.558	0.005	1.020	0.912
Tertiary education	1.988	0.012	0.175	0.039	0.201	0.000	0.889	0.664
Education attendance	0.020	0.000	0.178	0.000	0.077	0.000	0.041	0.000
Semi-skilled	1.166	0.037	48.662	0.000	3.511	0.000	1.813	0.000
High-skilled	1.051	0.875	***	.	3.859	0.000	2.247	0.002
Age	1.241	0.000	1.039	0.317	1.260	0.000	1.342	0.000
Age squared/100	0.723	0.000	0.946	0.276	0.751	0.000	0.673	0.000
Ability to write	1.882	0.000	0.928	0.839	1.217	0.285	1.430	0.029
Afrikaans language	1.222	0.450	3.913	0.029	2.415	0.007	1.674	0.039
English language	0.925	0.800	***	.	0.595	0.220	1.374	0.243
Agri. Activity	1.108	0.323	12.751	0.000	3.005	0.000	0.653	0.001
Western Cape	1.034	0.811	2.262	0.120	1.153	0.435	1.947	0.000
Eastern Cape	0.907	0.352	1.299	0.504	0.839	0.229	0.807	0.066
Northern Cape	1.251	0.175	4.024	0.009	0.922	0.714	1.513	0.012
Freestate	1.195	0.131	2.848	0.008	1.018	0.914	0.871	0.296
Kwazulu-Natal	1.060	0.541	1.779	0.125	0.867	0.282	1.071	0.503
North-West	0.882	0.230	1.201	0.663	0.686	0.009	0.678	0.001
Mpumalanga	1.521	0.000	1.656	0.226	1.642	0.002	1.693	0.000
Limpopo	0.874	0.275	0.759	0.520	0.666	0.015	0.977	0.869
Urban location	1.399	0.000	0.995	0.982	0.862	0.104	0.605	0.000
Homeland location	0.736	0.000	1.269	0.228	0.818	0.048	0.315	0.000
Household size	0.979	0.253	0.988	0.829	0.745	0.000	0.782	0.000
Number of male	1.046	0.067	0.971	0.656	1.056	0.173	1.071	0.026
Children (number)	0.975	0.379	1.027	0.732	1.292	0.000	1.199	0.000
Children 0-1 yr old	1.019	0.888	1.440	0.270	0.987	0.946	0.901	0.498
Children 1-3 yrs old	1.019	0.850	1.264	0.382	0.682	0.011	0.598	0.000
Children 4-5 yrs old	0.950	0.642	1.172	0.580	0.676	0.015	0.597	0.000
Children 6-9 yrs old	0.823	0.036	1.042	0.865	0.533	0.000	0.461	0.000
Children 9-15 yrs old	0.909	0.298	0.578	0.036	0.544	0.000	0.611	0.000
Disability grant in the hh	1.190	0.124	1.723	0.067	1.560	0.017	0.899	0.472
OAP**** in the hh	0.900	0.104	1.492	0.051	0.849	0.197	0.631	0.000
Family allow. in the hh	1.191	0.160	1.508	0.216	1.386	0.085	0.792	0.168
Capital income**	0.993	0.198	0.976	0.390	1.000	0.783	0.987	0.003
Gov. transfers to the hh**	1.008	0.521	0.901	0.155	0.777	0.000	0.868	0.000
Predicted inactive inc.*	1.00005	0.000	1.00005	0.000	***	.	***	.
Predic. informal earnings*	***	.	***	.	1.00005	0.000	***	.
Predic. formal earnings*	***	.	***	.	***	.	1.00005	0.000

Note: *in Rand per year **in 1000's Rand per year ***fixed to zero **** Old Age and War Pension

Table 6: Estimated relative risk ratios of the selection models for married men

Sample size: 10,757 obs.	Unemployed		Subsistence agriculture		Informal sector		Formal sector	
	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level
Coloured	1.284	0.565	4.844	0.230	1.369	0.390	1.602	0.142
Asian	1.432	0.517	***	.	3.197	0.018	1.833	0.133
White	0.352	0.039	0.610	0.748	0.327	0.006	0.798	0.499
Primary education	1.390	0.152	0.915	0.818	0.915	0.627	0.914	0.592
Standard 8	1.492	0.106	0.576	0.212	0.736	0.127	0.918	0.641
Standard 10	1.525	0.126	0.458	0.090	0.521	0.005	1.098	0.651
Tertiary education	1.232	0.643	***	.	0.179	0.000	0.783	0.401
Education attendance	0.487	0.113	***	.	0.182	0.000	1.037	0.899
Semi-skilled	1.087	0.527	112.721	0.000	2.544	0.000	1.568	0.000
High-skilled	0.792	0.424	***	.	0.929	0.753	1.520	0.024
Age	1.362	0.000	1.159	0.043	1.178	0.000	1.311	0.000
Age squared/100	0.674	0.000	0.810	0.008	0.803	0.000	0.711	0.000
Ability to write	1.014	0.953	0.729	0.423	1.054	0.777	1.229	0.227
Afrikaans language	0.473	0.082	0.308	0.385	0.764	0.453	0.792	0.456
English language	0.667	0.404	0.178	0.234	0.225	0.000	0.788	0.487
Agri. Activity	0.783	0.119	7.806	0.000	1.669	0.000	0.736	0.008
Western Cape	0.667	0.118	1.517	0.551	0.877	0.554	1.153	0.461
Eastern Cape	0.668	0.047	0.716	0.530	0.675	0.027	0.685	0.017
Northern Cape	1.125	0.697	0.214	0.187	0.850	0.545	1.186	0.471
Freestate	0.962	0.861	2.122	0.150	0.829	0.359	1.539	0.017
Kwazulu-Natal	0.790	0.214	0.711	0.509	0.619	0.005	0.805	0.153
North-West	0.886	0.569	0.721	0.567	0.674	0.039	1.065	0.711
Mpumalanga	0.992	0.973	0.715	0.567	1.031	0.880	1.684	0.004
Limpopo	0.785	0.287	0.578	0.318	0.766	0.171	1.042	0.817
Urban location	1.143	0.318	0.430	0.004	0.654	0.000	0.564	0.000
Homeland location	0.592	0.000	0.662	0.149	0.594	0.000	0.282	0.000
Household size	1.009	0.828	0.859	0.074	0.862	0.000	0.840	0.000
Number of male	1.037	0.503	1.077	0.488	1.058	0.237	1.062	0.161
Children (number)	0.879	0.030	0.928	0.513	1.041	0.452	1.060	0.218
Children 0-1 yr old	1.749	0.019	2.720	0.038	1.356	0.154	1.530	0.028
Children 1-3 yrs old	1.456	0.049	2.056	0.068	1.182	0.319	1.178	0.280
Children 4-5 yrs old	1.430	0.095	2.611	0.023	1.159	0.432	1.218	0.243
Children 6-9 yrs old	1.237	0.260	1.880	0.096	1.060	0.722	1.053	0.721
Children 9-15 yrs old	1.234	0.278	1.982	0.069	1.093	0.589	1.063	0.674
OAP**** in the hh	0.935	0.671	1.211	0.533	0.569	0.000	0.534	0.000
Family allow. in the hh	1.138	0.604	1.842	0.147	0.844	0.462	0.506	0.002
Capital income**	1.000	0.726	0.999	0.523	0.999	0.169	0.995	0.000
Gov. transfers to the hh**	0.999	0.936	0.971	0.578	0.876	0.000	0.887	0.000
Partner live here	4.412	0.001	1.239	0.787	4.220	0.000	2.531	0.005
Primary education (partner)	1.230	0.274	0.989	0.972	0.971	0.844	1.136	0.356
Standard 8 (partner)	1.021	0.919	0.701	0.330	0.757	0.095	0.889	0.447
Standard 10 (partner)	1.098	0.707	0.461	0.104	0.747	0.160	1.106	0.592
Tertiary education (partner)	1.361	0.588	0.692	0.814	1.329	0.535	2.118	0.049
Age (partner)	0.967	0.000	0.985	0.350	0.970	0.000	0.967	0.000
Semi-skilled (partner)	0.603	0.232	0.402	0.552	0.552	0.092	0.453	0.005
High-skilled (partner)	0.770	0.066	1.150	0.552	0.939	0.602	0.985	0.890
Predicted inactive inc.*	1.00002	0.000	1.00002	0.000	***	.	***	.
Predic. informal earnings*	***	.	***	.	1.00002	0.000	***	.
Predic. formal earnings*	***	.	***	.	***	.	1.00002	0.000

