
Tax Policies and Informality in South Africa

Document de Travail
Working Paper
2015-22

Eliane El Badaoui
Riccardo Magnani



UMR 7235

Université de Paris Ouest Nanterre La Défense
(bâtiment G)
200, Avenue de la République
92001 NANTERRE CEDEX

Tél et Fax : 33.(0)1.40.97.59.07
Email : nasam.zaroualete@u-paris10.fr

université
Paris Ovest

Nanterre La Défense

Tax Policies and Informality in South Africa*

Eliane El Badaoui[†]

Riccardo Magnani[‡]

Abstract

We use a micro-macro simulation model to evaluate the effects of labor income tax policies in South Africa. The country is characterized by a high unemployment rate while employment in the informal sector is relatively low. Our approach is based on the aggregation of the preferences of individuals who choose among (i) working in the formal sector, (ii) working in the informal sector, and (iii) not working. We quantify the effects of different tax policies on the individual's labor supply choices (formal/informal employment, work/leisure) and at the macro level (GDP, equilibrium wages, size of the informal sector, and unemployment rate). We further analyze the effects in terms of income distribution, poverty and inequality. We find that the introduction of a negative income tax system reduces inequality and poverty but, at the same time, discourages people from participating in the labor market and working in the formal sector and, consequently worsens the equity-efficiency trade-off. We find however that the equity-efficiency trade-off can be improved by introducing (i) a (revenue-neutral) tax system which combines a sufficiently high flat rate (25%) and a lump-sum transfer paid only to formal workers; (ii) an expansionary policy (a reduction in total taxes or an increase in public expenditures). Interestingly, even though they increase the informal sector, these reforms reduce inequality and poverty through a reduction in the level of unemployment.

JEL Classification: J22, H24, O17, C68

Key-words: Tax Reform, Informal Sector, Labor Supply, CGE Model, Microsimulation, South Africa

*This research received financial support from the French National Research Agency, under the grant INEG_CONSO (ANR-11-INEG-004-01).

[†]EconomiX - Université Paris Ouest Nanterre La Défense, 200 Avenue de la République, 92001 Nanterre, France. E-mail: eliane.badaoui@u-paris10.fr. Tel: +33 1 40 97 78 25.

[‡]CEPN - Université de Paris 13, 99 Avenue Jean-Baptiste Clément, 93430 Villetaneuse, France. E-mail: riccardo.magnani@univ-paris13.fr. Tel: +33 1 49 40 33 54.

1 Introduction

Unlike the advanced countries, the developing countries have a large informal sector that fails to comply with tax authorities. Developing countries face difficult challenges when they attempt to establish efficient tax systems, given their weak direct and indirect tax capacity. The tax revenue-GDP ratio is a good measure of direct taxation efficiency, and this ratio is usually higher in developed countries than in poor countries (McLaren, 1998).¹ In this context, it is of interest to study how a tax change modifies the behavior of persons in the labor market, the level of poverty and inequality and the amount of government revenues.

The fundamental distinction between the formal and the informal sector is related to tax payment, where the informal sector is the tax-free option. Informality has a direct impact on public revenues, and thus, it accounts for fiscal policy analysis. Informality was traditionally viewed as a low-productive and low-paid activity, i.e., an option for workers who cannot find a formal job (Fields, 1975; Hart, 1973). However, recent economic analysis suggests that informal employment may be voluntary. Maloney (1999) considers that workers self-select into the informal sector because of the various benefits and opportunities that it can offer, implying the existence of comparative advantages in the informal sector. Subject to job availability, workers choose the sector of activity that maximizes their utility. Obviously, wages are important determinants of the utility. However, non-monetary advantages as autonomy, flexibility, distance to work and working hours also exist, which affect utility, determine job satisfaction and make informal employment desirable.² Thus, depending on their characteristics, certain workers choose the informal sector (Günther and Launov, 2012). Using data from Argentina, Pratap and Quintin (2006) conclude the absence of a formal sector wage premium, and thus, reject the segmentation hypothesis in the labor market.³ The authors further conclude on the

¹For instance, this ratio is equal to 37.9% for OECD countries and 18.2% for developing countries over the period 1995-1997 (Tanzi and Zee, 2000). Tanzi (1987) focuses on the high reliance of low-income countries on indirect taxation (61% of total tax revenue on average), whereas indirect taxation represents 36.4% of total tax revenue in high-income countries.

²See, e.g., Mulinge and Mueller (1998) and Saavedra and Chong (1999).

³See Fields (1980) for a review of the literature. In contrast, Maloney (2004) argues that significant formal-

voluntary aspect of the transitions from one sector to another. [El Badaoui *et al.* \(2008\)](#) find similar results for South Africa. Evidence from Mexico ([Gong and van Soest, 2002](#); [Gong *et al.*, 2004](#)) is also consistent with such a view of the informal activity as an attractive option.

Tax policy modifies the relative (formal/informal) wage, and thus, plays a significant role by affecting the individual's labor supply decision ([Fugazza and Jacques, 2004](#); [Johnson *et al.*, 1998](#)) and the size of the informal sector in the economy. In the literature, several studies analyzed the link between the informal sector and taxation. According to [Gutmann \(1977\)](#), higher taxes increase the size of the informal economy. [Loayza \(1996\)](#) shows that the informal sector arises when governments with weak enforcement capacity impose excessive taxes and regulations. [Ihrig and Moe \(2004\)](#) propose a simple dynamic model that captures the negative and convex relationship between the size of the informal sector and the per capita real GDP. Moreover, the study shows that a reduction in tax rates attracts people out of the informal sector, and thus, improves the standard of living in the economy. The authors show for Sri Lanka that raising the tax rate from 9.3% to 10% and holding enforcement constant, generates a 2.3% increase in the informal employment in steady state. [Saracoglu \(2008\)](#) proposes a dynamic general equilibrium model with heterogeneous goods and an endogenous price for the informal sector good. The model shows that the informal sector diminishes over time as the economy grows and, more importantly, a lower tax on employment in the formal sector reduces the size of the informal sector. For high-income countries, [Fugazza and Jacques \(2004\)](#) show that lower taxation stimulates participation in the formal sector. In [Lemieux *et al.* \(1994\)](#), an increase in the tax rate drives reallocation of labor from the formal to the informal sector. Using a data set collected in Québec City in Canada, they found that this relationship is significant for particular groups of the population, such as social-welfare claimants.

In this paper, we analyze the economic consequences of several reforms in South Africa, both at the macro and micro levels. We first simulate a (revenue-neutral) negative income tax

informal wage differentials are not necessarily proof of segmentation.

that combines a flat rate equal to 20% with a lump-sum transfer. Second, we analyze the effects of different tax policies that combine a flat rate and a transfer paid only to formal workers. Finally, we consider two expansionary fiscal policies that consist of reducing the total amount of taxes and increasing public expenditures, respectively.

South Africa is an interesting case study. First, by international standards, South Africa has an exceptionally low rate of informal employment and an exceptionally high rate of unemployment (Kingdon and Knight, 2004). According to the 2000 Labour Force Survey, 24.4% of employed workers operate in the informal sector in South Africa, and the official unemployment rate is 23.3%. Agreement is widespread that the main factor responsible for such a situation in the labor market is the existence of barriers to entry in the informal sector.⁴ Capital and experience are cited as common barriers to entrepreneurship (Maloney, 2004) and lead to lower levels of informal salaried employment. South Africa, in contrast to other countries, further suffers from the repercussions of the apartheid system that repressed the informal activities of Black South Africans and inhibited the development of entrepreneurial and social skills. Second, the South African government has been particularly successful in collecting direct taxes, through a progressive tax system, in the form of corporate and personal income taxes. The tax represented 24% of the South African GDP in 2000, the highest among middle-income countries for which the average ratio is 15% over the period 1997-2002.⁵

Our analysis is carried out using a micro-macro simulation approach that combines a Microsimulation model and a CGE (Computable General Equilibrium) model. The integration of these two types of models, which are widely used in policy analysis, allows to evaluate the individual and the macro effects by taking into account both the individual heterogeneity present in a micro dataset and the general equilibrium effects produced at the macro level. In particular, we use a new micro-macro simulation approach developed by Magnani and Mercenier (2009) based on the exact aggregation theory of Anderson *et al.* (1992), which permits to aggregate

⁴See, e.g., Kingdon and Knight (2004) and Davies and Thurlow (2009).

