Centre d'économie du développement

IFReDE - GRES - Université Bordeaux IV

Document de travail

DT/114/2005

Building and Linking a Microsimulation Model To a CGE Model: The South African Microsimulation Model

par

Nicolas Hérault

PhD student – CED / IFReDE-GRES – Université Montesquieu-Bordeaux IV

Avenue Léon Duguit - 33608 Pessac (France) - tél : 0556842938 - fax : 0556848506

ced@u-bordeaux4.fr

http://ced.u-bordeaux4.fr

Sommaire

1. Introduction **3**

2. Microsimulation Modelling **3**
2.1. Overview of the microeconomic data 4
2.2. Key characteristics of the model 6

3. CGE Modelling **9**
3.1. Background 9
3.2. Key characteristics 10

4. Linking CGE and Microsimulation modelling **12**
4.1. Application of the "top-down" approach 13
4.2. Micro-macro consistency equations 14

5. A first simulation: the removal of all tariffs **17**
5.1. Macro-results from the CGE model 17
5.2. Micro-results from the microsimulation model 20

6. Conclusion 22

References ***23***

Appendix ***25***

Building and Linking a Microsimulation Model to a CGE Model: The South African Microsimulation Model[1]

Nicolas Hérault[2]
PhD student

Centre d'Économie du Développement (IFReDE-GRES)
Université Montesquieu Bordeaux IV

Abstract

This paper describes the project of building a micro-macro model for South Africa. The aim is to deal with the links between globalisation and poverty or inequality, explaining the effects of trade liberalisation on poverty and inequality. The main issue of interest is the effect of international trade on households (especially their income); some changes may contribute to reduce poverty while other changes could work against the poor. The approach presented in this paper relies on combining a macro-oriented CGE model and a microsimulation model. Combining these two models the microeconomic effects (on poverty and inequality) of a macroeconomic policy (trade liberalisation) can be analysed. The paper gives details about the microsimulation model and the "top-down" approach used to link the microsimulation model and the CGE model. In addition, the methodology discussed is applied to South African data and a selection of preliminary results using this approach are presented and discussed. The main concern regarding poor households is whether the decrease in real (or nominal) earnings for formal low-skilled and skilled workers is offset by the upward trend in formal employment levels. This appears to be the case implying a decrease in poverty due to trade liberalisation. Although whites emerge as the main winners, the increase in inter-group inequality is more than compensated by the decrease in intra-group inequality.

Résumé

Ce papier décrit le projet d'élaboration d'un modèle micro-macro pour l'Afrique du Sud. L'objectif est d'examiner les liens entre la mondialisation et la pauvreté ou l'inégalité, en expliquant les effets de la libéralisation commerciale sur ces deux indicateurs de progrès social. La préoccupation principale concerne l'effet du commerce international sur les ménages (particulièrement, leur revenu), certains changements pouvant contribuer à réduire la pauvreté, tandis que d'autres étant susceptibles d'aggraver les privations. L'approche présentée dans cet article est fondée sue la combinaison d'un modèle CGE orienté-macro, et d'un modèle de micro-simulation. En combinant ces deux modèles, les effets micro-économiques (sur la pauvreté et l'inégalité) d'une politique macro-économique (libéralisation commerciale) peuvent être analysés. L'étude spécifie le modèle de micro-simulation et l'approche "top-down", employés pour relier les modèles de micro-simulation et CGE. En outre, la méthodologie discutée est appliquée aux données sud-africaines, et des résultats préliminaires, fondés sur cette approche, sont présentés et discutés. Un élément central de l'analyse concernant les ménages pauvres est d'examiner si la diminution des revenus réels (ou nominaux) des ouvriers qualifiés ou faiblement qualifiés du secteur formel est compensée par la tendance à la hausse de l'emploi formel. L'étude montre que cela semble être le cas, ce qui implique une diminution de la pauvreté due à la libéralisation commerciale. Bien que les bénéficiaires principaux soient les « blancs », l'augmentation de l'inégalité inter-groupes est plus que compensée par la diminution de l'inégalité intra-groupes.

---

[1] 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.

[2] I would like to thank Jean-Pierre Lachaud, Daniela Casale, James Thurlow, Colette Muller, John Cockburn, Ramos Mabugu, Margaret Mabugu and Kalie Pauw for their suggestions. 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.

1. Introduction

This paper aims to deal with the link between globalisation and poverty or inequality through simulation of the effects of trade liberalisation on poverty and inequality in South Africa. This is achieved by building a micro-macro model for South Africa.

The main issue of interest is the effect of international trade on households (especially their income); some changes may contribute to reduce poverty while other changes could work against the poor or the effect may be ambiguous by being beneficial for some groups of poor people but not for others. 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, we are 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,[3] they do not address the question adequately because they miss the important micro-macro links.

The approach presented in this paper combines a macro-oriented Computable General Equilibrium (CGE) model and a microsimulation (MS) model following Robilliard et al. (2001). However, this approach is extended by also taking into account the changes in capital returns at the household level. This is similar to what has been done in some more complex CGE-MS models.[4] Combining these two models the microeconomic effects (on poverty and inequality) of a macroeconomic policy (trade liberalisation) can be predicted.

The second section gives details about the micro-data and the microsimulation model to be implemented, focussing on the computation of labour market responses. The third section of the paper describes both the macro-data and the CGE model. The fourth section provides details on the "top-down" approach used to link the MS model and the CGE model. In addition, the paper applies the methodology discussed in the previous sections to South African data. A selection of results from the CGE and the MS model are presented and discussed in section 5. Finally, conclusions are provided in section 6.

2. Microsimulation Modelling

Guy Orcutt is known as the originator of microsimulation modelling (see Orcutt 1957, and Orcutt et al. 1976). MS modelling allows 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.

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

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

[4] For instance, see the integrated model by Cororaton and Cockburn (2005).

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. Then, the more complex behavioural MS models are more appropriate because this type of model allows individuals to adjust their behaviour in response to the simulated policy change.

In addition, 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.[5] In a static framework, the size and demographic characteristics of the population from the original cross-sectional dataset used in the simulation are fixed. The dynamic framework, however, also considers the demographic phenomena that affect the original population, such as changes in the mortality and fertility rates or in the time taken out for education.

Finally, MS models usually are partial equilibrium models, focussing on the household side of the economy. As a result, 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.

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 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.[6]

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

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

[6] 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).

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*

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. 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 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.

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 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]

[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.

[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).

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*

2.2. Key characteristics of the model

The South African MS model is a static MS model allowing for behavioural responses. 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.