Note: *in Rand per year **in 1000's Rand per year

An additional feature has been added, which makes the model probabilistic. Following Creedy et al. (2002), this means that the model "does not identify a particular (…) [labour market choice] for each individual after the policy change, but generates a probability distribution over the (…) [labour market choices] used". This is achieved by drawing, for each individual, a set of error terms u_{ij}, representing the random utility, from the extreme value distribution (here 100 error terms are drawn). These error terms are drawn in such a way that only those which preserve the observed labour market choice as the optimal choice when adding the error term to the deterministic part of the utility function (U_{ij} with u_{ij} left out) are selected. After a macro-change, only the deterministic part of the utility function is recomputed. Then, by adding the random error terms, previously drawn, to the recomputed deterministic utility components, a probability distribution over the labour market choices is generated for each individual. This implies that the model does not assign every individual from the sample to one particular labour market choice after the policy change. Instead, it gives the individual probabilities of choosing each of the five sectors. Combining this information with the sampling weights from the household survey allows us to generate the new distribution of the South African population over the labour market choices incorporated in the model.

2.2.2. The regression model for earnings

The regression model is used to predict individual gross earnings in each of the five labour market categories. This variable influences the labour market choices made in the selection model. Furthermore, predicted earnings are used at the end of the microsimulation process when household real net incomes are calculated. Equation (3) below expresses the (log) predicted gross earnings $PGE_{i,j}$ of individual i in labour market j as a linear function of individual characteristics X_{ij}, the inverse Mills ratio imr_{ij} and unobserved earnings determinants v_{ij}:

$$PGE_{ij} = \alpha_{h(i),j} + X_{ij}.\beta_{h(i),j} + \delta_{h(i),j}.imr_{ij} + v_{ij} \qquad (3)$$

The set of individual characteristics X_{ij} essentially includes age, education, province of residence, race, skill level, language and ability to write. The inverse Mills ratios are estimated using a benchmark run of the selection model, where earnings are excluded from the explanatory

variables. The inverse Mills ratios control for the selection bias of observing an individual's earnings (Maddala, 1983).

The regression model, like the selection model, is estimated separately for the four demographic groups *h(i)*.[13] Since formal and informal sectors are the only labour market segments where individuals receive earnings, the results of the regression model are used to assign predicted formal and informal earnings to all individuals in the model. However, for those individuals who were formal or informal workers at the time of the survey, observed earnings are used rather than predicted earnings. The estimated parameters of the regression model are presented in table 7 for women and table 8 for men.

Table 7: Estimated parameters of the regression models for women (log earnings)

	Single women				**Married women**			
	Informal earnings		**Formal earnings**		**Informal earnings**		**Formal earnings**	
(Sample size)	(2503 obs.)		(3065 obs.)		(2124 obs.)		(2953 obs.)	
	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level
Coloured	0.281	0.126	-0.022	0.872	0.504	0.013	-0.008	0.942
Asian	1.244	0.003	0.028	0.870	1.736	0.000	0.006	0.957
White	1.724	0.000	0.335	0.021	2.184	0.000	0.347	0.001
Primary education	0.016	0.870	-0.097	0.321	-0.008	0.948	-0.282	0.003
Standard 8	0.253	0.020	0.077	0.444	0.521	0.000	-0.017	0.861
Standard 10	0.631	0.000	0.442	0.000	0.725	0.000	0.369	0.000
Tertiary education	2.404	0.000	0.728	0.000	2.017	0.000	0.545	0.000
Education attendance	-0.014	0.925	0.441	0.000	0.488	0.043	0.335	0.000
Semi-skilled	-0.077	0.221	0.210	0.000	-0.263	0.000	0.285	0.000
High-skilled	0.918	0.001	0.562	0.000	0.102	0.652	0.576	0.000
Age	0.029	0.080	0.064	0.000	0.046	0.058	0.044	0.005
Age squared/100	-0.033	0.117	-0.051	0.001	-0.056	0.058	-0.033	0.092
Ability to write	0.052	0.602	0.282	0.003	0.257	0.028	0.607	0.000
Afrikaans language	-0.153	0.397	0.092	0.504	-0.187	0.342	0.045	0.672
English language	0.055	0.761	0.295	0.035	-0.232	0.358	0.169	0.111
Western Cape	0.028	0.793	-0.122	0.053	0.048	0.727	-0.090	0.117
Eastern Cape	-0.481	0.000	-0.113	0.056	-0.341	0.004	-0.171	0.003
Northern Cape	-0.478	0.000	-0.131	0.115	-0.495	0.001	-0.113	0.117
Freestate	-0.529	0.000	-0.418	0.000	-0.668	0.000	-0.377	0.000
Kwazulu-Natal	-0.237	0.002	-0.239	0.000	-0.099	0.351	-0.098	0.061
North-West	-0.145	0.117	-0.056	0.361	0.088	0.460	-0.021	0.726
Mpumalanga	-0.275	0.003	-0.206	0.002	-0.018	0.871	-0.081	0.223
Limpopo	0.104	0.307	-0.225	0.001	0.231	0.066	-0.168	0.010
Urban location	0.160	0.004	0.292	0.000	0.320	0.000	0.386	0.000
Homeland location	-0.071	0.267	0.169	0.000	0.088	0.274	0.251	0.000
Agri. Activity	-1.078	0.000	-0.225	0.000	-0.991	0.000	0.049	0.407
Inverse Mills ratios	-0.811	0.000	-0.404	0.000	-1.183	0.000	-0.368	0.000
Constant	8.939	0.000	7.842	0.000	8.705	0.000	7.963	0.000

[13] Chow tests confirm that the coefficients are significantly different in the four demographic groups.