⁵World Bank: <http://data.worldbank.org/indicator/GC.TAX.TOTL.GD.ZS?page=2>.

the preferences of individuals who make discrete choices into explicit labor supply functions that we introduce into a CGE model. The CGE model is subsequently used to evaluate the macro effects of different tax policies that are later introduced into the Microsimulation model to evaluate the effects at the individual level and on income distribution, inequality and poverty. In our paper, we assume that individuals face a discrete choice problem and choose among three alternatives: working in the formal sector, working in the informal sector, and not working.⁶ We assume that individuals compare the costs and benefits of each alternative and choose the option that maximizes their utility given their preferences and job characteristics.

We further distinguish between voluntary and involuntary unemployment. The voluntarily unemployed, as inactives, have high reservation wages and choose not to work. As previously mentioned, unemployment is high in South Africa, whereas informal employment is relatively low, and the unemployed are likely to be involuntary (Kingdon and Knight, 2004) because of the existence of barriers to entry in the informal sector. In this particular situation in the labor market, we assume that the involuntarily unemployed are better off in informal jobs and that the weakness in aggregate demand prevents some individuals from becoming entrepreneurs. Consequently, at the country level, entrepreneurship is poorly developed which limits the number of informal vacancies. The weakness in aggregate demand is thus responsible for the high level of unemployment and the low level of informality.⁷ In contrast, an increase in the aggregate demand increases the expected revenue compared with the high entry costs, which induces certain individuals to become entrepreneurs in the informal sector and to create vacancies. As a consequence, involuntary unemployment decreases and employment in the informal sector increases.

⁶Self-employment is not considered an option for individuals, and the underlying reasons are developed in the following sections.

⁷The weakness in aggregate demand is not the only factor that could be involved in explaining the small informal salaried sector and the high level of unemployment. A number of other factors exist and are involved, e.g., spatial separation of potential workers from employers, local government regulations that discourage self-employment (Skinner, 2008), unwillingness of some unemployed to take up jobs that they view as beneath them, and availability of government support through children's allowances and "granny grants". In contrast, the rigidity of formal wages cannot be considered as a determinant of unemployment since empirical studies conclude on the flexibility of wages in South Africa (Fourie, 2011; Kingdon and Knight, 2006).

The paper is organized as follows. In Section 2, we specify the individual discrete choice problem and present the microeconomic results. The CGE model used in this paper is presented in Section 3. Section 4 offers a short description of the South African tax system. In Section 5, we analyze the micro and macro effects of the simulated reforms, and the conclusions are given in Section 6.

2 The Individual Discrete Choice Problem

2.1 Model Specification

In this paper, we estimate a discrete choice labor supply model in which individuals are assumed to choose among three options of economic activity: (1) to work in the formal sector, (2) to work in the informal sector, and (3) not to work. We specify a nested logit model (McFadden, 1981) with two nests B_k .

The decision concerning labor market status is based on utility comparisons. Individual preferences are described by the following utility function:

$$U_{hjk} = V_{hjk} + Z_{hk} + \epsilon_{hjk} \quad (1)$$

where h indexes the individuals, $k \in \{l, L\}$ denotes the labor (l) and the leisure (L) status, and $j \in \{1, 2, 3\}$ indexes formal/informal employment activity and not working, respectively. Thus, each labor market option is uniquely identified by a double index jk . Z_{hk} represents the upper nest specific component, V_{hjk} is the alternative specific component, and ϵ_{hjk} is the error term that is assumed to follow a generalized extreme value (GEV) distribution.

We assume that choices depend on net annual wages. For each alternative jk , the alternative

specific component V_{hjk} is defined as follows:

$$\begin{aligned} V_{hjk} &= \beta \ln W_{hjk} + \delta_{jk} X_h^V & \text{if } j \in B_I \\ &= 0 & \text{if } j \in B_L \end{aligned} \quad (2)$$

where W_{hjk} represents the net annual wage of individual h in the formal and informal sector, and X_h^V is a vector of individual characteristics that affect the formal/informal options.⁸

The upper nest specific component Z_{hk} is defined as follows:

$$Z_{hk} = \gamma_k X_h^Z \quad (3)$$

where X_h^Z is a vector of characteristics that affect the labor/leisure options.

To account for involuntary unemployment, the labor supply model is combined with the following latent equation:

$$Invol_h^* = \theta X_h^I + \eta_h \quad (4)$$

where X_h^I is a vector of characteristics that affect the probability of being involuntarily unemployed. The indicator variable for involuntary unemployment is equal to one if and only if $Invol_h^* > 0$ (for people who want to work and do not find a job) and to zero otherwise (for people who find a job).

The micro data give information on wages earned by individuals who are employed in the formal sector or the informal sector, and obviously, the wages of non-working individuals are not observed. Moreover, given that an individual is perceived in one option in the labor market, we observe at most one wage for each person. Thus, we have to impute the unobserved wages for all non-chosen options.

⁸Considering this utility maximization problem, informality would not disappear even if taxation were set to zero. Although the effect of an increase in the net wage on the probability of choosing the formal sector is positive, all persons will not modify their labor market choices in favor of formality given that preferences include non-monetary criteria and unobservable characteristics. For the same reason, even if taxation were set to zero, not all of the inactive and voluntarily unemployed persons will decide to participate in the labor market.

The wage equations for formal and informal workers can be represented as follows:

$$\ln w_{hjk} = \alpha_{jk}X_{hjk} + \xi_{hjk} \quad \text{if } j \in B_l \quad (5)$$

where w_{hjk} is the net hourly wage determined by observable personal and job characteristics X_{hjk} and a zero-mean normally distributed error term ξ_{hjk} . Equation 5 is a censored regression because we observe w_{hjk} only if individual h is employed in the formal or the informal sector.

To take into consideration the possible correlation between ϵ_{hjk} and ξ_{hjk} , we estimate the wage equations using Lee's (1983) method to correct for selection bias, i.e., selection specified as a nested logit.⁹ Therefore, we first estimate the nested model (Equation 1) without considering the wage as an explanatory variable. Next, for each individual, we compute the probability of choosing alternative j in nest B_l as well as the value of the sample selection correction variables $-\frac{\phi(\Phi^{-1}(\tilde{P}_{hjk}))}{\tilde{P}_{hjk}}$, where Φ^{-1} represents the inverse of the standard normal distribution, ϕ is the standard normal pdf, and \tilde{P}_{hjk} is the probability for an individual h of choosing option $j \in B_l$.¹⁰ Finally, we write the conditional mean of the wage equations as follows (Lee, 1983):

$$\ln w_{hjk} = \alpha_{jk}X_{hjk} + \sigma_{hjk}\rho_{jk} \left(-\frac{\phi(\Phi^{-1}(\tilde{P}_{hjk}))}{\tilde{P}_{hjk}} \right) + \mu_{hjk} \quad \text{if } j \in B_l \quad (6)$$

We estimate Equation 6 separately for formal and informal workers. Note that negative values of the coefficients of the selection correction variables imply positive selectivity, i.e., persons who choose to work in a given sector obtain, *ceteris paribus*, a higher wage than the average of the total population. Once we obtain the parameter estimates of the wage equations, we can impute the wages for each individual in the non-chosen options. More precisely, for workers in the formal (informal) sector, we calculate a predicted wage in the informal (formal) sector, and

⁹Lee (1983) proposes a consistent two-step procedure based on the conditional logit model. His approach is a generalization of the two-step selection bias correction method introduced by Heckman (1979) and an extension to the case where selectivity is modeled as a multinomial logit. Since it is likely that there will be unobserved similarities among subsets, a generalization of the Lee's approach to a less restrictive nested logit is appropriate (Falaris, 1987).

¹⁰ \tilde{P}_{hjk} are obtained from the estimation of Equation 1 without considering the wages as explanatory variables.

for those persons who do not work, we compute the potential wages in both the formal and the informal sectors.

In the final step, we re-estimate Equation 1 including the individual annual wage, observed for the chosen option and imputed for the non-chosen options. We use P_{hjk} to denote the probability of choosing alternative j in nest B_l . Because $P_{hjk} = P_{hj|B_l} \cdot P_{hB_l}$ (McFadden, 1981), we can write the probability of choosing alternative j (formal/informal sector employment) given that an alternative in nest B_l is chosen and the marginal probability of choosing one alternative in nest B_l :

$$P_{hj|B_l} = \frac{W_{hjk}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{jk} X_h^V}{\lambda}\right)}{\sum_{j' \in B_l} W_{hj'k}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{j'k} X_h^V}{\lambda}\right)} \quad (7)$$

$$P_{hB_l} = \frac{\exp(\gamma_l X_h^Z + \lambda I_{hl})}{\exp(\gamma_L X_h^Z) + \exp(X_h^Z \gamma_l + \lambda I_{hl})} \quad (8)$$

where $I_{hl} = \ln \sum_{j \in B_l} \left[W_{hjk}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{jk} X_h^V}{\lambda}\right) \right]$ is the *Inclusive Value* (IV) of nest B_l , i.e., the quantity that links the upper and the lower nests. This IV enters the upper nest as an explanatory variable. The parameter λ is a measure of the degree of independence in the unobserved utility among the alternatives in nest B_l . Note that a higher value of λ means greater independence and lower correlation. In particular, $\lambda = 1$ implies complete independence within the nest, and thus, the absence of correlation. Moreover, λI_{hl} represents the expected utility that individual h receives from choosing to work, i.e., nest B_l .