The model simulates the new labour market choices after changes in individual characteristics, such as wages (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]

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:

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

This utility function is defined separately for the four demographic groups considered in the model: single women, married women, single men and married men. 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

[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 and incomes are fixed at the observed levels. Moreover, all recipients of the Old Age and War pension, as well as full-time students aged between 15 and 18 years, are excluded from the simulation.

net earnings,[10] skill level,[11] 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 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).[12]

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} 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 policy 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.

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

[11] 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.

[12] The estimated parameters of the selection model are presented in Appendix A.

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}:

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

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.

2.2.3. Calculation of household income

After running the selection and the regression models described above, individual earnings are added to other (observed) 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 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

[13] Chow tests confirm that the coefficients are significantly different in the four demographic groups. The estimated parameters of the regression model are presented in Appendix B.

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

3. CGE modelling

The development of CGE models can be traced to Johansen (1960), Harberger (1962) and Scarf (1973). The main innovation introduced by these 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, openness to trade 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

In 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 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.[16]

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

[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 results are 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 always found to increase.

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

[16] About CGE closures, see Sen (1963) and Robinson (2003).

neoclassical hypotheses did not match with the complexity of the real economy. For more than fifteen years, 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).[17] 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".

3.2. Key characteristics

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)[18] which can be classified as a neoclassical-structuralist model in the tradition originally introduced in Dervis et al. (1982). The aim of this section is not to provide the reader with a complete description of the model but rather with a basic background on the macro data and the functioning of the model. Table 3 presents the structure of the South African gross domestic production.

Table 3: 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

[17] 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).

[18] 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.

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 1990's, 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 4.

These constraints determine to a large extent the functioning of the modelled economy. Table 4 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 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

[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.

[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).

Table 4: Closure rules

	Keynes 1	**Keynes 2**	**Neoclassical**
Factor market:	fixed nominal earnings; mobile factors;	fixed real earnings; mobile factors;	
Skilled and low-skilled labour	flexible supply	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)

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

As mentioned in the previous sections, CGE and microsimulation (MS) models have relatively long histories, which go back to the early 1960's. The idea of linking these two approaches was envisaged for the first time by Dervis, de Melo and Robinson (1982). Although, this idea was realised only by the end of the 1990's (see Decaluwé et al. 1999, Cogneau 1999), 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]

[22] See Savard (2003, 2004), and Cororaton and Cockburn (2005) for a more extensive survey of the literature.

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 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

This paper presents an application of the "top-down" approach to the South African case. The focus is on poverty and inequality impacts of trade liberalisation policies.

The first step consists of running the CGE model to simulate the complete removal of import tariffs. 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

[23] Our model borrows heavily from Robilliard et al. (2001), especially with regard to the methodology being employed.
[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).

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. 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 the 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 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]

---

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

[27] The unity coefficient in front of ΔFE is based on the assumption that the elasticity of individual informal earnings with respect to formal earnings is equal to one. The elasticity of individual informal earnings with respect to 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

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:

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

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.

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 a 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:

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

---

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.

[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.

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.

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.

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.

---

[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.

To conclude, the 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 defined 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. The following section uses the CGE-MS model described above to assess the effects of trade liberalisation in South Africa.

5. A first simulation: the removal of all tariffs

Since the early 1990's, South Africa has been involved in a trade liberalisation process. As a result, the decline in tariffs was uninterrupted throughout the 1990's and the weighted average tariff decreased by one third between 1993 and 2000. Moreover, the highly complex tariff regime was simplified. In particular, South Africa committed itself to eliminate, or convert into bound ad valorem rates, all quantitative restrictions by 1998. The process even went beyond the Uruguay Round commitments announced in 1994 (see Jonsson et al., 2001).[35]

South African trade policy relies on unilateral, bilateral and multilateral trade liberalisation in the context of the World Trade Organisation (WTO). The free trade area agreement with the European Union, by far the most important trading partner of South Africa, was signed in 1999 and due to be progressively implemented from 2000 onwards. Unilateral trade liberalisation is also on top of the South African political agenda, mainly because of its commitments under the WTO and the Growth, Employment and Redistribution Programme. This section uses the CGE-MS model described above to assess the short-run effects of the full elimination of tariffs. The first section presents the macroeconomic effects while the second section focuses on poverty and inequality impacts.

5.1 Macro-results from the CGE model

The ratio of import duties collected to the value of imports is only 3.6% but there are some important disparities between the commodities.[36] For instance, petroleum products enter virtually duty free while the tariff on rubber is 24.2%. The total amount of collected duties represents 8.2 billion rand, which is equivalent to 3.6% of government income. Consequently, the removal of tariffs only implies a limited loss of revenue for the government. However, it also affects the economy through the declining prices of imported commodities. Table 5 below summarises the results of the CGE model under three different closures.

The initial impact is due to the lowering of import prices, which causes a shift towards imported goods and away from domestic production. As the value of imports rises and trade balance deteriorates, the exchange rate depreciates which promotes exports and contributes to maintain the current account balance. Regarding the government account, the loss of import duties implies an increase in the government deficit between 0.58 and 0.73 percentage point of

[35] South Africa is a member of the WTO.

[36] All estimates presented refer to the year 2000. The CGE model uses the value of the collected duties rather than the official tariff rates in order to account for the numerous rebates (see Thurlow and van Seventer, 2002).

Table 5: Simulation results from the elimination of import tariffs

	BASE	Keynes 1	Keynes 2	Neoclassical
	Values[(a)]	Percentage change from base year[(b)]		
Real GDP	R888	0.59	0.37	0.30
CPI	-	1.15	-0.57	-0.24
Real exchange rate	-	0.74	0.22	0.57
Nominal exchange rate	-	2.16	-	0.59
Exports (volume)	R249	1.81	0.95	1.41
Imports (volume)	R225	2.01	1.98	1.56
Trade balance	5%	-0.72	-0.48	-0.73
Private savings	R154	3.54	1.15	1.35
Government deficit	-2%	-0.58	-0.69	-0.73
Foreign savings	R4	2.16	57.05	0.59
Investment	15%	-0.21	-0.19	-0.45
Total real household consumption	R556	0.94	0.97	0.86
Factor real returns				
Capital	-	2.59	1.89	1.76
Low-skilled labour	-	-1.32	-	-
Skilled labour	-	-1.13	-	-
High-skilled labour	-	0.55	0.57	0.48
Factor demand				
Low-skilled labour	3,596	1.42	0.65	0.50
Skilled labour	2,718	1.34	0.74	0.64
High-skilled labour	1,118	0.27	0.29	0.24

Note: (a) Values in billions of rand, percentage of GDP and thousands of workers (b) Changes for base values expressed as a percentage of GDP are expressed in percentage points of GDP. *Source: author's calculations*

GDP. Finally, the mechanisms invoked to bring back the balance between savings and investment depend on the closure.