Table 8: Estimated parameters of the regression models for men (log earnings)

	Single men				Married men			
	Informal earnings		Formal earnings		Informal earnings		Formal earnings	
(Sample size)	(1332 obs.)		(3453 obs.)		(1710 obs.)		(7139 obs.)	
	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level
Coloured	0.193	0.492	0.064	0.542	0.324	0.101	0.015	0.835
Asian	0.821	0.147	0.265	0.046	0.154	0.641	0.091	0.265
White	1.523	0.000	0.710	0.000	1.173	0.000	0.720	0.000
Primary education	-0.252	0.119	-0.006	0.946	-0.018	0.866	0.079	0.079
Standard 8	-0.176	0.310	0.206	0.016	0.313	0.012	0.299	0.000
Standard 10	0.305	0.094	0.460	0.000	0.672	0.000	0.572	0.000
Tertiary education	0.717	0.119	0.682	0.000	0.925	0.002	0.822	0.000
Education attendance	0.037	0.826	0.307	0.000	0.656	0.107	0.300	0.000
Semi-skilled	-0.750	0.000	0.133	0.000	-0.340	0.000	0.166	0.000
High-skilled	0.102	0.687	0.687	0.000	0.924	0.000	0.462	0.000
Age	0.048	0.009	0.083	0.000	0.053	0.015	0.089	0.000
Age squared/100	-0.062	0.010	-0.085	0.000	-0.059	0.017	-0.087	0.000
Ability to write	0.474	0.003	0.221	0.005	0.237	0.036	0.203	0.000
Afrikaans language	-0.153	0.588	0.088	0.395	-0.071	0.725	-0.015	0.824
English language	0.512	0.131	0.098	0.372	0.405	0.131	0.169	0.020
Western Cape	0.204	0.244	-0.162	0.003	-0.061	0.675	-0.046	0.231
Eastern Cape	-0.526	0.000	-0.169	0.002	-0.485	0.000	-0.160	0.000
Northern Cape	0.049	0.812	-0.360	0.000	-0.303	0.075	-0.054	0.238
Freestate	-0.516	0.001	-0.226	0.000	-0.538	0.000	-0.135	0.000
Kwazulu-Natal	0.195	0.102	-0.091	0.043	-0.140	0.223	-0.082	0.016
North-West	0.200	0.144	0.153	0.003	-0.017	0.895	0.087	0.016
Mpumalanga	-0.102	0.467	-0.085	0.140	-0.047	0.703	-0.035	0.360
Limpopo	0.351	0.021	-0.273	0.000	-0.085	0.502	-0.033	0.444
Urban location	0.297	0.000	0.387	0.000	0.240	0.001	0.331	0.000
Homeland location	-0.349	0.000	0.192	0.000	-0.089	0.321	0.206	0.000
Agri. Activity	-1.166	0.000	0.083	0.263	-0.548	0.000	-0.067	0.060
Inverse Mills ratios	-0.787	0.000	-0.271	0.000	-0.262	0.208	-0.541	0.000
Constant	9.024	0.000	7.482	0.000	8.145	0.000	7.493	0.000

Regardless of the sector and the demographic group, being white (or to a lesser extent Asian) is always associated with higher earnings. Age has a positive effect while age squared has a negative impact on earnings, indicating that the effect of age is non-linear. The results show that as the educational qualifications increase, formal and informal earnings go up. Although not all coefficients are significant, living in Gauteng (the reference group) and in an urban location generally has a positive impact on earnings. However, the general finding is that coefficients, and especially their magnitudes, vary substantially across sectors and demographic groups.

2.2.3. Calculation of household income

After running the selection and the regression models described above, individual earnings are added to other non-labour income to generate the updated household incomes Y_h. Equation (4) describes how the real net income of household h is computed:

$$Y_h = [\sum_{i \epsilon h} PGE_{i,FS}.FW_i + PGE_{i,IS}.IW_i + y_h - taxes_h]/CPI_h \quad (4)$$

FW_i is a dummy variable that equals one if individual i is found to be a formal worker and zero otherwise. Likewise, IW_i indicates if individual i is predicted to be an informal worker. Therefore, the earnings in the formal sector $PGE_{i,FS}$ are summed only over household members actually engaged as formal workers. Correspondingly, informal earnings $PGE_{i,IS}$ are added to the household income only if individual i is found to be an informal worker. The model also accounts for non-labour incomes of the household, y_h. The majority of non-labour income consists of incomes from capital, and transfer incomes from other South African households, from abroad and from the government.

The sum of labour and non-labour incomes is the gross household income from which we subtract the income tax payable ($taxes_h$). However, only earnings from the formal sector are considered as being taxable. Income taxes are computed at the individual level by applying the official tax rates from the South African Revenue Service.[14] Then, household net income is deflated by a household-specific consumer price index CPI_h. Two sources of data are used to compute the household *CPI*: the budget shares of household h are observed in the micro-data while the price changes are derived from the CGE model. As a result, the household real net income Y_h takes into account household-specific expenditure patterns.[15] This particular characteristic of the model is of interest when the focus is on poverty and inequality impacts of policy changes because expenditure patterns are significantly different for low-income and high-

[14] Observed income taxes from micro- and macro-data are not used for various reasons. First and foremost because tax and income data do not refer to the same period. Moreover, there is no consistency between micro and macro tax data. Therefore, results, as well as base values, are computed using official tax rates rather than those derived from observed taxes. The model is reasonably robust to the use of observed tax rates instead of official tax rates. The differences are driven by the evident underestimation of tax rates and the absence of progressivity when using observed taxe rates. As a consequence, the use of observed taxes places a light downward pressure on poverty measures while inequality is scaled upward.

[15] However, the model does not allow for a change in consumption patterns due to relative price changes.

income households. Moreover, the price changes derived from the CGE model can vary substantially amongst the 43 goods included in the model.

Finally, since MS models usually are partial equilibrium models, focussing on the household side of the economy. They miss out on a significant part of the story: the general equilibrium effects. When dealing with substantial policy changes, it is essential to take into account the macroeconomic effects of these changes since they are likely to influence the microeconomic outcomes.