Finally, the probability of being involuntarily unemployed, for an individual who chooses to participate, is:

$$P_h^{Invol} = \frac{\exp(\theta X_h^I)}{1 + \exp(\theta X_h^I)} \quad (9)$$

2.2 Data and Descriptive Statistics

We apply the theoretical framework presented above using data from the South African Labour Force Survey (LFS), which is a biannual household survey conducted from February 2000 to September 2007 and is designed to measure the dynamics of employment and unemployment in the country. The survey measures a variety of issues related to the labor market, and most importantly enables us to obtain a good measure of the informal sector. Among the different available surveys, we adopt the September 2000 LFS for several reasons. First, the input-output table necessary to build the macro CGE model is available for the year 2000. Second, the information on the amount of household non-labor income (which is not provided by the LFS although necessary to perform income distribution, inequality and poverty analysis) is available from the September 2000 Income and Expenditure Survey (IES), which uses the same sample of persons with the same identifiers as the 2000 LFS. In addition, the September 2000 LFS contains information that we use as proxy for a person's wealth.¹¹

The 2000 LFS includes a population base of 105,371 on a sample of 26,648 households. Because we are interested in the outcome associated with employment status (formal and informal), the important information required from our data is that of remuneration. All persons in paid employment are explicitly asked for the salary amount in their main job for the week preceding the survey. More precisely, the survey provides a worker's weekly, monthly or annual income and hours worked in the previous week in their main activity. This information allows us to compute hourly wage rates.¹² Another important piece information required for our study is that associated with the distinction between formal and informal employment. The LFS explicitly asks workers whether their main job is in the formal or the informal sector. We use

¹¹One way to improve the specification of the model would be to use panel techniques in order to estimate the wage equations and the discrete labor market choices. Next, the dynamic labor supply functions could be introduced in a dynamic CGE model. These extensions are beyond the scope of our paper. Moreover, data constraints make the implementation of these extensions very difficult.

¹²We note that 2.6% of paid workers (salaried and self-employed) reported their salaries in income categories rather than exact values. For these categories, we use the midpoints of the intervals and the minimum of the open-ended interval.

this information to define the informal employment dummy. Moreover, the survey gives other information related to employment status that allows us to better identify the activity sector of the workers. Among this information and in addition to the worker's answer, we control for whether the worker has a written contract and whether the company for which he/she works is registered and pays social contributions for him/her. Only earnings from the formal sector are considered as taxable. Income tax is computed on the individual level by applying the official tax rates for the year 2000 from the South African Revenue Service (SARS). The tax system is progressive, and the tax rates range from 18% to 42%. Given that unemployment remains high in South Africa and because a subset of the unemployed are found to be involuntary (Kingdon and Knight, 2004), it is crucial for our study to distinguish between voluntary and involuntary unemployment. It is further important to consider involuntary unemployment to avoid biased labor supply elasticities that potentially induce a misleading evaluation of the simulated policy reforms (Bargain et al., 2010).¹³ The dataset provides information on the reasons that prevent the unemployed from working. In particular, we consider unemployment as involuntary if the unemployed person reports having a lack of skills or qualification for available jobs, being seasonal or contract worker, or being recently retrenched. The existence of barriers to entry in the informal sector prevents the involuntarily unemployed from obtaining informal salaried employment. This situation means that involuntarily unemployed are better off in informal jobs, and we thus consider that these people would prefer to work in the informal sector. Given the constraint on the demand side of the labor market, these workers cannot find a job. In contrast, the voluntarily unemployed, as the inactives, choose not to work.

We reduce our sample to people aged 16-64 for males and 16-59 for females. We drop observations of non-paid workers and the self-employed.¹⁴ After controlling for missing values,

¹³Bargain et al. (2010) enumerate several reasons behind biased elasticities from not considering involuntary unemployment: (i) participation bias, which leads to an upward bias of the labor supply elasticities; (ii) preference bias, which acts in the opposite direction; and (iii) specification bias, which has an uncertain effect.

¹⁴The self-employed represent 21.4% of total workers. We exclude them from the analysis because their earnings would be expected to contain measurement error and incorporate returns to risk, capital and entrepreneurship that would not be included in wages of employees. Additionally, although comparing self-employed informal

we retain a total of 52,638 observations. Table 1 presents the summary statistics separately for formal salaried workers, informal salaried workers, the involuntarily unemployed and people who decide not to work (inactives and voluntarily unemployed); these subsamples include 14,138 observations, 2,916 observations, 7,689 observations and 35,584 observations, respectively. As shown in Table 1, the log hourly net wage is higher on average in the formal than the informal sector, and men are more likely to be employed in the formal sector than women. Furthermore, the inactives and unemployed are more likely to have higher education levels than workers in the informal sector.

2.3 Estimation Results

First, we estimate the individual's labor supply choice using the nested logit model presented in Equation 1. In our model formulation, we assume that individuals first choose whether to work on a full-time basis or to stay out of the labor market. Next, conditional on participation, the individuals choose between informal sector and formal sector employment. The estimates are obtained in two stages (see Table 2).

Second, we estimate the wage equations separately for the formal sector and the informal sector, using OLS with correction for selection bias using Lee's method. Table 3 gives the estimation results. The dependent variables are the logarithm of the net hourly wage in the formal and the informal sector, respectively. The estimates show that wages increase with age in both the formal and the informal sectors. Black workers obtain lower wages than Whites in the formal sector. Furthermore, wages in the formal sector increase with firm size. However, apart from firms with 10-19 employees, the firm size has no effect on wages in the informal sector. In both the formal and the informal sectors, wages are higher on average for men and

sector workers with their formal sector counterparts might be of interest in its own right, one could argue that the decision of whether to register one's own enterprise is likely to be determined by different criteria than attempting to obtain a formal sector job. Apart from this observation, the data do not contain information on the salaries and/or the working hours for almost half of the self-employed. Bargain and Kwenda (2011) show that for the period from September 2001 to March 2007, 3% of the self-employed in South Africa paid contributions to social security and that only 10.5% of self-employed workers own firms with more than five employees.

workers in urban areas. For Lee’s selection variables, which are computed from the probabilities obtained in the first step, the estimates yield a negative parameter, thus implying the existence of positive selectivity.

Given the estimates of the wage equations, we impute the hourly net wages in the non-chosen options, and we compute the annual net wages for all individuals in both the formal (W_1) and informal (W_2) sectors.¹⁵ Finally, we re-estimate the formal-informal employment decision by taking into account the individual formal-informal relative wage¹⁶ and the leisure-work decision. The results are presented in Table 4. The main result of the estimates is that the relative wage (W_1/W_2) positively affects participation in the formal salaried sector. In other words, wages are significant determinants of sectoral choice in that the probability of choosing a formal job increases (decreases) if wages in the formal (informal) sector increase. Moreover, men and highly-educated people are more likely to choose the formal employment sector and are also more likely to participate in the labor market. Acting as head of the household increases the probability that the person will choose to work and to work in the formal sector. The results also show that large firms and highly-educated people are more likely to operate formally. For the work-leisure decision, the existence of at least one person in the household who receives pension benefits negatively affects the participation decision. Finally, the parameter estimate of the inclusive value is positive and significantly different from unity, which provides evidence that the nested model is a realistic way to describe the labor market decisions analyzed in this paper.

Following [Bargain et al. \(2010\)](#), we assume that the error terms in Equations 1 and 4 are independent, which allows us to estimate the unemployment probability separately from the labor supply model.¹⁷ The results of this estimation, as reported in Table 5, show that the

¹⁵For non-working persons, we assume that the total working hours per week are equal to the average weekly working hours of the total salaried population.

¹⁶Observed wages are used for the employed persons for the chosen option. For all non-chosen options, we impute potential wages.

¹⁷These probabilities are used in our simulations to identify, after the implementation of a tax reform, people who exit and enter unemployment. In particular, if the unemployment rate decreases (increases) a the national

probability of being involuntarily unemployed decreases with age. Whites and highly-educated people are less likely to be involuntarily unemployed with respect to Blacks and under-educated people.