In the first Keynesian closure, money illusion amongst low-skilled and skilled workers[37] implies that the increase in the consumer price index causes their real earnings to decrease. The lower cost of these two production factors, along with the diminished price of imported intermediate goods and the stimulation of exports through the real exchange rate depreciation, is at the root of real GDP growth. This GDP growth, together with the induced job creation, supports the rise in aggregate income. In turn, this allows for an increase in private savings which finances the widening government deficit. A growing economy also means more demand for capital and high-skilled labour. Given the limited supply of these factors, their real returns are found to increase significantly. There is a boost in total real household consumption as the decline in real earnings for low-skilled and skilled workers is more than offset by the expansion in employment and by the increased real returns of capital and high-skilled labour.

Under the second Keynesian closure, there is no money illusion amongst low-skilled and skilled workers and the nominal exchange rate is fixed. Consequently, the deteriorating trade balance is funded by an increase in foreign savings. The expansion in exports is halved because the real depreciation is much less important than in the previous closure. The drop in the consumer price index, resulting from falling import prices, causes nominal earnings of skilled and low-skilled workers to go down. Therefore, their relative competitiveness improves, which

[37] Since nominal earnings are fixed for these workers.

results in a downward pressure on unemployment. The induced economic growth, which is also boosted by the decline in imported intermediate inputs, calls for more use of the two scarce production factors. As a result, real returns of capital and high-skilled labour rise substantially. Given that in this closure there is no contraction in real earnings of low-skilled and skilled workers, and since inflation has now turned into deflation, the increase in total real household consumption is found to be slightly higher than in the previous closure, despite a lower real GDP growth.

The Neoclassical closure implies both fixed real earnings and a flexible exchange rate but the level of investment is now free to adjust to the changes in savings. As in the previous closures, two phenomena are at the origin of real GDP growth: the boost in exports generated by the exchange rate depreciation and the lower prices of imported inputs. However, the expansion of the government deficit is not balanced with an increase in foreign or private savings. Instead, there is a crowding-out of investment, which explains why the recorded real GDP growth is less than what is observed under the Keynesian closures. Once more, trade liberalisation implies a decrease in the consumer price index, which translates into more demand for low-skilled and skilled workers since their real earnings are fixed. The increase in total real household consumption stems from a combination of the lower level of unemployment and the growth in real returns of capital and high-skilled labour.

To summarise, the strongest effects are found when there is money illusion, like in the first Keynesian closure, while the smallest effects occur when allowing for a crowding-out effect of investment, like in the Neoclassical closure. More generally, the effects are greater under the Keynesian closures because savings and investment adjust through the new incomes generated by the fall in unemployment and because there is no crowding-out effect.

More disaggregated data (not shown in Table 5), regarding the level of activity in the 43 sectors of the model, show that there are both losers and winners from trade liberalisation. The sectors benefiting most from trade liberalisation are those orientated towards exports or household final consumption and with a low initial level of protection like transport, communication, trade and catering services. The sectors with the highest initial levels of protection are the systematic losers. They include the following sectors: footwear, rubber, glass, non-metal minerals, metal products and electrical machinery. Depending on the closure, this list can be extended to leather and textile industries. Finally, trade liberalisation seems to be more favourable to services than to manufacturing sectors. This explains why the boost in demand for skilled labour is found to be systematically higher than the surge in demand for low-skilled labour, because manufacturing is more intensive in low-skilled labour than the rest of the economy.

Regarding household incomes, all results converge towards a gain. Nevertheless, given the contrast between the changes in real returns from capital and high-skilled labour on the one hand, and real earnings from low-skilled and skilled labour on the other hand, a worsening of the inter-group income distribution can be expected.[38] Moreover, regarding low-income households, the MS model will have to determine whether the decrease in real (or nominal) earnings for formal low-skilled and skilled workers is offset by the upward trend in formal employment levels.

[38] High-skilled labour and capital are the most important sources of income for rich (mainly white) households whereas (black and coloured) poor households' incomes depend essentially on the earnings of low-skilled and skilled workers.

5.2 Micro-results from the microsimulation model

Table 6 presents the effects of trade liberalisation at the household level for the entire South African population, as well as for each racial group.[39] The effects prove to be substantially different depending on the racial group, especially regarding the changes in labour market status. As will be explained later in further detail, the impacts amongst racial groups are mainly driven by the differences in the subpopulation distribution by skill level and labour market sector. Given that the expansion in formal sector employment is stronger for low-skilled and skilled workers (see table 5), blacks (and, to a lesser extent, coloureds) are found to benefit the most from formal job creations. Indeed, there is a substantial stock of unemployed and inactive low-skilled black workers while Asians and whites have higher skill levels and higher levels of employment. Accordingly, in both closures, more than half of the new formal workers are blacks who were formerly unemployed or inactive. The trend is thus toward a rather significant decrease in poverty incidence.[40] The increase in formal employment is the main force driving poverty alleviation, largely offsetting the negative impact of decreasing formal real earnings under the first Keynesian closure. The larger growth in formal employment under this closure explains the larger poverty reduction despite a lower increase in real income per capita. The dampening effect is even stronger on the poverty gap index and the poverty severity index than on the poverty incidence. This is due to the fact that the poor households moving out of poverty, because one (or more) members find a formal job, were more deeply in poverty than the very few new households moving into poverty because of lower formal earnings. It is observed that poor households with one (or more) members in formal employment tend not to be the poorest households. Another interesting result is that, despite a minor increase in real income, coloureds experience a more important decrease in poverty than blacks. This can be attributed to the much higher share of formal workers amongst coloureds than amongst blacks: despite a smaller percentage increase in the number of formal workers (starting from a larger base value), the percentage decreases in the other labour market categories are stronger for coloureds than for blacks. This explanation also holds for Asians and whites when compared to blacks.

Another result that becomes evident from the table is that the number of informal workers amongst whites and Asians exhibits an upward trend, especially for whites. Whites and Asians are less in demand than blacks (and coloureds) to fill the new formal jobs. This is due to their higher skill levels and the fact that they are already more likely to work in formal employment. Thus, the number of individuals entering the informal sector because of higher earnings is greater than the number of informal workers moving in the direction of the formal sector. Moreover, following the tightening gap between formal and informal earnings, there is a significant flow of white high-skilled workers moving from the formal to the informal sector, where earnings are not taxed. This phenomenon is the main contributor to the vigorous growth of informal white earnings. It also strengthens the rise in average informal earnings. However, at the country level, the latter increase is mainly caused by the trickle down effect of the formal sector development

[39] In order to present a table of a reasonable size, results from the second Keynesian closure have been omitted because they are similar in many respects to the results obtained with the Neoclassical closure. However, the decreases in poverty indicators are slightly more important under the Keynesian closure because the presence of the Keynesian multiplier effect implies more job creations.