3. CGE modelling

The CGE model used in this paper was developed by Hérault (2004) based on the 2000 South African social accounting matrix (SAM). It is a static model similar in many aspects to the model presented in Thurlow and van Seventer (2002)[16] which can be classified as a neoclassical-structuralist model in the tradition originally introduced in Dervis et al. (1982). The main innovation introduced by CGE models is the explicit consideration of the general equilibrium effects. This built-in characteristic explains why CGE models have been so intensively used in the last decade, especially for the assessment of trade liberalisation policies in developing countries. As a matter of fact, trade liberalisation cannot be satisfactorily evaluated under a partial equilibrium framework. The liberalisation of trade in a given sector is more than likely to have some effects beyond this single sector. The most straightforward examples may occur if the output of this sector is used as an intermediate input in other sectors or if the sector of interest is one of the main buyers of any domestic product.

3.1. Background

The South African CGE model shares some characteristics common to most of the CGE models. Producers maximise profits subject to their existing production technology and households maximise utility subject to a budget constraint, but the government does not have an

[16] The author is very grateful to James Thurlow and the International Food Policy Research Institute for their technical support during (and after) the advanced CGE workshop at the University of Cape Town in January 2004.

objective function. The latter is one of the reasons why more constraints are necessary in order to "close" the model and define how the economic system operates. In practice, the aim is to have the same number of equations and endogenous variables in the model. The required additional constraints are often referred to as the closure rules or system constraints and are of particular interest because they determine, to a large extent, the way in which the model behaves. They ensure equilibrium is reached by answering such questions as: is the budget deficit fixed or flexible? Is the exchange rate or the current account fixed? These closure rules also enforce the *ex post* equality of investments and savings.[17]

The first CGE models were mainly applied to developed countries and were directly derived from the neoclassical general equilibrium theory. These models could be classified as *walrasian*, assuming the full employment of production factors, and they primarily focused on the optimal allocation of resources. Nevertheless, it became rapidly obvious that the usual neoclassical hypotheses did not match with the complexity of the real economy. The South African CGE model belongs to the last generation of CGE models where these hypotheses have been progressively relaxed in order to allow for more and more market imperfections such as unemployment, price rigidities, and imperfect competition..

However, regarding the estimation of poverty and inequality measures, the main drawback of CGE models remains the unavoidably limited number of representative household groups (RHG).[18] This means that exogeneity assumptions about the within-RHG income distribution have to be made that do not always fit the historical evidence. In addition, this implies that the within-RHG income distribution has to be held fixed unless another assumption can be made regarding the way it varies. Robilliard et al. (2001) noted that the limited number of RHG is leading to "a systematic underestimation of the impact on inequality, (…) [and] an underestimation of the impact on poverty".

[17] About CGE closures, see Sen (1963) and Robinson (2003).
[18] For more details on the limitations of CGE models see Round and Whalley (2002), Iqbal and Siddiqui (2001), Hérault (2003), Dervis et al. (1982), Thurlow and van Seventer (2002) and Thurlow (2003).

3.2. Key characteristics

The aim of this section is to provide the reader with a basic background on the macro data and the functioning of the model. Table 9 presents the structure of the South African gross domestic production.

Table 9: Structure of Gross Domestic Production (2000)[(a)]

	Value (Billions of Current Rand)	Share of GDP (Market Prices)
Private consumption	556.7	59.8
Fixed investment	131.8	14.1
Inventory changes	8.7	0.9
Government consumption	209.9	22.5
Exports	249.1	26.7
Imports	-224.6	-24.1
GDP (market prices)	931.6	100
Net indirect taxes	100	10.7
GDP (factor cost)	831.6	89.3

Note: The average exchange rate in 2000 was $US 1 = R6.94. *Source: 2000 South African SAM*

The low share of investment (14.1% of GDP) remains a structural characteristic of the South African economy and it is remarkably stable since it was already 14.1% of the GDP in 1993. South Africa is classified by the World Bank (2001) as an upper middle-income country with a GDP per capita of $2,620. The country can be thought of as an economic leader in the region, Sub-Saharan Africa, which is primarily composed of low-income and least developed countries. Indeed, the South African GDP accounts for more than one third of the Sub-Saharan Africa global GDP and its GDP per capita is more than five times higher than the regional figure. International trade (i.e. exports plus imports) makes up almost half of the GDP, where the European Union is, by far, the most important trading partner. The share of international trade in GDP is still below the average for Sub-Saharan Africa but it has been increasing rapidly since the early 1990s, partly due to the implementation of economic policies directed towards more openness. Unlike the neighbouring countries, manufactured products represent the bulk of South African exports.

The South African CGE model is based on the 2000 SAM, which includes 43 sectors and 4 factors of production: high-skilled labour, skilled labour, low-skilled labour and one type of

capital. The South African trade has been regionally disaggregated as in Thurlow (2003).[19] Following the Armington specification (Armington, 1969), imported and domestic products are imperfect substitutes which allow for intra-branch trade. There is no data available at the macro level regarding the informal sector, which is therefore not represented in the SAM. The results from the CGE model are available under three different sets of closures: two Keynesian closures and one neoclassical closure.[20] Each of these closures is based on a number of different assumptions regarding some key components of the economy, as discussed in Table 10.

Table 10: Closure rules

	Keynes 1	**Keynes 2**	**Neoclassical**
Factor market: Skilled and low-skilled labour	fixed nominal earnings; mobile factors; flexible supply	fixed real earnings; mobile factors; flexible supply	
High-skilled labour	flexible earnings; mobile factor; flexible supply		
Capital	flexible activity-specific returns; immobile factor; fixed supply		
Savings-Investment	fixed investment; fixed savings rates		flexible investment; fixed savings rates
Current account	flexible exchange rate; fixed foreign savings	fixed exchange rate; flexible foreign savings	flexible exchange rate; fixed foreign savings
Government account	flexible government savings; fixed direct tax rates; fixed real government consumption		
Numeraire price	flexible consumer price index; flexible producer price index; fixed earnings (=numeraire)	flexible consumer price index; flexible producer price index; fixed exchange rate (=numeraire)	flexible consumer price index; fixed producer price index (=numeraire)

These constraints determine to a large extent the functioning of the modelled economy. Table 10 describes briefly how each of the closures affects the factor market, the savings-investment account, the current account and the government account.[21] In each closure, unemployment amongst low-skilled and skilled workers is allowed by assuming either nominal or real earnings to be fixed. All closures have a short term focus since capital is not mobile across

[19] There are 10 trading partners in the model, namely: Southern African Development Community (SADC), Rest of Africa (ROA), United States of America (US), Southern Cone Common Market (Mercosur), European Union (EU), India, China, Japan, Rest of East Asia (ROEA), and Rest of World (ROW).