2.4 The representative Agent Formulation

Our micro-macro simulation approach is based on aggregation of the preferences of individuals who make discrete choices. In this context, we assume that the population is partitioned into $s = 1, \dots, S$ cells according to selected characteristics.¹⁸ Following [Magnani and Mercenier \(2009\)](#), we assume that in each cell, there is a sufficiently large set N_s of statistically identical and independent individuals, each of which has a total time endowment normalized to one. In particular, individuals belonging to the same cell are supposed to be identical according to some observed characteristics, whereas they differ only in the GEV error terms in the utility function. Thus, the individuals belonging to the same cell have the same probability of choosing option j and can be aggregated into a single representative agent.

The aggregate labor supplied by all individuals belonging to cell s for each option $j \in B_L$ ($L_{sj} = P_{sj|B_l} \cdot P_{sB_l} \cdot N_s$) is therefore:

$$L_{sj} = \frac{W_{sj}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{sj}}{\lambda}\right)}{\sum_{j'=1}^2 W_{sj'}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{sj'}}{\lambda}\right)} \cdot \frac{\exp(\gamma_{sl}) \cdot \left[\sum_{j'=1}^2 W_{sj'}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{sj'}}{\lambda}\right)\right]^\lambda}{\exp(\gamma_{sL}) + \exp(\gamma_{sl}) \cdot \left[\sum_{j'=1}^2 W_{sj'}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{sj'}}{\lambda}\right)\right]^\lambda} \cdot N_s \quad \text{if } j \in B_l \quad (10)$$

where W_{sj} represents the average within-cell annual wage for option $j \in B_L$.

Equation 10 represents the labor supply functions that aggregate the preferences of individuals belonging to the same cell, and thus, can be introduced into the CGE model. As shown by

level, we identify who go out of (into) involuntary unemployment.

¹⁸In our model, we consider $S = 32$ cells according to the following characteristics: age category, sex, education, race, and area of residence (see Appendix B.2).

Magnani and Mercenier (2009), these labor supply functions can be derived from the resolution of the optimization problem of the representative agent from each cell.¹⁹

3 The CGE Model

The macro model used in our analysis is a static and multisectoral CGE model based on the South African input-output dataset from the year 2000 provided by the OECD. The input-output table, which includes 48 industries, is aggregated into 10 industries reported in Table A.1 in Appendix A.²⁰ The construction of the SAM (*Social Accounting Matrix*), which is necessary to calibrate our CGE model, is completed using data from national accounts with respect to the balance of payments and the government accounts. The elasticities used in our CGE model come from the GTAP model.

The South African input-output table gives information on the remuneration of labor and capital used by each sector only in the formal sector.²¹ For calibration of the informal sector, we determine the remuneration of the informal labor by applying, for each industry, the weight of the informality, expressed in terms of labor remuneration, observed in the micro dataset (see Table A.1). Given the lack of data, the remuneration of the informal capital for each industry is determined by assuming that the weight of the informality for capital in each industry is the same as that observed for labor.²² Finally, given that GDP is equal to the remuneration of formal factors only and given that in our model we also include the remuneration of informal factors, it is necessary to correspondingly increase the aggregate demand. We assume that the remunerations of the informal labor and capital (which are perceived by the representative

¹⁹Note that this aggregation result allows us to avoid the iterations between the CGE and the microsimulation models.

²⁰The aggregation of industries is carried out with respect to the ten industries that are considered in the micro dataset.

²¹For sector 10, “Private households”, no information is available on labor and capital remuneration in the input-output dataset. We determine the factors’ remunerations by taking into account that the labor remuneration in sector 10 from the micro dataset represents 0.56% of the total formal labor remuneration.

²²A sensitivity analysis is carried out to prove that this hypothesis is not fundamental for policy simulations. Using a weight of informality for capital equal to half of the weight observed for labor, we find that the assumption used does not alter the simulation results. The results are available from the authors upon request.

agent) are entirely consumed.

On the production side of the model, we assume that each industry produces two different types of goods: a formal good and an informal good. In each industry, the formal good is produced using intermediate goods, formal labor, and formal capital, whereas the informal good is produced using informal labor and informal capital.²³ The formal and informal goods are assumed to be qualitatively different, implying the existence of two distinct markets, and thus, the equilibrium prices are also different. Labor units used to produce formal and informal goods are assumed to be non substitutes, implying that the wage in the formal sector is different from the wage in the informal sector. In particular, the relative formal/informal wage affects the individual labor participation choices in the formal and informal sectors. Moreover, only the formal good is subject to indirect taxation. We exclude the self-employed; otherwise, it would be necessary to add (besides the two production functions for formal and informal enterprises) two more production functions for self-employed workers in both the formal and the informal sectors. In addition, the macro data do not allow us to identify self-employed enterprises, to know the quantity of goods and services that they produce and the quantity of production factors that they use. However, the exclusion of the self-employed has no effect on the consistency of the micro-macro analysis given that their remuneration does not enter the labor remuneration but the capital remuneration in the input-output table.

On the demand side, in our model individuals are grouped according to five characteristics: age category, sex, education, race, and area of residence. In particular, we consider two age groups (people aged less than 40 and people aged 40 and more), two education groups (highly-educated and under-educated), two race groups (Black and non-Black), and two areas of residence (urban and rural). Thus, we consider 32 cells. Each cell supplies formal and informal labor according to the labor supply functions that aggregate the preferences of all individuals who belong to the same cell. In our CGE model, we consider one representative household that

²³Given the lack of data, we assume that informal goods are produced without use of intermediate inputs.

earns the net labor (formal and informal) income earned by the 32 cells, capital income, and exogenous transfers from the government. An exogenous fraction of disposable income is saved and the complementary fraction is consumed. The consumption choice relative to formal and informal goods produced by each industry is made by the representative consumer to maximize his/her utility given his/her budget constraint.

With respect to the government, revenues are given by indirect taxes on the production of formal goods and direct taxes on formal labor incomes, whereas expenditures are given by the public consumption of goods and services, the interest on the public debt and transfers to families. We assume that the ratio between public savings and GDP is held constant and that the total public expenditure is endogenously determined to satisfy the government budget constraint.²⁴

Aggregate investments are not determined by aggregate savings, as in the neoclassical closure that is traditionally used in CGE models. Instead, we use a Keynesian closure, which consists of holding constant the level of aggregate investments. This choice allows us to endogenize the unemployment rate, which is subsequently determined in a such a way as to verify the macroeconomic equilibrium condition between aggregate investments and aggregate savings. Thus, in our model, unemployment is explained by the weakness in aggregate demand. The unemployment rate for each cell is determined by assuming that the share of unemployed in each cell with respect to the total unemployed remains unchanged.

Finally, we assume that domestic prices and wages are perfectly flexible and adjust to guarantee the equilibrium in each market, i.e., the markets of formal and informal goods and services of the ten industries, the two labor markets, and the capital market.

²⁴A sensitivity analysis is carried out to prove that this hypothesis is not fundamental for policy simulations. We use different fiscal rules: (i) the public deficit is exogenous, whereas public expenditures are endogenously determined to satisfy the government's budget constraint; (ii) the public deficit-GDP ratio is held constant, whereas public expenditures are endogenous; and (iii) public expenditures are held constant, whereas the public deficit is endogenous. We find that the hypothesis used in our model does not alter the simulation results. The results are available from the authors upon request.

4 The Personal Income Tax System in South Africa

For the purpose of our study, we focus our analysis on personal income taxation in South Africa, and in this section, we summarize its most important features. Given the year in which the sample is collected, we apply the 2000/2001 tax rates.²⁵ The South African tax system is progressive. The tax system consists of six brackets ranging from a rate of 18% applied to incomes less than or equal to 35,000 Rands to 42% for incomes greater than 200,000 Rands (Table A.2 in Appendix A). The primary tax rebate, for which all taxpayers are eligible, amounts to 3,800 Rands, and thus, the tax threshold for persons under 65 years is equal to 21,111 Rands.

The taxable unit is not the family, but the individual. In other words, the individual income tax rate structure is applicable to all persons regardless of marital status.²⁶ However, this method implies that in the presence of a progressive tax system, fiscal discrimination remains between a one-income-earner family and a family with two members earning the same taxable income together. Obviously, because of the tax progressivity, this discrimination increases with total family income. The South African tax legislation aims to reduce the tax gap and increase tax collections. Although the child rebate has been removed, the primary rebate continues to increase on an annual basis. The majority of registered individual taxpayers are men given that participation of men is higher than that of women and men are more likely to have formal employment activity. In 2000/2001, the personal income tax accounted for 86,478 millions of Rands (9.1% of GDP) and represented 39.2% of total tax revenue.