[40] When interpreting poverty changes at the national level, it should be borne in mind that blacks account for more than 95% of all poor.

Table 6: Simulation Results from the MS model (percentage change from the base values as presented in tables 1 and 2)

	Keynes 1					Neoclassical				
	All	Blacks	Coloureds	Asians	Whites	All	Blacks	Coloureds	Asians	Whites
Inactive[(a)]	-0.16	-0.12	-0.28	-0.46	-0.36	-0.07	-0.04	-0.11	-0.19	-0.20
Unemployed	-0.91	-0.86	-1.39	-1.13	-1.35	-0.36	-0.33	-0.55	-0.62	-0.67
Subsistence agriculture	-0.23	-0.23	-0.18	-0.58	-0.26	-0.10	-0.10	-0.06	-0.58	-0.11
Informal workers	-0.25	-0.34	-0.70	0.30	2.64	-0.14	-0.14	-0.29	-0.02	0.36
Formal workers	1.21	1.54	1.23	1.02	0.36	0.51	0.61	0.49	0.46	0.28
Informal earnings[(b)]	3.94	2.36	2.61	2.24	7.34	0.97	0.69	0.74	0.58	1.64
Formal earnings[(b)]	-1.43	-1.71	-1.48	-0.97	-0.65	-0.13	-0.19	-0.11	0.01	0.08
Real income per capita[(c)]	0.69	0.61	0.21	0.43	0.91	0.65	0.63	0.49	0.57	0.72
P0[(d)]	-1.55	-1.54	-1.75	0.00	0.00	-0.95	-0.96	-0.71	0.00	0.00
P1[(d)]	-3.18	-3.05	-7.42	-12.26	-13.20	-1.57	-1.51	-3.49	-4.60	-4.59
P2[(d)]	-4.25	-4.08	-9.58	-28.08	-21.59	-2.03	-1.96	-4.64	-10.66	-9.73
Gini	-0.14	-0.42	-0.41	-0.19	0.36	-0.04	-0.11	-0.11	-0.08	0.12

Note: (a) Including 652,000 "unspecified" workers (b) Average real per capita earnings (c) Average real disposable income per capita in Rand (d) The poverty line is the international $2/day poverty line (R174/month/capita in 2000 prices).

on informal earnings. Indeed, the growth in formal employment and total formal earnings is indirectly responsible for the increase in informal per capita earnings.[41]

Under the Neoclassical closure, the labour market changes are relatively similar albeit smaller than with the first Keynesian closure. This is not surprising given the smaller changes observed at the macro level (see table 5). Here, the decrease in formal per capita earnings is less important because real formal earnings are fixed at the macro level for low-skilled and skilled workers (whereas they were flexible and decreasing under the first Keynesian closure). The only downward pressure on formal per capita earnings now arises from the fact that the new formal workers tend to have a lower skill level, and thus lower earnings, than the former formal workers. At the macro level, the expansion of the formal sector is smaller in terms of employment and total earnings, which is why informal per capita earnings progress to a lesser extent. However, combined with deflation and formal job creations, this is enough to generate a reduction in poverty incidence. That is the decrease in nominal formal earnings of low-skilled and skilled workers is not sufficient to offset the positive impacts of the formal employment growth and the falling consumer price index.

Regarding inequality, the results using either closure converge towards a small decline at the country level despite a visible worsening of the inter-group income distribution. The growth in real income per capita is systematically higher for whites (the richest racial group) than for the poorest groups (blacks and coloureds). However, the apparent increase in inter-group inequality is more than offset by the reductions in intra-group inequalities for blacks, coloureds and Asians due to the labour market changes.

The upward trend in inter-group inequality is mostly caused by the surge in real capital returns and to a lesser extent by the differential rates of growth in formal earnings by skill level. The distribution of capital income is extremely unequal, with a Gini coefficient of more than 0.95. Moreover, it is highly concentrated amongst whites, which explains the increase in inequality amongst this racial group. For whites, price effects are more important than labour market changes: the increase in inequality is driven by the fact that high-income households gain

[41] A secondary reason is that most of the workers leaving the informal sector are low-skilled or skilled workers (who are moving to the formal sector), which pushes up the average skill level and thus average earnings.

from trade liberalisation because of the increase in real capital returns and real earnings from high-skilled labour. In contrast, labour market changes are more important than price changes for blacks and coloureds (and to a lesser extent for Asians) since capital income and high-skilled workers' earnings account for a much smaller share of their incomes. The decrease in intra-group and total inequality is due to the fact that the poorest households are the main winners from the expansion in formal job creation because the latter is biased towards inactive and unemployed people with a low skill level.

6. Conclusion

In this paper, a CGE model and a MS model are combined in a sequential approach in order to assess the effects of trade liberalisation on South African households. Since there is no strong consensus emerging from the CGE literature regarding the most relevant closure to adopt when assessing trade liberalisation, three different closures with a short-run focus are used in this paper. The aim was to choose some closures either based upon economic theory, CGE literature or backed up by empirical evidence. One Neoclassical and two Keynesian macroeconomic closures of the CGE model are considered. Although the magnitude of the results can vary depending on the closure, a clear pattern emerges regarding the direction of the effects. Trade liberalisation appears to be pro-poor and to have a limited dampening effect on inequality. No matter the closure, the positive impact of formal job creations is always the main force leading to a decrease in poverty. Regarding inequality, the Gini coefficient shows a decreasing trend, independent of the closure used in the CGE model. There is an apparent increase in inter-group inequality, which is due to the fact that the two production factors showing the highest growth in their returns are skilled labour and capital. These two factors are the two main income sources of high-income (mainly white) households. However, the decrease in intra-group inequalities driven by the labour market changes is not completely offset by the more unequal income distribution due to increasing inter-group inequality. The results show that labour market changes are the most important factor placing a downward pressure on inequality and poverty while the increase in capital earnings tends to increase income inequality.

Finally, the sensitivity analyses carried out in this paper reveal that the results are 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.