[20] These designations are only indicative. In the CGE literature, the differences between closures are usually more pronounced. Here, the main distinction concerns the savings-investment account whereas the closures of the other accounts are more in concordance with empirical observations rather than economic theories or the short-run focus of the model.

[21] More information on the closures of the South African CGE can be found in Hérault (2004), Thurlow and van Seventer (2002) and Thurlow (2003).

sectors. Most important when dealing with CGE models is how savings and investment adjust in each closure.

In the Neoclassical closure, investment is savings-driven. That is, the investment level is the endogenous variable that adjusts *ex post* to the level of savings. In this specification, an increase in the government deficit (equivalent to a decrease in government savings) depresses investment because of the implicit higher interest rate. Hence, a crowding-out of investment follows any growing government deficit. The long-term empirical study by Nell (2003) supports this point of view in the case of South Africa.

The adjustment mechanism is less obvious in the Keynesian closures because both investment and savings are assumed to be fixed and both producer and consumer prices are flexible. In the first Keynesian closure, nominal earnings of skilled and low-skilled workers are fixed. If, after an exogenous shock, savings happen to be insufficient then a price increase is required to lower real wages. Consequently, this generates a boost in production and employment until savings reach the desired level of investment. The implicit assumption is that unions are too weak (for instance because of the high level of unemployment) to obtain any improvement in earnings when prices go up. The main concern with this closure is that it does not allow for any expansion of employment without a fall in real earnings, and that may not be consistent with the economic reality (see Robinson, 2003). Moreover, the empirical study over the period 1997-2001 by Muller et al. (2004) supports the view that real earnings are fixed in the South African formal sector.

The second Keynesian closure alleviates the latter concern by considering real rather than nominal earnings to be fixed. However, this is achieved at the expense of the current account closure where the nominal exchange rate now has to be fixed because of the need for a numeraire. As a result, foreign savings adjust to ensure the balance of the current account. In this specification, a shortfall in savings is basically offset by an increase in foreign savings. Indeed, the rise in earnings following any increase in prices implies a real appreciation and thus, a deterioration of the trade balance which can only be counterbalanced by an increase in foreign savings.

4. Linking CGE and Microsimulation modelling

This paper applies a "top-down" approach to link the South African CGE and MS models. The idea of linking these two types of model was envisaged for the first time by Dervis, de Melo and Robinson (1982). This idea was realised by the end of the 1990s (see Decaluwé et al. 1999, Cogneau 1999) and the literature on this subject has been flourishing since then. The aim of combining these models is to exploit the advantages of CGE and MS modelling. This approach offsets the respective drawbacks of each type of model, which are primarily the lack of general equilibrium effects in MS models and the limitations arising from the representative household assumptions in CGE models.

By providing a comprehensive picture of the economy, micro-macro models (the combination of a CGE and a MS model) allow an analysis at the microeconomic level of macroeconomic policy simulations. There are several ways of linking these two types of model.[22]

A first approach consists of increasing the number of representative households in the CGE model. These so-called integrated CGE-MS models can incorporate as many households as found in household surveys (see Cockburn 2002, Cororaton and Cockburn 2005, Cogneau and Robilliard 2000). The heterogeneity in household behaviours, as well as in expenditure and income patterns, is thus captured by the model in a general equilibrium framework. Although this seems to be the ideal approach, the data requirements can prove to be large and full reconciliation between micro and macro data is essential. Moreover, the size of the model can quickly become problematic and force the modeller to impose some simplifications either on the complexity of microeconomic household behaviours or on the size of the CGE model in terms of the number of sectors and factors of production.

The "top-down" approach used in this paper is an alternative to integrated models. It relies on using a CGE and a MS model in a sequential way: first the CGE model is run, followed by a second step in which the changes in some selected variables are passed on to the MS model (see Robilliard et al., 2001).[23] As is the case with integrated models, "top-down" models have the advantage of avoiding the use of representative agent assumptions, while accounting for general equilibrium effects. Since both models are run separately, nothing prevents the use of fairly

[22] See Savard (2003, 2004), and Cororaton and Cockburn (2005) for a more extensive survey of the literature.
[23] Our model borrows from Robilliard et al. (2001) with regard to the methodology being employed.

comprehensive models. This method also has the advantage of not formally requiring full reconciliation of micro and macro data. However, it also implies a lack of theoretical consistency because nothing guarantees coherence between the CGE and the MS models

Savard (2003) presents a method capable of overcoming this latter problem by using a so-called "top-down bottom-up" (TDBU) approach. In a CGE-TDBU model aggregate results from the MS model are incorporated in the CGE model. A loop is used to run both models iteratively until convergence is obtained. However, the existence of a converging solution is not guaranteed.

4.1 Application of the "top-down" approach

The first step consists of running the CGE model to simulate the policy change or the macro-shock to be analysed. The model returns the new macro-structure of the economy after the "shock", while taking into account the interactions between the various sectors of the economy. In the context of the "top-down" approach, three sets of variables are of particular interest: prices, returns from capital and labour, and employment levels. In a second step, the changes in these variables are passed on to the MS model. With regard to the prices, this procedure is relatively straightforward, because prices are exogenous to the MS model. The original expenditure items from the IES have been mapped to the corresponding commodity groups according to the Standard Industrial Classification (SIC) codes.[24] The 96 commodity groups obtained from this mapping have been further aggregated to match the 43 goods of the CGE model. The 43 price changes computed by the CGE model are simply passed on to the MS model.[25]

The procedure is more complex for the two other sets of variables. The changes from the CGE model cannot be directly transmitted to the MS model since the MS model is based on microeconomic data whereas the CGE model only returns macro numbers. The two sets of numbers are not automatically consistent with each other. In the "top-down" approach, the macro outcomes are imposed on the micro model. That is, coefficients of the MS model have to be modified in such a way that it reproduces the macro numbers obtained from the CGE model, while allowing for the price and factor return changes which may affect individuals' behaviours.

[24] See PROVIDE project (2005) for more details.
[25] The new prices are used to compute a household-specific consumer price index (see section 2.2.3).

This is achieved by applying micro-macro consistency equations, as explained in the next subsection.

4.2 Micro-macro consistency equations

The CGE model provides us with the total earnings from wages and salaries, and total returns from capital after the policy change. Given that three types of labour are incorporated in the CGE model, it means that it returns three earnings; that is the earnings for high-skilled, semi-skilled and low-skilled workers. These earnings concern only the formal workers since the informal sector is not represented in the CGE model. In contrast, capital returns potentially concern all households.