²⁵The South African fiscal year spans from the first of March to the end of February.

²⁶The earlier South African income tax system sustains the belief that households with two income earners are better off than households with one income earner, assuming that altruism prevails within the household. Thus, the second earner (i.e., the wife) was taxed more heavily. In 1994, the *Katz Commission* was charged to evaluate the appropriateness of the tax system and make some recommendations to improve it. The main purpose of this commission was to establish equity between men and women. Following the recommendations of the Katz Commission, the South African government introduced several tax policy changes since 1994 among which the introduction of a unified structure for all individuals, the adjustments of the tax rates and income brackets, and the reduction of the number of income brackets.

5 Simulation Results

In this section, we analyze the effects of three reforms in South Africa using our micro-macro simulation model. In the first simulated reform, we replace the current progressive tax system with a Negative Income Tax (NIT) system that combines a flat tax rate and a transfer paid to all individuals. Next, we analyze the effects of different tax policies that combine a flat rate and a transfer paid only to formal workers.²⁷ Finally, we simulate the effects of two expansionary fiscal policies: (i) a reduction in the total amount of taxes; and (ii) an increase in the public expenditures.

For each of the mentioned reforms, we analyze the macroeconomic effects as well as the effects at the individual level. The main economic mechanisms produced by the introduction of a tax policy are described as follows. Tax policies produce a direct effect on the labor supply because they modify the average tax rate, and thus, the opportunity cost of working and of choosing the employment sector. Second, a general equilibrium effect on wages is produced. Accordingly, the change in the labor supply might be absorbed by firms only through an adjustment in wages, which in turn affects the labor market choices. Third, a change in the unemployment rate computed at the macro level provokes a change at the micro level in the number of people who are involuntarily unemployed.²⁸

5.1 Negative Income Tax System

A negative income tax system, which combines a flat rate and a lump-sum transfer, supports maintenance of the progressivity of the tax system at a notably low administrative cost. In our analysis, the flat tax rate is set to 20% and the transfers paid to all individuals are computed to guarantee that the pre-reform total amount of taxes remains unchanged. The lump-sum

²⁷Peichl (2009) simulates a revenue-neutral flat tax reform of the German tax system using a linked microsimulation-CGE model. De Mooij (2008) uses a CGE model with a detailed classification of households and analyses the effects of a flat tax reform in the Netherlands.

²⁸In contrast with Bargain et al. (2010), we take into account the effects on the demand side of the labor market that affect the involuntary unemployment.

transfers are equal to 408 Rands. Consequently, non-workers and informal workers receive transfers from the government, whereas formal workers receive a net transfer only if their income is lower than the transfer divided by the flat rate, which implies that the tax becomes negative.

Table 6 presents the macro effects of our simulations. Given that the flat rate used in our simulation (20%) is higher than the initial (pre-reform) average tax rate paid by formal workers (12.6%), it is straightforward that such a tax policy produces an incentive to leave the formal sector and to reduce labor participation. In particular, we find that the number of workers decreases in the formal sector by 6.1%, whereas it increases by 14.1% in the informal sector. The overall effect on the number of workers is negative (-2.6%). However, the labor supply depends not only on the number of workers but also on their productivity. The number of units of effective labor decreases by 3.7% in the formal sector but increases by 13.8% in the informal sector. Thus, in the formal sector, the reduction in the number of units of effective labor is lower than the reduction in the number of workers, and vice versa in the informal sector. This result implies that people who decide to move from the formal to the informal sector are “less productive” on average. The wages in the formal and informal sectors adjust to guarantee equilibrium in the two labor markets. In particular, given the changes in the labor supply expressed in units of effective labor, the equilibrium wage in the formal sector increases by 1.9% and decreases in the informal sector by 4.1%. The reduction in formal labor incomes implies a reduction in the aggregate income tax. This effect has a positive impact on the private disposable income and consequently on the consumption and the aggregate demand, which permits a reduction in the unemployment rate by 0.6 percentage points (p.p.). The value added created at the national level in the formal sector, which represents real GDP, decreases by 1.7%, and the value added created in the informal sector increases by 4.2%. The introduction of a NIT system produces negative effects in terms of efficiency given that the total production decreases by 1.3%. Table 7 presents the effect of the NIT system on the size of the informal sector measured in terms of production, number of workers and number of units of effective

labor. The results confirm that the size of the informal sector increases.

To analyze the effects of the tax policies on the individual labor market choices, we first generate 300 GEV conditional error terms for each individual starting from the estimation of the nested logit model. In our analysis, the individual labor market choices are determined by taking into account not only the change in the taxation rules but also the change in the equilibrium wages and in the unemployment rate determined in the CGE model. Thus, the individual effects are computed by taking into account the general equilibrium effects.

Tables 8 and 9 present the effects on the work/leisure decision and on the formal/informal employment decision for the total sample and for different groups by age category, sex, education, race, and area of residence.²⁹ At the aggregate level, we find that both the fraction of people who want to work and the fraction of people who want to work in the formal sector decrease by 1.6 p.p. In the results for different groups, as previously mentioned in the discussion of the effects at the macroeconomic level, it is possible to observe that these effects apply to all categories and to a greater extent, to individuals belonging to “less productive” groups (young adults, females, under-educated, Blacks and persons living in rural areas). In fact, with a flat rate equal to 20%, people belonging to these groups reduce their labor participation and their participation in the formal sector because they face a strong increase in the marginal tax rate. The reduction in the unemployment rate determined at the macro level implies that some individuals who were previously involuntarily unemployed (and who present the lowest probability of being unemployed) find a job. Table 10 shows that the unemployment rate decreases by 0.6 p.p. and that this reduction is beneficial to all the categories.

It is further possible to analyze the effects on income inequality and poverty and to do this, we first compute the post-reform net annual wage for each individual, which depends on the labor participation, the choice of the employment sector, the equilibrium level of wages in the formal and informal sectors, and the taxation rule. Next, we compute the net annual

²⁹Tables A.3, A.4 and A.5 in Appendix A present the detailed results for the 32 cells of our model.

labor income at the household level to which we add the household non-labor incomes. The inequality and poverty analyses are carried out by computing the equivalent incomes using the OECD equivalence scales. Table 11 presents the effects on the Gini index for the total sample and for the different groups. We find a strong reduction in inequalities. The Gini index decreases from 0.604 to 0.589. This positive effect applies to all categories (except for non-Blacks), and in particular, to people belonging to “less productive” groups. Table 12 presents the effects on poverty for the entire sample and for all groups by computing the fraction of poor individuals (headcount ratio).³⁰ We find that the fraction of poor individuals decreases for the whole sample (from 34.4% to 32.6%) and for all categories, especially for “less productive” groups.

In terms of income distribution, Table 13 confirms that the introduction of a NIT system improves the economic situation of poor individuals, especially for those who belong to “less productive” groups. The net annual income of the tenth percentile strongly increases (+140% on average). In contrast, the net annual income of the fiftieth and the ninetieth percentiles slightly decrease (-2.7% and -1.5%, respectively). Finally, it is possible to determine who wins and who loses from the introduction of a NIT system. Table 14 shows that 64.1% of the population obtains a higher net income with an average gain equal to 7.4%. In contrast, 35.9% of the population loses and the average loss is equal to 7%.

Our simulation confirms that the introduction of a NIT system worsens the equity-efficiency trade-off. The lump-sum transfers, which are paid to both non-workers and informal workers, allow a reduction in income inequality and poverty. However, the reform discourages people from participating in the labor market and from working in the formal sector, which produces negative effects on employment and production.

³⁰In our analysis, an individual is defined as poor if his/her equivalent income is lower than the poverty line, which is defined as half of the median equivalent income.

5.2 Alternative Tax Policies

The negative income tax system can be criticized on the grounds that people receive a transfer without making efforts to obtain a job. Therefore, in this section, we consider the introduction of a flat tax rate combined with transfers paid only to formal workers. In our simulations, we consider three flat rates, i.e., 15%, 20% and 25%, and the transfers to formal workers are computed to guarantee that the pre-reform total amount of taxes remains unchanged and are 486, 1,513 and 2,540 Rands, respectively.

With a simulated flat rate equal to 15%, the average tax rate paid by formal workers significantly increases for people who belong to “less productive” groups (especially for the under-educated, Blacks and people living in rural areas) and decreases for “more productive” groups (especially for the highly-educated, non-Blacks and people living in urban areas). In contrast, a flat rate equal to 25% is beneficial to people who belong to “less productive” groups. Interestingly, the average tax rates are little affected by the introduction of a flat rate equal to 20%. In particular, for people living in urban areas and people living in rural areas, the average tax rates remain essentially unchanged.