In conclusion, the combination of a micro- and a macro-model allows the assessment of macro-policies beyond their economy-wide effects. Even though some micro-impacts can be inferred from the macro-results, the use of a micro-model has proved to be useful in order to evaluate more precisely these impacts as well as to carry out some disaggregated analyses. In this paper, the microeconomic effects of trade liberalisation are found to be quite diverse depending on the racial groups. Although whites emerge as the main winners, the poor are not left behind because they benefit from the creation of new formal jobs. Some preliminary results from a more disaggregated analysis, allowing for the distinction between rural and urban areas at the provincial level, also reveal some different impacts for rural versus urban areas. Poor urban households seem to benefit more from trade liberalisation than poor rural households.

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 P. Posel (2004). "Two millions net new jobs": A reconsideration of the rise in employment in South Africa, 1995-2003. University of KwaZulu-Natal.

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.

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.

Orcutt, G.H. (1957). A new type of socio -economic system. *Review of Economics and Statistics*, 58, 773-797.

Orcutt, G.H., S. Caldwell and R. Wertheimer (1976). Policy Explorations Through Microanalytic Simulation. The Urban Institute, Washington D.C.

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: SELECTION MODELS

Table A.1: Estimated parameters of the selection models for single women

	Unemployed		Subsistence agriculture		Informal sector		Formal sector	
	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level
Coloured	0.239	0.360	-0.508	0.530	-0.466	0.071	1.003	0.000
Asian	-0.222	0.565	(dropped)	0.000	-1.699	0.001	0.497	0.144
White	-0.330	0.324	(dropped)	0.000	-2.654	0.000	0.558	0.056
Primary education	-0.070	0.664	0.515	0.212	0.069	0.629	0.099	0.554
Standard 8	0.014	0.931	-0.033	0.940	-0.170	0.258	0.379	0.028
Standard 10	0.429	0.010	-1.070	0.013	-0.327	0.040	0.600	0.001
Tertiary education	0.186	0.519	(dropped)	0.000	-6.260	0.000	0.406	0.131
Education attendance	-3.371	0.000	-1.336	0.000	-2.094	0.000	-2.238	0.000
Semi-skilled	0.164	0.012	3.476	0.000	-0.052	0.494	0.601	0.000
High-skilled	-0.791	0.061	(dropped)	0.000	-0.603	0.174	1.165	0.000
Age	0.277	0.000	0.023	0.629	0.385	0.000	0.407	0.000
Age squared/100	-0.421	0.000	-0.063	0.326	-0.462	0.000	-0.506	0.000
Ability to write	0.439	0.006	-0.370	0.381	0.347	0.013	0.178	0.276
Afrikaans language	-0.624	0.017	0.039	0.962	0.214	0.404	-0.285	0.274
English language	-0.840	0.005	(dropped)	0.000	-0.224	0.441	-0.662	0.016
Agri. Activity	0.220	0.017	3.803	0.000	1.075	0.000	0.192	0.063
Western Cape	0.293	0.027	1.607	0.042	0.755	0.000	0.849	0.000
Eastern Cape	-0.150	0.130	-0.382	0.495	0.082	0.476	-0.185	0.098
Northern Cape	-0.172	0.256	0.917	0.324	0.299	0.080	-0.207	0.218
Freestate	-0.024	0.829	0.620	0.275	-0.155	0.235	0.444	0.000
Kwazulu-Natal	0.035	0.697	0.603	0.263	0.250	0.015	0.350	0.000
North-West	-0.194	0.051	-1.766	0.047	-0.158	0.178	-0.253	0.026
Mpumalanga	0.240	0.026	0.195	0.733	0.359	0.004	0.467	0.000
Limpopo	-0.238	0.043	-1.414	0.016	-0.262	0.050	0.062	0.631
Urban location	0.648	0.000	-0.360	0.181	0.279	0.000	0.072	0.315
Homeland location	-0.206	0.004	0.446	0.031	-0.330	0.000	-0.425	0.000
Children (number)	-0.084	0.001	-0.060	0.356	0.188	0.000	0.085	0.005
Household size	0.047	0.000	0.021	0.587	-0.168	0.000	-0.101	0.000
Children 0-1 yr old	-0.297	0.005	0.248	0.454	-0.793	0.000	-0.735	0.000
Children 1-3 yrs old	0.090	0.312	0.192	0.504	-0.531	0.000	-0.457	0.000
Children 4-5 yrs old	-0.229	0.023	-0.027	0.932	-0.556	0.000	-0.431	0.000
Children 6-9 yrs old	-0.120	0.195	0.372	0.197	-0.516	0.000	-0.394	0.000
Children 9-15 yrs old	-0.009	0.927	-0.405	0.230	-0.365	0.000	-0.265	0.008
Disability grant in the hh	0.177	0.076	-0.702	0.036	-0.550	0.001	-0.269	0.043
OAP in the hh	-0.084	0.157	0.201	0.218	-0.723	0.000	-0.493	0.000
Family allow. in the hh	0.289	0.006	-0.614	0.084	0.090	0.499	-0.076	0.567
Capital income**	-0.002	0.578	-0.073	0.126	0.003	0.243	-0.009	0.074
Gov. transfers to the hh**	0.009	0.085	-0.002	0.852	-0.072	0.002	-0.019	0.124
Predicted inactive inc.*	4.57E-05	0.000	4.57E-05	0.000	(dropped)	0.000	(dropped)	0.000
Predic. informal earnings*	(dropped)	0.000	(dropped)	0.000	4.57E-05	0.000	(dropped)	0.000
Predic. formal earnings*	(dropped)	0.000	(dropped)	0.000	(dropped)	0.000	4.57E-05	0.000
Constant	-5.267	0.000	-6.575	0.000	-7.127	0.000	-8.400	0.000