The earnings changes are transmitted to the MS model by applying the resulting earnings changes per employee from the CGE model to the predicted earnings of each individual in the formal sector. Equation 5 expresses this relation:

$$PGE_{i,j,k} = PGE_{i,j,k}.(1 + \Delta W_k) \qquad k=LS,SK,HS, \quad j=formal\ sector \qquad (5)$$

Depending on the skill level k of the considered individual i, the corresponding percentage change in earnings ΔW_k, derived from the CGE model, is added to the predicted gross earnings of individual i as a formal worker $PGE_{i,j,k}$.[26] Given that the structure of the labour force by skill level is very similar in both macro- and micro-data, the MS model is thus able to reproduce the changes in the aggregated figures from the CGE model concerning the formal earnings.

Regarding informal earnings, the assumption is that their change depends on the changes in formal employment levels and total formal earnings as represented by equation 6:

$$PGE_{i,j,k} = PGE_{i,j,k}.(1+\Delta FE + \tfrac{1}{2}*\Delta E_k) \qquad k=LS,SK,HS, \quad j=informal\ sector \qquad (6)$$

where ΔFE is the percentage change in total formal earnings and ΔE_k is the percentage change in formal employment by skill level. The underlying assumption is that working in the informal

[26] More precisely, the predicted gross earnings are a mix of observed and predicted earnings (see section 2.2 for further details).

sector is a survival strategy and that formal workers are the main consumers of informal goods. Consequently, when the formal sector is expanding one would expect fewer people to rely on the informal sector while at the same time there should be more demand for informal goods. As a result, individual informal earnings are expected to depend positively on total formal earnings and formal employment.[27] This might prove to be unrealistic but other better-informed choices are limited due to the lack of empirical studies.[28]

The change in capital returns is transmitted to the MS model at the household level since income from capital is a household level variable in the micro-data. The resulting percentage change in capital returns predicted in the CGE is applied to the total income from capital of each household, as reported in equation 7:

$$K_h = K_h.(1+\Delta K) \qquad (7)$$

The capital income K_h of each household h is increased (or decreased) by ΔK, where ΔK is the percentage change in capital returns from the CGE model.

In addition to capital income, other components of household non-labour income are also affected by the changes from the CGE model. Transfer incomes from other South African households are updated using changes in formal and informal earnings in accordance with the sectoral distribution of each racial group.[29] Transfer incomes from abroad are assumed to be fixed in the foreign currency and are consequently adjusted following changes in the exchange rate. Since real government expenditures are fixed in the CGE model, transfer incomes from the government are indexed to the CPI. However, the model could be easily extended to account for changes in government transfers.

[27] The unity coefficient in front of ΔFE is based on the assumption that the elasticity of individual informal earnings with respect to total formal earnings is equal to one. The elasticity of individual informal earnings with respect to total formal employment is assumed to be 0.5, which is derived from preliminary results of the MS model which have shown that the formal employment elasticity of informal employment is approximately 0.5.

[28] Another alternative was explored in which individual informal earnings vary in the same way as individual formal earnings by skill level. Although the direction of the results remained largely unaffected, all impacts were lower because of the implied cut in the growth of informal earnings. This was particularly true under the first Keynesian closure because of decreasing real earnings for low-skilled and skilled formal workers.

[29] The underlying assumption is that the majority of these transfers is generated by workers in the formal or informal sector, who live outside the household.

Although this was not explicitly mentioned in section 2.2, household income from capital and transfers is included in the selection model.[30] As a result, any change in capital income or in the transfer incomes of a given household can potentially affect the labour market choices of its members.

Even though the changes in the predicted earnings and the capital returns already imply that some people will switch from one sector to another, this is not sufficient to ensure full consistency between the two models as far as employment levels are concerned. Changes in the number of formal workers by skill level in the MS model must match those same changes in the CGE model. This can be done by modifying some specific coefficients of the selection model. The most straightforward choice is to target the coefficients associated with the skill level in the equation defining the utility level in the formal sector. As defined in equations (1) and (2) (see section 2.2), these coefficients are *B1* for skilled individuals, *B2* for high-skilled individuals and the constant *a* for low-skilled individuals, since the latter group is the reference category. Moreover, given that the selection model is estimated separately for four demographic groups, it means that to solve the consistency problem, a new set of 12 parameters ($a_{h(i),FS}$, $B1_{h(i),FS}$, $B2_{h(i),FS}$) is required such that the numbers of formal low-skilled, skilled and high-skilled workers equal the corresponding numbers obtained from the CGE. This constraint is described by:

$$\Sigma_i \Sigma_q \; Ind[U_{i,FS}(a_{h(i),FS},\; B1_{h(i),FS},\; B2_{h(i),FS})+u_{i,FS,q} = Max_j \; (U_{i,j}(a_{h(i),j},\; B1_{h(i),j},\; B2_{h(i),j}))+u_{ijq} \text{ and } skill_i=k] \; /Q= FSW_k.(1+\Delta E_k) \text{ with } k=LS,SK,HS \qquad (8)$$

where $U_{i,FS}$ defines the utility of individual i as a formal worker. $u_{i,FS,q}$ and u_{ijq} are the q^{th} draws from the error distribution for individual i and category FS or category j and Q is the total number of draws selected (see section 2.2.1). FSW_k is the benchmark number of formal workers, with skill level ($skill_i$) equal to k, in the MS model and ΔE_k is the percentage variation in the number of formal workers predicted by the CGE model.

[30] Transfer incomes from South African households and from abroad are included in the predicted earnings while capital income and government transfers are considered as separate independent variables.

To summarise, the consistency problem now consists of solving a system of three equations and 12 variables. Consequently, further assumptions are necessary. The choice made in this paper is to impose the same percentage changes on the parameters for the four demographic groups:[31]

$$a_{h(i),FS} = a'_{h(I),FS}.(1+\Delta_a), \qquad h(i)=sw, mw, sm, mm \qquad (9)$$

$$B1_{h(i),FS} = B1'_{h(i),FS}.(1+\Delta_{B1}), \qquad h(i)=sw, mw, sm, mm \qquad (10)$$

$$B2_{h(i),FS} = B2'_{h(I,),FS}.(1+\Delta_{B2}), \qquad h(i)=sw, mw, sm, mm \qquad (11)$$

where ($a'_{h(i),FS}$, $B1'_{h(i),FS}$, $B2'_{h(i),FS}$) are the parameter values as estimated using the micro data, and Δ_a, Δ_{B1} and Δ_{B2} are the percentage changes imposed respectively on these parameter values in all demographic groups. This particular choice implies that the MS model is allowed to determine which individuals, amongst the entire population, will fill the need for more formal workers if their number is to increase in the CGE model. On the contrary, if the number of formal workers is found to decrease, then the MS model will freely choose the individuals with the highest probability to lose their job, amongst all formal workers.[32] Another option would have been, for instance, to keep constant the proportions of formal workers by demographic group. Our choice restricts the utility of being in the formal sector due to an individual's specific skill level to increase with the same percentage across the four demographic groups.[33]

Now, there are three endogenous variables (Δ_a, Δ_{B1}, Δ_{B2}) to solve the three consistency equations presented in (8). This is done through an iterative process. The first step is to find the appropriate change (Δ_a) in the constant to obtain the appropriate number of low-skilled formal workers. Effectively, since the latter group is the reference category, this number is not affected by changes in *B1* and *B2*.[34] The second step consists of finding the appropriate change (Δ_{B1}) in *B1* that results in the required number of skilled formal workers. In the last step, the endogenous variable (Δ_{B2}) is adjusted to obtain the desired number of high-skilled formal workers.