Given that the introduction of a flat rate equal to 20% implies that the average tax rates are little affected, the macro and micro effects are quite small, implying that this tax policy allows a reproduction of the current progressive tax system. In contrast, the introduction of a flat rate equal to 15% implies that the system becomes less progressive, whereas the introduction of a flat rate equal to 25% implies that the system becomes more progressive.

Concerning the macro effects of our simulations, Table 6 shows that the number of workers in the formal sector decreases if the flat rate is 15% (-1.7%), remains essentially unchanged if the flat rate is 20% (-0.2%), and increases if the flat rate is 25% (+1.3%). The number of workers in the informal sector increases if the flat rate is 15% (+3.2%), remains essentially unchanged if the flat rate is 20% (0.4%) and decreases if the flat rate is 25% (-1.4%). The overall effect on the

number of workers is negative if the flat rate is 15% (-0.8%), almost nil if the flat rate is 20%, and positive if the flat rate is 25% (+0.9%). By analyzing the effects on the number of units of effective labor in the formal and informal sectors, if the flat rate is 15%, it is possible to conclude that people who decide to move from the formal to the informal sector are “less productive” on average. These individuals decide to move to the informal sector because if they remain in formal employment, they will have to pay additional taxes. If the flat rate is 25%, people who decide to move from the informal to the formal sector are “less productive” on average. These individuals decide to move to the formal sector as the formal sector becomes relatively more attractive for “less productive” individuals because the system becomes more progressive. The equilibrium wage in the formal sector slightly increases if the flat rate is 15% (+0.1%), whereas it remains essentially unaffected if the flat rate is 20% or 25%. In contrast, the equilibrium wage in the informal sector decreases if the flat rate is 15% (-0.8%), remains essentially unaffected if the flat rate is equal to 20% and increases if the flat rate is equal to 25% (+0.4%). The value added created at the national level in the formal sector slightly decreases if the flat rate is 15%, but it remains essentially unchanged if the flat rate is 20% and 25%. The value added created in the informal sector increases if the flat rate is 15% (+0.9%), remains essentially unaffected if the flat rate is equal to 20% and decreases if the flat rate is equal to 25% (-0.3%). The total production remains essentially unchanged with all these reforms. In the case of a flat rate equal to 25%, the unemployment rate decreases by 0.2 percentage points. This result is explained by the fact that with this reform, the number of workers in the formal sector increases, which implies an increase in the total amount of transfers to formal workers, and thus, a reduction in the total income tax. The increase in private disposable income stimulates consumption and aggregate demand, which reduces unemployment in our demand-driven model. In contrast, with a flat rate equal to 15% and 20%, the total income tax remains essentially unchanged, and consequently, the unemployment rate is unaffected. Table 7 confirms that the size of the informal sector increases if the flat rate is 15%, remains essentially unchanged if the flat rate is

equal to 20%, and decreases if the flat rate is equal to 25%.

For the effects on the individual choices, Table 8 shows that, at the aggregate level, the fraction of people who want to work decreases if the flat rate is 15% (-0.4 p.p.) and increases if the flat rate is 25% (+0.3 p.p.). Table 9 shows that the fraction of people who would choose to work in the formal sector decreases if the flat rate is 15% (-0.5 p.p.) and increases if the flat rate is 25% (+0.4 p.p.). With respect to the results for different groups, it is possible to observe that these effects primarily concern individuals who belong to “less productive” groups (young adults, females, under-educated, Blacks and persons living in rural areas). In fact, if the flat rate is 15%, people belonging to “less productive” groups reduce their labor participation and their participation in the formal sector because they face an increase in the average tax rate.³¹ If the flat rate is 20%, the labor participation and the participation in the formal sector are affected neither at the aggregate level, nor at the group level. If the flat rate is 25%, people belonging to “less productive” groups increase their labor participation and their participation in the formal sector because they benefit from a reduction in the average tax rate. As previously mentioned in the analysis of the macro effects, Table 10 shows that the unemployment rate is significantly affected only by the introduction of the flat rate of 25%. In this case, the unemployment rate at the aggregate level decreases by 0.2 percentage points. Table 10 also shows that the reduction in the unemployment rate is beneficial to people belonging to “less productive” groups. As shown in Table 11, income inequality increases if the flat rate is 15% at the aggregate level and for all categories. This result is clearly explained by the fact that the reduction in the level of progressivity implies that high income workers are less heavily taxed, whereas low income workers are more heavily taxed. If the flat rate is 20%, inequalities slightly increase both at the aggregate level and for all categories. In contrast, if the flat rate is 25%, inequalities decrease at the aggregate level and for the “less productive” groups. Table 12 presents the effects of the tax reforms on poverty. We find that the number of poor individuals increases if the tax rate

³¹Only non-Blacks increase their labor participation because they benefit from a significant reduction in the average tax rate.

is 15% at the aggregate level and for all categories (except for non-Blacks), remains essentially unchanged if the tax rate is 20% and decreases if the tax rate is 25% (excepted for non-Blacks and people living in urban areas). Table 13 shows that the net annual income of the tenth percentile decreases for all categories (except for non-Blacks) if the flat rate is 15%, is not affected if the flat rate is 20%, and increases (for “less productive” groups) if the flat rate is 25%. The net annual income of the fiftieth percentile decreases for all categories (except for non-Blacks) if the flat rate is 15%, slightly decreases (-1.1%, on average) if the flat rate is 20%, and increases (+1.6%, on average) if the flat rate is 25%. The net annual income of the ninetieth percentile increases if the flat rate is 15% (+1.9%, on average), remains essentially unchanged if the flat rate is 20%, and decreases if the flat rate is 25% (-0.8%, on average). In terms of winners and losers, Table 14 shows that the fraction of the population that wins is only 13.1% if the flat rate is 15%, 16.7% if the flat rate is 20%, and 32.1% if the flat rate is 25%. In contrast, the fraction of the population that loses is 45.8% if the flat rate is 15%, 42.3% if the flat rate is 20%, and 27.6% if the flat rate is 25%.

Our results imply that the introduction of a flat rate equal to 25% combined with a transfer of 2,540 Rands to the formal workers permits improvement of the equity-efficiency trade-off. Employment increases, and income inequality and poverty decrease.³²

5.3 Expansionary Fiscal Policies

As previously mentioned, South Africa is characterized by a high level of involuntary unemployment and a low level of informality, with respect to other developing countries. This situation is explained by the presence of barriers to entry in the informal sector. In our paper, we assume that the weakness in aggregate demand is responsible for both the high unemployment and the low level of informality. Thus, the government has the possibility to stimulate the aggregate demand through a reduction in taxation or an increase in public expenditures, which permits

³²De Mooij (2008) finds similar results for the Netherlands, where the equity-efficiency trade-off can be improved by combining a flat rate with in-work tax credits.

a reduction in the unemployment rate, and given our assumption that only people who want to work in the informal sector can be involuntarily unemployed, an increase in the size of the informal sector.

Here, we simulate the economic consequences of two expansionary fiscal policies. In the first, a flat tax equal to 20% is applied and a transfer, paid to all formal workers, is computed such that the pre-reform total amount of taxes decreases by 10%. The transfer is equal to 1,772 Rands which is higher than the amount received by formal workers with the tax policy presented in Section 5.2. In the second, the public expenditures/GDP ratio increases by 1 percentage point, from 18.1% to 19.1%.

Table 6 shows that the number of workers increases in both the formal sector (+1.2% in the case of a reduction in taxes and +1% in the case of an increase in public expenditures) and in the informal sector (+6.7% and +15%, respectively). At the aggregate level, the number of workers increases by 2.2% and 3.4%, respectively. The value added increases respectively by 0.5% and 0.6% in the formal sector and by 2.7% and 5% in the informal sector. The total production increases respectively by 0.6% and 0.9%. The strong increase in the size of the informal sector (see Table 7) is explained by the fact that both fiscal policies allow for a significant reduction in the unemployment rate (respectively -1.4 p.p. and -2.8 p.p.).