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

Table A.2: Estimated parameters of the selection models for married women

	Unemployed		Subsistence agriculture		Informal sector		Formal sector	
	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level
Coloured	-0.618	0.022	1.093	0.365	-0.738	0.001	0.141	0.514
Asian	-1.440	0.000	(dropped)	0.000	-2.312	0.000	-0.718	0.004
White	-1.267	0.000	0.533	0.669	-2.609	0.000	-0.917	0.000
Primary education	0.113	0.530	1.643	0.000	0.172	0.202	-0.077	0.657
Standard 8	0.228	0.226	1.122	0.013	-0.072	0.626	0.089	0.617
Standard 10	0.353	0.077	-0.549	0.271	-0.183	0.269	0.355	0.058
Tertiary education	0.203	0.587	0.161	0.892	-1.679	0.000	0.523	0.043
Education attendance	-1.401	0.000	-0.435	0.359	-0.712	0.001	-0.267	0.065
Semi-skilled	0.107	0.210	4.740	0.000	0.160	0.041	0.535	0.000
High-skilled	0.482	0.113	(dropped)	0.000	0.701	0.013	0.872	0.000
Age	0.164	0.000	0.065	0.291	0.180	0.000	0.275	0.000
Age squared/100	-0.266	0.000	-0.095	0.208	-0.207	0.000	-0.339	0.000
Ability to write	0.001	0.996	-1.664	0.000	-0.102	0.445	0.186	0.292
Afrikaans language	-0.401	0.131	-2.596	0.035	0.220	0.320	0.176	0.411
English language	-0.563	0.060	-2.584	0.031	-0.018	0.945	-0.180	0.418
Agri. Activity	-0.303	0.005	2.737	0.000	0.740	0.000	-0.150	0.128
Western Cape	-0.466	0.004	1.233	0.081	0.015	0.918	0.314	0.008
Eastern Cape	-0.476	0.000	0.441	0.403	-0.360	0.004	0.016	0.889
Northern Cape	-0.418	0.025	0.090	0.923	0.026	0.870	-0.228	0.117
Freestate	-0.140	0.259	0.610	0.258	-0.021	0.862	0.277	0.021
Kwazulu-Natal	-0.357	0.003	0.616	0.227	-0.207	0.072	0.172	0.107
North-West	-0.401	0.001	-0.195	0.741	-0.483	0.000	-0.193	0.097
Mpumalanga	0.101	0.418	0.686	0.191	-0.036	0.772	0.067	0.599
Limpopo	-0.331	0.016	-0.751	0.168	-0.633	0.000	0.045	0.726
Urban location	0.618	0.000	-1.148	0.000	-0.058	0.421	-0.090	0.212
Homeland location	-0.155	0.101	-0.437	0.047	-0.333	0.000	-0.569	0.000
Children (number)	-0.152	0.000	-0.151	0.077	-0.073	0.037	-0.190	0.000
Household size	0.079	0.000	-0.022	0.675	0.013	0.508	0.020	0.341
Children 0-1 yr old	-0.406	0.003	0.198	0.574	-0.569	0.000	-0.077	0.548
Children 1-3 yrs old	-0.039	0.730	-0.026	0.933	-0.340	0.001	0.011	0.916
Children 4-5 yrs old	-0.149	0.239	-0.116	0.721	-0.333	0.004	-0.042	0.716
Children 6-9 yrs old	0.056	0.631	0.385	0.197	-0.119	0.244	0.243	0.018
Children 9-15 yrs old	0.054	0.662	-0.161	0.605	-0.157	0.137	0.025	0.807
OAP in the hh	0.082	0.477	-0.293	0.229	-0.465	0.000	-0.133	0.230
Family allow. in the hh	0.241	0.218	0.563	0.131	0.223	0.190	-0.086	0.688
Capital income**	-0.014	0.029	-0.169	0.001	-0.002	0.024	-0.005	0.002
Gov. transfers to the hh**	-0.031	0.049	0.008	0.811	-0.020	0.172	-0.065	0.000
Partner live here	0.840	0.002	0.893	0.152	0.500	0.036	1.190	0.000
Primary education (partner)	0.195	0.118	-0.015	0.951	-0.014	0.875	0.027	0.815
Standard 8 (partner)	0.147	0.268	-0.240	0.424	-0.239	0.025	0.000	0.999
Standard 10 (partner)	0.185	0.205	0.174	0.647	-0.580	0.000	-0.064	0.630
Tertiary education (partner)	-0.084	0.760	1.297	0.113	-0.879	0.006	-0.124	0.489
Age (partner)	-0.019	0.000	-0.013	0.254	-0.014	0.002	-0.026	0.000
Semi-skilled (partner)	-0.226	0.220	-0.767	0.242	-0.031	0.865	0.169	0.134
High-skilled (partner)	0.048	0.553	0.255	0.185	0.307	0.000	0.062	0.382
Predicted inactive inc.*	4.08E-05	0.000	4.08E-05	0.000	(dropped)	0.000	(dropped)	0.000
Predic. informal earnings*	(dropped)	0.000	(dropped)	0.000	4.08E-05	0.000	(dropped)	0.000
Predic. formal earnings*	(dropped)	0.000	(dropped)	0.000	(dropped)	0.000	4.08E-05	0.000
Constant	-3.394	0.000	-7.189	0.000	-3.675	0.000	-6.719	0.000