---

[31] sw, mw, sm and mm stand for single women, married women, single men and married men respectively.

[32] In fact, the process is slightly more complex. The MS model allows some people to find a formal job and others to lose their formal job independent of whether the CGE model predicts an increase or a decrease in the aggregate number. The consistency constraints concern only the aggregate results of the MS model since the CGE model only returns numbers at the macro level.

[33] Given that already more than two thirds of the married men were formal workers at the time of the household survey, it seems sensible to allow for a variation in their share of all formal workers, especially if the total number of formal workers is found to increase after a policy change.

[34] *However, it is essential to solve first for* Δ_a since the changes in the constant also have an impact on the utility functions of skilled and high-skilled workers.

Regarding the informal workers, no constraint is imposed on the macro outcomes of the MS model since this segment of the labour market is not included in the CGE model. Indeed, only the macro outcomes concerning the formal sector, which accounts for 70% of paid workers (see table 1), are imposed on the MS model. As a result, the number of people in the four other sectors (inactive, unemployed, subsistence agriculture and informal sector) is entirely determined by the MS model as a function of individual characteristics and as a function of the required changes in formal employment.

5. Conclusion

In this paper, a CGE model and a MS model are combined in a sequential approach in order to build an effective tool to assess the effects of various macroeconomic policies and shocks on South African households. The CGE model is use to simulate the changes in prices, capital returns, earnings and employment after the policy change. In a second step, these changes are passed on to the MS model. Micro-macro consistency equations, along with the direct transmission of prices, ensure that changes in prices, in earnings from wages and salaries, in returns from capital, and in employment levels are fully transmitted from the CGE to the MS model. Given any change in the macroeconomic structure of the economy predicted by the CGE model, the MS model predicts how individual agents modify their behaviours and how their incomes are affected, while accounting for individual heterogeneity. Therefore, it provides us with an updated picture of the economy at the microeconomic level taking into account the simulated changes in macroeconomic policies.

Finally, the sensitivity analyses carried out in this paper reveal that the model is fairly robust to changes in both the CGE closures and the MS model's assumptions made upon the links between the formal and the informal sector. However, additional sensitivity analyses may still prove useful. At a later stage, the plan is to extend the MS model to allow for discrete hours choices so changes in labour market choices can be modelled more precisely, but this is outside the scope of the current paper.

The "top-down" CGE-MS model presented in this paper is designed to offset the respective drawbacks of each type of model, which are primarily the lack of general equilibrium effects in MS models and the limitations arising from the representative household assumptions in CGE

models. Such a model can be particularly useful when the microeconomic effects of macroeconomic policies or shocks are at stake. This CGE-MS model could be used to analyse the microeconomic impacts of various policies and shocks like trade liberalisation, expansionary monetary policy leading to a large depreciation of the South African Rand, expansionary or restrictive fiscal policy, a sudden rise or fall in the prices of some imported and/or exported commodities, a shortage or boom in foreign investment, an increase in government transfers to the poor, an increase or decrease in productivity.

References

Armington, P. (1969). A theory of Demand for Products Distinguished by Place of Production. IMF Staff Papers, 16, 159-178.

Casale, D., C. Muller and D. Posel(2004). "Two million net new jobs": A reconsideration of the rise in employment in South Africa, 1995 –2003. *South African Journal of Economics*, 72(5): 978-1002.

Cockburn, J. (2002). Trade Liberalisation and Poverty in Nepal: A Computable General Equilibrium Micro Simulation Analysis. CREFA, Université Laval, CSAE WPS/2002-11.

Cogneau, D. (1999). Labour Market, Income Distribution and Poverty in Antananarivo: A General Equilibrium Simulation. mimeo, DIAL, Paris.

Cogneau, D. and A.S. Robilliard (2000). Growth, Distribution and Poverty in Madagascar: Learning from a Microsimulation Model in a General Equilibrium Framework. International Food Policy Research Institute, Trade and Macroeconomics Division, discussion paper no. 61.

Cororaton, C.B. and J. Cockburn (2005). Trade Reform and Poverty in the Philippines: a Computable General Equilibrium Microsimulation Analysis. Centre interuniversitaire sur le risque, les politiques économiques et l'emploi, Working Paper 05-13

Creedy, J., A.S. Duncan, M. Harris, and R. Scutella (2002). Microsimulation Modelling of Taxation and The Labour Market: The Melbourne Institute Tax Transfer Simulator. Cheltenham: Edward Elgar.

Decaluwé, B., J.-C. Dumont and L. Savard (1999). Measuring Poverty and Inequality in a Computable General Equilibrium Model. Working paper 99-20, CREFA, Université Laval, Quebec, Canada.

Dervis, K., J. de Melo and S. Robinson (1982). General Equilibrium Models for Development Policy. New York: Cambridge University Press.

Hérault, N. (2003). Mondialisation et pauvreté : les faiblesses des modèles d'équilibre général calculables. Centre d'Économie du Développement (IFReDE-GRES), Université Montesquieu Bordeaux IV, Document de travail n°87.

— (2004). Un modèle d'équilibre général calculable (MEGC) pour évaluer les effets de l'ouverture au commerce international : le cas de l'Afrique du Sud. Centre d'Économie du Développement (IFReDE-GRES), Université Montesquieu Bordeaux IV, Document de travail n°102.

— (2005). Trade Liberalisation, Poverty and Inequality in South Africa: A CGE-Microsimulation analysis. Melbourne Institute Working Paper, WP No. 17/05, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne.

Hoogeveen, J.G. and B. Özler (2004). Not Separate, Not Equal Poverty and Inequality in Post-Apartheid South Africa. World Bank.

Human Sciences Research Council (HSRC). (2004). Fact Sheet: Poverty in South Africa.

Iqbal, Z. and R. Siddiqui (2001). Critical Review of Literature on Computable General Equilibrium Models. MIMAP technical paper series no. 9.

Jonsson, G. and A. Subramanian (2001). Dynamic Gains from Trade: Evidence from South Africa. IMF Staff Papers, Vol. 48, No. 1.