For the effects at the individual level on labor market choices (see Tables 8 and 9), we find that the fraction of people who choose to work in the formal sector increases with both reforms (+0.6 p.p. with the reduction in taxes and by 0.9 p.p. with the increase in public expenditures). These effects are explained by the change in the relative formal/informal wage.³³ The changes in formal and informal wages provoke a reduction in the labor participation by 0.3 p.p. in the case of an increase in public expenditures, whereas it remains unaffected in the case of a reduction

³³Both expansionary policies permit a reduction in the unemployment rate, implying a strong increase in the supply of informal labor, which in turn reduces the wage in the informal sector. Thus, the relative formal/informal wage increases, which induces people to prefer to work in the formal sector. In addition, in the case of the reduction in taxes, the amount of transfers paid to formal workers increases, which implies an increase in the net formal wage.

in taxes. However, it is worth noting that in both simulations, the number of workers increases given that the unemployment rate decreases (see Table 10), and the number of informal workers increases because unemployed people who find a job will work in the informal sector. Tables 11 and 12 show that the expansionary fiscal policies also produce positive effects on income inequality and poverty. In particular, the reduction in taxation permits a decrease in inequality among people belonging to “less productive” groups (under-educated, Blacks and people living in rural areas), whereas the increase in public expenditures allows a decrease in inequalities for all categories. The fraction of poor individuals decreases from 0.344 to 0.336 with both fiscal policies, with significant improvements for Blacks and people living in rural areas. Table 13 shows that the net annual income increases with both expansionary reforms for the tenth, the fiftieth and the ninetieth percentiles. However, in terms of winners and losers, Table 14 shows that 24.3% of the population wins in the case of a reduction in taxation (with an average gain of 5%) and that 45.5% wins in the case of an increase in public expenditures (with an average gain of 2.1%). In contrast, 35.5% of the population loses in the case of a reduction in taxation (with an average loss of 1.2%), and 11.9% loses in the case of an increase in public expenditures (with an average loss of 3.7%).

To conclude, our simulations show that both expansionary fiscal policies improve the equity-efficiency trade-off. However, it is important to highlight that these policies are assumed to be financed by public borrowing. Consequently, the public deficit would increase from 2.1% of GDP to 2.7% in the case of a reduction in taxes and to 3.1% in the case of an increase in public expenditures.

6 Conclusions

In this paper, we apply a new micro-macro simulation approach to analyze the impact of different tax policies in developing countries characterized by a large informal sector and a low level of

participation in the labor market. Our micro-macro model is built using the approach developed by [Magnani and Mercenier \(2009\)](#) based on the aggregation theory of [Anderson *et al.* \(1992\)](#) that permits aggregation of preferences of individuals facing discrete choices. We use data from South Africa, and we assume that individuals decide whether not to work, to work in the formal sector or to work in the informal sector.

We use our micro-macro simulation model to analyze the effects of different types of tax reform. In particular, we quantify the effects at the macro level (on GDP, equilibrium prices and wages, and unemployment rate) and at the individual level (on the labor market choice of each individual, income distribution, and inequality and poverty) by taking into account the general equilibrium effects provoked at the macro level. We first show that the introduction of a negative income tax system worsens the equity-efficiency trade-off. This type of reform reduces inequalities and poverty but, at the same time, produces negative effects on employment and production because the transfers paid to all individuals discourage people from participating in the labor market and working in the formal sector. Next, we replace the current progressive tax system with a flat tax system combined with transfers paid only to formal workers. We show that the equity-efficiency trade-off can be improved if the flat tax rate and the transfer paid to the formal workers are sufficiently high (25% and 2,540 Rands, respectively). In this case, we show that employment increases, whereas income inequality and poverty decrease. Finally, we show that another method that allows improvement of the equity-efficiency trade-off is the implementation of expansionary fiscal policies. A reduction in the total amount of taxes or an increase in public expenditures allows an increase in the aggregate demand and a reduction in the exceptionally high unemployment rate in South Africa. Consequently, employment and real GDP increase, and inequalities and poverty decrease. One interesting result of our simulations concerns the relationship between the size of the informal sector and the level of income inequality and poverty. In general, it is reasonable to think that a higher size of the informal sector implies a higher level of inequality and poverty. Such a relationship

is verified if the increase in the size of the informal sector is due to people moving from the (high-paid) formal to the (low-paid) informal sector. However, if the increase in the size of the informal sector is due to the reduction in the level of unemployment, both inequality and poverty may decrease. This result is obtained in our simulations in the case of the introduction (i) of a (revenue-neutral) negative income tax system which combines a flat rate equal to 25% and a lump-sum transfer, and (ii) of an expansionary policy (reduction in total taxes and increases in public expenditures). These reforms produce a reduction in the unemployment rate at the macro level. Given our assumption that involuntarily unemployed would choose to work in the informal sector, the reduction in unemployment provokes, from one hand, an increase in the size of the informal sector and, from the other hand, a reduction in income inequality and poverty given that some poor people now find a job, although it is in the informal sector.

This paper constitutes one possible application of this new micro-macro simulation approach which offers the advantage of ease to implement compared with other micro-macro simulation techniques because it allows aggregation of individual choices into explicit functions that can be introduced into the macro model. Many other applications could be contemplated in the study of other discrete choices such as migration, education, and profession.

References

- Anderson, S., de Palma, A., Thisse, J-F., 1992. *Discrete Choice Theory of Product Differentiation*. Cambridge: MIT Press.
- Bargain, O., Kwenda, P., 2011. Earnings Structures, Informal Employment, And Self-Employment: New Evidence From Brazil, Mexico, And South Africa. *Review of Income and Wealth*, 57, S100-S122.
- Bargain, O., Caliendo, M., Haan, P., Orsini, K., 2010. Making work pay in a rationed labor market. *Journal of Population Economics*, 23, 323-351.

- Davies, R. and J. Thurlow. 2009. Formal-Informal Economy Linkages and Unemployment in South Africa. IFPRI Discussion Paper 00943, December, 2009.
- De Mooij, R.A., 2008. Reinventing the Dutch tax-benefit system: exploring the frontier of the equity-efficiency trade-off. *International Tax and Public Finance*, 15, 87-103.
- El Badaoui, E., Strobl, E., Walsh, F., 2008. Is There an Informal Employment Wage Penalty? Evidence from South Africa. *Economic Development and Cultural Change*, 56, 683-710.
- Falaris, E.M., 1987. A Nested Logit Migration Model with Selectivity. *International Economic Review*, 28, 429-443.
- Fields, G.S., 1980. *Poverty, Inequality, and Development*. New York: Cambridge University Press
- Fields, G.S., 1975. Rural-urban migration, urban unemployment and underemployment, and job-search activity in LDCs. *Journal of Development Economics*, 2, 165-187.
- Fourie, F. 2011. The South African Unemployment Debate: Three Worlds, Three Discourses?. SALDRU Working Papers No. 63.
- Fugazza, M., Jacques, J-F., 2004. Labor market institutions, taxation and the underground economy. *Journal of Public Economics*, 88, 395-418.
- Gong, X., van Soest, A., 2002. Wage differentials and mobility in the urban labour market: A panel data analysis for Mexico. *Labour Economics*, 9, 513-529.
- Gong, X., van Soest, A., Villagomez, E., 2004. Mobility in the urban labor market: A panel data analysis for Mexico. *Economic Development and Cultural Change*, 53, 1-36.
- Günther, I., Launov, A., 2012. Informal employment in developing countries: Opportunity or last resort? *Journal of Development Economics*, 97, 88-98.
- Gutmann, P.M., 1977. The subterranean economy. *Financial Analysts Journal*, 33, 26-34.

- Hart, K., 1973. Informal income opportunities and urban employment in Ghana. *Journal of Modern African Studies*, 11, 61-89.
- Heckman, J., 1979. Sample Selection Bias as a Specification Error. *Econometrica*, 47, 153-161.
- Ihrig, J., Moe, K., 2004. Lurking in the shadows: the informal sector and government policy. *Journal of Development Economics*, 73, 541-557.
- Johnson, S., Kaufmann, D., Zoido-Lobaton, P., 1998. Regulatory Discretion and the Unofficial Economy. *American Economic Review*, 88, 387-392.
- Kingdon, G., Knight, J., 2006. How Flexible Are Wages in Response to Local Unemployment in South Africa? *Industrial and Labor Relations Review*, 59, 471-495.
- Kingdon, G., Knight, J., 2004. Unemployment in South Africa: The Nature of the Beast. *World Development*, 32, 391-408.
- Lee, L-F., 1983. Generalized Econometric Models with Selectivity. *Econometrica*, 51, 507-512.
- Lemieux, T., Fortin, B., Fréchet, P., 1994. The effect of Taxes on Labor Supply in the Underground Economy. *American Economic Review*, 84, 231-254.
- Loayza, N.V., 1996. The economics of the informal sector: a simple model and some empirical evidence from Latin America. *Carnegie-Rochester Conference Series on Public Policy*, 45, 129-162.
- Magnani, R., Mercenier, J., 2009. On Linking Microsimulation and Computable General Equilibrium Models using Exact Aggregation of Heterogeneous Discrete-choice Making Agents. *Economic Modelling*, 26, 560-570.
- Maloney, W.F., 2004. Informality Revisited. *World Development*, 32, 1159-1178.
- Maloney, W., 1999. Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico. *World Bank Economic Review*, 13, 275-302.