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

Table A.3: Estimated parameters of the selection models for single men

	Unemployed		Subsistence agriculture		Informal sector		Formal sector	
	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level
Coloured	-0.389	0.144	-0.978	0.127	-0.263	0.421	0.125	0.615
Asian	-0.176	0.644	-0.448	0.667	-1.048	0.095	-0.057	0.867
White	-1.117	0.000	-0.847	0.247	-2.287	0.000	-0.790	0.005
Primary education	-0.121	0.460	0.171	0.669	0.030	0.876	-0.128	0.452
Standard 8	0.156	0.356	-0.183	0.669	-0.024	0.908	-0.194	0.276
Standard 10	0.351	0.041	-1.972	0.000	-0.583	0.005	0.020	0.912
Tertiary education	0.687	0.012	-1.744	0.039	-1.604	0.000	-0.118	0.664
Education attendance	-3.929	0.000	-1.728	0.000	-2.562	0.000	-3.206	0.000
Semi-skilled	0.154	0.037	3.885	0.000	1.256	0.000	0.595	0.000
High-skilled	0.050	0.875	(dropped)	0.000	1.350	0.000	0.810	0.002
Age	0.216	0.000	0.039	0.317	0.231	0.000	0.294	0.000
Age squared/100	-0.324	0.000	-0.056	0.276	-0.287	0.000	-0.396	0.000
Ability to write	0.632	0.000	-0.075	0.839	0.196	0.285	0.357	0.029
Afrikaans language	0.200	0.450	1.364	0.029	0.882	0.007	0.515	0.039
English language	-0.078	0.800	(dropped)	0.000	-0.519	0.220	0.318	0.243
Agri. Activity	0.102	0.323	2.546	0.000	1.100	0.000	-0.426	0.001
Western Cape	0.033	0.811	0.816	0.120	0.142	0.435	0.666	0.000
Eastern Cape	-0.098	0.352	0.261	0.504	-0.175	0.229	-0.215	0.066
Northern Cape	0.224	0.175	1.392	0.009	-0.081	0.714	0.414	0.012
Freestate	0.178	0.131	1.047	0.008	0.018	0.914	-0.138	0.296
Kwazulu-Natal	0.058	0.541	0.576	0.125	-0.142	0.282	0.069	0.503
North-West	-0.126	0.230	0.183	0.663	-0.377	0.009	-0.389	0.001
Mpumalanga	0.419	0.000	0.504	0.226	0.496	0.002	0.526	0.000
Limpopo	-0.135	0.275	-0.276	0.520	-0.407	0.015	-0.023	0.869
Urban location	0.336	0.000	-0.005	0.982	-0.148	0.104	-0.502	0.000
Homeland location	-0.307	0.000	0.238	0.228	-0.201	0.048	-1.156	0.000
Household size	-0.022	0.253	-0.012	0.829	-0.295	0.000	-0.246	0.000
Number of male	0.045	0.067	-0.030	0.656	0.054	0.173	0.068	0.026
Children (number)	-0.025	0.379	0.026	0.732	0.256	0.000	0.182	0.000
Children 0-1 yr old	0.019	0.888	0.365	0.270	-0.013	0.946	-0.104	0.498
Children 1-3 yrs old	0.019	0.850	0.234	0.382	-0.383	0.011	-0.514	0.000
Children 4-5 yrs old	-0.051	0.642	0.159	0.580	-0.391	0.015	-0.516	0.000
Children 6-9 yrs old	-0.195	0.036	0.041	0.865	-0.630	0.000	-0.774	0.000
Children 9-15 yrs old	-0.095	0.298	-0.549	0.036	-0.609	0.000	-0.493	0.000
Disability grant in the hh	0.174	0.124	0.544	0.067	0.445	0.017	-0.107	0.472
OAP in the hh	-0.105	0.104	0.400	0.051	-0.164	0.197	-0.461	0.000
Family allow. in the hh	0.175	0.160	0.411	0.216	0.326	0.085	-0.233	0.168
Capital income**	-0.007	0.198	-0.024	0.390	0.000	0.783	-0.013	0.003
Gov. transfers to the hh**	0.008	0.521	-0.105	0.155	-0.253	0.000	-0.142	0.000
Predicted inactive inc.*	4.90E-05	0.000	4.90E-05	0.000	(dropped)	0.000	(dropped)	0.000
Predic. informal earnings*	(dropped)	0.000	(dropped)	0.000	4.90E-05	0.000	(dropped)	0.000
Predic. formal earnings*	(dropped)	0.000	(dropped)	0.000	(dropped)	0.000	4.90E-05	0.000
Constant	-3.821	0.000	-5.725	0.000	-3.964	0.000	-4.019	0.000

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

Table A.4: Estimated parameters of the selection models for married men

	Unemployed		Subsistence agriculture		Informal sector		Formal sector	
	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level	Estimated coefficient	Significance level
Coloured	0.250	0.565	1.578	0.230	0.314	0.390	0.471	0.142
Asian	0.359	0.517	(dropped)	0.000	1.162	0.018	0.606	0.133
White	-1.045	0.039	-0.494	0.748	-1.118	0.006	-0.226	0.499
Primary education	0.330	0.152	-0.089	0.818	-0.089	0.627	-0.090	0.592
Standard 8	0.400	0.106	-0.552	0.212	-0.307	0.127	-0.086	0.641
Standard 10	0.422	0.126	-0.780	0.090	-0.653	0.005	0.094	0.651
Tertiary education	0.208	0.643	(dropped)	0.000	-1.718	0.000	-0.244	0.401
Education attendance	-0.719	0.113	(dropped)	0.000	-1.703	0.000	0.036	0.899
Semi-skilled	0.084	0.527	4.725	0.000	0.934	0.000	0.450	0.000
High-skilled	-0.233	0.424	(dropped)	0.000	-0.074	0.753	0.419	0.024
Age	0.309	0.000	0.147	0.043	0.164	0.000	0.271	0.000
Age squared/100	-0.395	0.000	-0.211	0.008	-0.219	0.000	-0.341	0.000
Ability to write	0.014	0.953	-0.316	0.423	0.052	0.777	0.206	0.227
Afrikaans language	-0.749	0.082	-1.177	0.385	-0.269	0.453	-0.233	0.456
English language	-0.405	0.404	-1.725	0.234	-1.493	0.000	-0.239	0.487
Agri. Activity	-0.245	0.119	2.055	0.000	0.512	0.000	-0.306	0.008
Western Cape	-0.405	0.118	0.417	0.551	-0.131	0.554	0.143	0.461
Eastern Cape	-0.403	0.047	-0.334	0.530	-0.393	0.027	-0.379	0.017
Northern Cape	0.118	0.697	-1.542	0.187	-0.162	0.545	0.170	0.471
Freestate	-0.039	0.861	0.752	0.150	-0.187	0.359	0.431	0.017
Kwazulu-Natal	-0.236	0.214	-0.341	0.509	-0.480	0.005	-0.216	0.153
North-West	-0.121	0.569	-0.327	0.567	-0.395	0.039	0.063	0.711
Mpumalanga	-0.008	0.973	-0.336	0.567	0.030	0.880	0.521	0.004
Limpopo	-0.242	0.287	-0.549	0.318	-0.267	0.171	0.041	0.817
Urban location	0.133	0.318	-0.845	0.004	-0.425	0.000	-0.572	0.000
Homeland location	-0.525	0.000	-0.412	0.149	-0.521	0.000	-1.264	0.000
Household size	0.009	0.828	-0.152	0.074	-0.148	0.000	-0.174	0.000
Number of male	0.036	0.503	0.074	0.488	0.057	0.237	0.060	0.161
Children (number)	-0.129	0.030	-0.075	0.513	0.040	0.452	0.058	0.218
Children 0-1 yr old	0.559	0.019	1.000	0.038	0.305	0.154	0.426	0.028
Children 1-3 yrs old	0.376	0.049	0.721	0.068	0.168	0.319	0.164	0.280
Children 4-5 yrs old	0.358	0.095	0.960	0.023	0.148	0.432	0.197	0.243
Children 6-9 yrs old	0.213	0.260	0.631	0.096	0.058	0.722	0.052	0.721
Children 9-15 yrs old	0.210	0.278	0.684	0.069	0.089	0.589	0.061	0.674
OAP in the hh	-0.067	0.671	0.192	0.533	-0.564	0.000	-0.627	0.000
Family allow. in the hh	0.129	0.604	0.611	0.147	-0.169	0.462	-0.681	0.002
Capital income**	0.000	0.726	-0.001	0.523	-0.001	0.169	-0.005	0.000
Gov. transfers to the hh**	-0.001	0.936	-0.029	0.578	-0.132	0.000	-0.120	0.000
Partner live here	1.484	0.001	0.214	0.787	1.440	0.000	0.929	0.005
Primary education (partner)	0.207	0.274	-0.011	0.972	-0.029	0.844	0.128	0.356
Standard 8 (partner)	0.021	0.919	-0.356	0.330	-0.279	0.095	-0.118	0.447
Standard 10 (partner)	0.093	0.707	-0.774	0.104	-0.292	0.160	0.101	0.592
Tertiary education (partner)	0.308	0.588	-0.368	0.814	0.284	0.535	0.750	0.049
Age (partner)	-0.034	0.000	-0.015	0.350	-0.030	0.000	-0.034	0.000
Semi-skilled (partner)	-0.507	0.232	-0.911	0.552	-0.595	0.092	-0.792	0.005
High-skilled (partner)	-0.261	0.066	0.140	0.552	-0.063	0.602	-0.015	0.890
Predicted inactive inc.*	2.04E-05	0.000	2.04E-05	0.000	(dropped)	0.000	(dropped)	0.000
Predic. informal earnings*	(dropped)	0.000	(dropped)	0.000	2.04E-05	0.000	(dropped)	0.000
Predic. formal earnings*	(dropped)	0.000	(dropped)	0.000	(dropped)	0.000	2.04E-05	0.000
Constant	-5.737	0.000	-5.795	0.001	-1.145	0.121	-2.012	0.002