Maddala, G.S. (1983). *Limited Dependent and Qualitative Variables in Econometrics*. New York: Cambridge University Press.

Muller, C. and J. Esselaar (2004). The Changing Nature of Work in South Africa: Evidence from Recent National Household Surveys. Paper presented at the EGDI and UNU-WIDER Conference *Unlocking Human Potential: Linking the Informal and Formal Sectors* 17-18 September 2004, Helsinki, Finland.

Nell, K. (2003). Long-run Exogeneity Between Saving and Investment: Evidence from South Africa. Working Paper 2-2003, Trade and Industrial Policy Strategies, Johannesburg, South Africa.

O'Donoghue, C. (2001). Introduction to the Special Issue on Dynamic Microsimulation Modelling. London School of Economics, University of Cambridge,Recife, December 15th.

PROVIDE Project (2005). Creating a 2000 IES-LFS Database in STATA. Technical Paper 2005:1.

Robilliard, A.-S., F. Bourguignon and S. Robinson (2001). Crisis and Income Distribution: A Micro-Macro Model for Indonesia. Paper presented at the OECD Development Center Conference, 9-10 December 2002, Paris, France.

Robinson, S. (2003). Macro Models and Multipliers: Leontief, Stone, Keynes, and CGE Models. International Food Policy Research Institute.

Round, J. and J. Whalley (2002). Globalization and Poverty: Implications of South Asian Experience for the Wider Debate. Paper presented at the OECD Development Center Conference, 9-10 December 2002, Paris, France.

Savard, L. (2003). Poverty and income distribution in a CGE-household sequential model. preliminary draft.

—, (2004). Poverty and Inequality Analysis within a CGE Framework: a Comparative Analysis of the Representative Agent and Micro-Simulation Approaches. Centre interuniversitaire sur le risque, les politiques économiques et l'emploi, Working Paper 04-12

Sen, A.K. (1963). Neo-Classical and Neo-Keynesian Theories of Distribution. *Economic Record*, Vol. 39, pp. 54-64.

Simkins, C. (2004). What Happened to the Distribution of Income in South Africa Between 1995 and 2001?. University of the Witwatersrand.

Thurlow, J. (2003). A Dynamic Computable General Equilibrium (CGE) Model for South Africa: Extending the Static IFPRI Model. International Food Policy Research Institute, Washington, D.C., L'Institut de Recherche pour le Développement, Paris, and University of Natal, Durban.

Thurlow, J. and D.E.N. van Seventer (2002). A Standard Computable General Equilibrium Model for South Africa. Trade and Macroeconomics Discussion Paper No. 100, International Food Policy Research Institute, Washington, D.C. and, Trade and Industrial Policy Strategies, Johannesburg.

Van der Berg, S. and M. Louw (2003). Changing patterns of South African income distribution: Towards time series estimates of distribution and poverty. Paper to the Conference of the Economic Society of South Africa, Stellenbosch, 17-19 September 2003, Department of Economics University of Stellenbosch.

Van der Berg, S., S. Nieftagodien and R. Burger (2003). Consumption patterns and living standards of the black population in perspective. Paper presented at the Biennial Conference of the Economic Society of South Africa, Somerset West.

World Bank, (2001). Global Economic Prospects and the Developing Countries 2001.

APPENDIX A: SUMMARY STATISTICS

Table A: Summary statistics by demographic group (weighted mean)

	Single women	Married women	Single men	Married men
Black	0.85	0.67	0.84	0.66
Coloured	0.09	0.11	0.08	0.11
Asian	0.02	0.05	0.02	0.05
White	0.05	0.17	0.06	0.18
No education & pre-school	0.07	0.10	0.05	0.10
Primary education	0.22	0.27	0.24	0.27
Standard 8	0.29	0.27	0.29	0.25
Standard 10	0.39	0.31	0.38	0.29
Tertiary education	0.03	0.05	0.03	0.08
Education attendance	0.17	0.03	0.19	0.03
Low-skilled	0.59	0.53	0.65	0.59
Semi-skilled	0.38	0.42	0.32	0.29
High-skilled	0.02	0.06	0.03	0.12
Age squared/100	10.50	15.11	8.72	18.26
Ability to write	0.93	0.90	0.94	0.91
Afrikaans language	0.10	0.19	0.10	0.20
English language	0.06	0.14	0.06	0.15
Indigenous	0.84	0.67	0.84	0.66
Agri. Activity	0.13	0.17	0.11	0.11
Western Cape	0.08	0.12	0.08	0.13
Eastern Cape	0.15	0.13	0.15	0.10
Northern Cape	0.02	0.02	0.02	0.02
Freestate	0.06	0.07	0.06	0.08
Kwazulu-Natal	0.24	0.18	0.22	0.17
North-West	0.09	0.07	0.09	0.08
Gauteng	0.19	0.21	0.20	0.27
Mpumalanga	0.07	0.07	0.06	0.07
Limpopo	0.11	0.12	0.11	0.09
Urban location	0.59	0.62	0.58	0.68
Homeland location	0.34	0.29	0.33	0.20
Household size	5.87	4.82	5.40	4.30
Number of male	2.24	2.30	3.08	2.22
Children (number)	1.96	1.72	1.40	1.38
Children 0-1 yr old	0.12	0.11	0.06	0.09
Children 1-3 yrs old	0.25	0.24	0.16	0.20
Children 4-5 yrs old	0.11	0.12	0.09	0.10
Children 6-9 yrs old	0.14	0.15	0.13	0.13
Children 9-15 yrs old	0.11	0.11	0.12	0.10
OAP*** in the hh	0.23	0.10	0.24	0.07
Family allow. in the hh	0.05	0.02	0.04	0.02
Capital income/1000**	1.07	4.05	1.32	4.07
Gov. transfers to the hh**	0.92	0.67	0.86	0.59
Partner live here	2.00	0.84	2.00	0.85
No education & pre-school (partner)	0.00	0.09	0.00	0.07
Primary education (partner)	0.00	0.21	0.00	0.21
Standard 8 (partner)	0.00	0.21	0.00	0.23
Standard 10 (partner)	0.00	0.25	0.00	0.30
Tertiary education (partner)	0.00	0.07	0.00	0.05
Age (partner)	0.00	35.93	0.00	31.99
Low-skilled (partner)	0.00	0.48	0.00	0.42
Semi-skilled (partner)	0.00	0.10	0.00	0.06
High-skilled (partner)	0.00	0.25	0.00	0.37
Predicted inactive inc.*	1,050	961	842	449
Predic. informal earnings*	9,299	11,633	9,361	32,320
Predic. formal earnings*	16,353	23,598	18,920	41,511

Note: *in Rand per year **in 1000's Rand per year *** Old Age and War Pension