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

APPENDIX B: REGRESSION MODELS

Table B.1: Estimated parameters of the regression models for women (log earnings)

	Single women				Married women			
	Informal earnings		Formal earnings		Informal earnings		Formal earnings	
	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	0.941
Asian	1.244	0.003	0.028	0.870	1.735	0.000	0.006	0.957
White	1.724	0.000	0.335	0.021	2.185	0.000	0.347	0.001
Primary education	0.016	0.870	-0.097	0.321	-0.007	0.956	-0.282	0.003
Standard 8	0.253	0.020	0.077	0.444	0.519	0.000	-0.017	0.863
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	-0.262	0.000	0.285	0.000
High-skilled	0.918	0.001	0.562	0.000	0.103	0.650	0.576	0.000
Age	0.029	0.080	0.064	0.000	0.045	0.060	0.044	0.005
Age squared/100	-0.033	0.117	-0.051	0.001	-0.055	0.060	-0.033	0.093
Ability to write	0.052	0.602	0.282	0.003	0.256	0.029	0.606	0.000
Afrikaans language	-0.153	0.397	0.092	0.504	-0.185	0.346	0.045	0.672
English language	0.055	0.761	0.295	0.035	-0.229	0.365	0.169	0.111
Western Cape	0.028	0.793	-0.122	0.053	0.047	0.737	-0.090	0.117
Eastern Cape	-0.481	0.000	-0.113	0.056	-0.349	0.003	-0.171	0.003
Northern Cape	-0.478	0.000	-0.131	0.115	-0.497	0.001	-0.113	0.118
Freestate	-0.529	0.000	-0.418	0.000	-0.670	0.000	-0.377	0.000
Kwazulu-Natal	-0.237	0.002	-0.239	0.000	-0.101	0.340	-0.098	0.061
North-West	-0.145	0.117	-0.056	0.361	0.087	0.460	-0.021	0.729
Mpumalanga	-0.275	0.003	-0.206	0.002	-0.022	0.848	-0.081	0.223
Limpopo	0.104	0.307	-0.225	0.001	0.227	0.069	-0.168	0.010
Urban location	0.160	0.004	0.292	0.000	0.319	0.000	0.386	0.000
Homeland location	-0.071	0.267	0.169	0.000	0.090	0.262	0.251	0.000
Agri. Activity	-1.078	0.000	-0.225	0.000	-0.991	0.000	0.049	0.408
Inverse Mills ratios	-0.811	0.000	-0.404	0.000	-1.184	0.000	-0.368	0.000
Constant	8.939	0.000	7.842	0.000	8.717	0.000	7.964	0.000

Table B.2: Estimated parameters of the regression models for men (log earnings)

	Single men				Married men			
	Informal earnings		Formal earnings		Informal earnings		Formal earnings	
	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	0.832
Asian	0.821	0.147	0.265	0.046	0.154	0.642	0.092	0.262
White	1.523	0.000	0.710	0.000	1.174	0.000	0.721	0.000
Primary education	-0.252	0.119	-0.006	0.946	-0.019	0.863	0.080	0.074
Standard 8	-0.176	0.310	0.206	0.016	0.313	0.012	0.300	0.000
Standard 10	0.305	0.094	0.460	0.000	0.674	0.000	0.575	0.000
Tertiary education	0.717	0.119	0.682	0.000	0.928	0.002	0.824	0.000
Education attendance	0.037	0.826	0.307	0.000	0.658	0.105	0.301	0.000
Semi-skilled	-0.750	0.000	0.133	0.000	-0.341	0.000	0.165	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	0.014	0.090	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	0.236	0.036	0.201	0.000
Afrikaans language	-0.153	0.588	0.088	0.395	-0.071	0.726	-0.015	0.825
English language	0.512	0.131	0.098	0.372	0.406	0.131	0.169	0.020
Western Cape	0.204	0.244	-0.162	0.003	-0.059	0.685	-0.046	0.231
Eastern Cape	-0.526	0.000	-0.169	0.002	-0.483	0.000	-0.161	0.000
Northern Cape	0.049	0.812	-0.360	0.000	-0.301	0.077	-0.054	0.236
Freestate	-0.516	0.001	-0.226	0.000	-0.536	0.000	-0.135	0.000
Kwazulu-Natal	0.195	0.102	-0.091	0.043	-0.138	0.230	-0.083	0.015
North-West	0.200	0.144	0.153	0.003	-0.015	0.905	0.086	0.017
Mpumalanga	-0.102	0.467	-0.085	0.140	-0.048	0.696	-0.036	0.348
Limpopo	0.351	0.021	-0.273	0.000	-0.084	0.506	-0.033	0.438
Urban location	0.297	0.000	0.387	0.000	0.242	0.001	0.330	0.000
Homeland location	-0.349	0.000	0.192	0.000	-0.088	0.326	0.205	0.000
Agri. Activity	-1.166	0.000	0.083	0.263	-0.551	0.000	-0.068	0.056
Inverse Mills ratios	-0.787	0.000	-0.271	0.000	-0.267	0.200	-0.537	0.000
Constant	9.024	0.000	7.482	0.000	8.145	0.000	7.486	0.000

APPENDIX C: SUMMARY STATISTICS

Table C.1: 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