- McFadden, D., 1981. Econometric Models of Probabilistic Choice. In C. F. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications*. Cambridge: MIT Press, 198-272.
- McLaren, J., 1998. Black Markets and Optimal Evadable Taxation. *Economic Journal*, 108, 665-679.
- Mulinge, M., Mueller, C.W., 1998. Employee Job Satisfaction in Developing Countries: the Case of Kenya. *World Development*, 26, 2181-2199.
- Peichl, A., 2009. The benefits and problems of linking micro and macromodels - Evidence from a flat tax analysis. *Journal of Applied Economics*, 12, 301-329.
- Pratap, S., Quintin, E., 2006. Are Labor Markets Segmented in Developing Countries? A Semiparametric Approach. *European Economic Review*, 50, 1817-1841.
- Saavedra, J., Chong, A., 1999. Structural Reform, Institutions and Earnings: Evidence from the Formal and Informal Sectors in Urban Peru. *Journal of Development Studies*, 35, 95-116.
- Saracoglu, D.S., 2008. The Informal Sector and Tax on Employment: A Dynamic General Equilibrium Investigation. *Journal of Economic Dynamics and Control*, 32, 529-549.
- Skinner, C., 2008. The Struggle for the Streets: Processes of Exclusion and Inclusion of Street Traders in Durban, South Africa. *Development Southern Africa*, 25, 227-242.
- Tanzi, V., 1987. Quantitative characteristics of the tax systems of developing countries. In Newbery and Stern.
- Tanzi, V., Zee, H., 2000. Tax Policy for Emerging Markets - Developing Countries. IMF Working Papers, No. 00/35.

Table 1: Summary Statistics

	FORMAL WORKERS		INFORMAL WORKERS		INVOL. UNEMPLOYED		INACTIVES	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Log hourly gross wage	2.023	(1.080)	0.887	(0.860)				
Log hourly net wage	1.968	(1.012)	0.887	(0.860)				
<u>INDIVIDUAL CHARACTERISTICS</u>								
Male	0.613	(0.487)	0.339	(0.473)	0.432	(0.495)	0.426	(0.494)
Age	37.056	(10.233)	38.261	(10.905)	30.547	(9.890)	28.379	(12.513)
Head of household	0.615	(0.486)	0.573	(0.495)	0.198	(0.398)	0.164	(0.370)
White	0.103	(0.304)	0.008	(0.090)	0.010	(0.101)	0.042	(0.201)
Black	0.681	(0.466)	0.876	(0.329)	0.918	(0.275)	0.842	(0.365)
English	0.100	(0.300)	0.015	(0.122)	0.015	(0.120)	0.044	(0.205)
Afrikaans	0.229	(0.420)	0.125	(0.331)	0.071	(0.257)	0.120	(0.325)
Single	0.592	(0.491)	0.442	(0.497)	0.288	(0.453)	0.248	(0.432)
Married	0.030	(0.171)	0.065	(0.247)	0.023	(0.149)	0.026	(0.160)
Read	0.925	(0.263)	0.830	(0.376)	0.919	(0.272)	0.927	(0.260)
Write	0.920	(0.272)	0.823	(0.382)	0.916	(0.277)	0.924	(0.265)
No education	0.071	(0.257)	0.164	(0.371)	0.075	(0.263)	0.067	(0.249)
Nursery	0.002	(0.045)	0.003	(0.052)	0.002	(0.048)	0.002	(0.043)
Primary	0.182	(0.386)	0.326	(0.469)	0.219	(0.413)	0.186	(0.389)
Secondary	0.577	(0.494)	0.499	(0.500)	0.690	(0.462)	0.708	(0.455)
NTC	0.008	(0.088)	0.001	(0.026)	0.002	(0.045)	0.005	(0.073)
Diploma	0.160	(0.366)	0.007	(0.086)	0.012	(0.109)	0.032	(0.017)
Urban area	0.644	(0.479)	0.586	(0.492)	0.537	(0.499)	0.546	(0.498)
House owner	0.386	(0.487)	0.540	(0.498)	0.769	(0.422)	0.736	(0.441)
Pension dummy	0.077	(0.267)	0.077	(0.267)	0.235	(0.424)	0.215	(0.411)
Household workers number	0.619	(0.863)	0.499	(0.807)	0.471	(0.704)	0.541	(0.771)
<u>PROVINCE</u>								
Western Cape	0.167	(0.373)	0.084	(0.277)	0.067	(0.250)	0.078	(0.269)
Eastern Cape	0.082	(0.274)	0.139	(0.346)	0.154	(0.361)	0.146	(0.353)
Northern Cape	0.058	(0.233)	0.055	(0.228)	0.018	(0.132)	0.043	(0.204)
Free State	0.092	(0.290)	0.083	(0.276)	0.054	(0.225)	0.078	(0.269)
Kwazulu-Natal	0.164	(0.370)	0.166	(0.372)	0.202	(0.402)	0.189	(0.392)
North West	0.093	(0.290)	0.123	(0.329)	0.113	(0.317)	0.109	(0.312)
Gauteng	0.180	(0.384)	0.156	(0.363)	0.168	(0.374)	0.133	(0.339)
Mpumalanga	0.085	(0.279)	0.099	(0.298)	0.087	(0.282)	0.091	(0.287)
Northern province	0.078	(0.269)	0.094	(0.292)	0.137	(0.344)	0.131	(0.338)
Regional unemployment rate	24.731	(5.176)	27.431	(4.748)	27.384	(4.515)	27.141	(4.625)
<u>FIRM SIZE</u>								
1 employee	0.039	(0.195)	0.613	(0.487)				
2-4 employees	0.098	(0.397)	0.255	(0.436)				
5-9 employees	0.138	(0.345)	0.064	(0.244)				
10-19 employees	0.169	(0.375)	0.036	(0.185)				
20-49 employees	0.192	(0.394)	0.018	(0.132)				
50+ employees	0.363	(0.481)	0.014	(0.116)				
<u>INDUSTRY DUMMIES</u>								
Industry 1	0.154	(0.361)	0.081	(0.272)				
Industry 2	0.087	(0.281)	0.003	(0.052)				
Industry 3	0.157	(0.364)	0.025	(0.155)				
Industry 4	0.010	(0.101)	0.001	(0.032)				
Industry 5	0.048	(0.215)	0.077	(0.266)				
Industry 6	0.156	(0.363)	0.059	(0.235)				
Industry 7	0.043	(0.204)	0.028	(0.165)				
Industry 8	0.078	(0.268)	0.006	(0.076)				
Industry 9	0.222	(0.416)	0.021	(0.143)				
Industry 10	0.042	(0.202)	0.701	(0.458)				
Observations	14,138		2,916		7,689		35,584	

Notes: (i) The sample study contains 52,638 observations. The statistics are based on information from the 2000 LFS and the amount of the household non-labor income comes from the 2000 IES; (ii) Industries refer to: 1- Agriculture, hunting, forestry and fishing, 2- Mining and quarrying, 3- Manufacturing, 4- Electricity, gas and water supply, 5- Construction, 6- Wholesale and retail trade, 7- Transport and communication, 8- Financial intermediation, insurance, real estate, 9- Public administration and 10- Private households.

Table 2: Nested Logit Model of Labor Participation

	<i>Work Dummy</i>		<i>Formal Dummy</i>	
Male dummy	0.329	(0.034)	-0.068	(0.079)
Age	0.436	(0.006)	0.132	(0.015)
Age squared	-0.005	(0.000)	-0.001	(0.000)
White	-0.263	(0.058)	0.554	(0.160)
Black	0.151	(0.092)	-0.021	(0.177)
English	-0.152	(0.098)	0.780	(0.194)
Afrikaans	0.190	(0.093)	0.526	(0.180)
Single	0.106	(0.027)	0.479	(0.052)
Married	-0.311	(0.064)	0.189	(0.122)
Read	0.302	(0.127)	0.277	(0.248)
Write	-0.064	(0.122)	-0.051	(0.237)
Nursery	0.194	(0.224)	0.376	(0.453)
Primary	0.038	(0.071)	0.143	(0.142)
Secondary	0.003	(0.073)	0.671	(0.146)
NTC	-0.433	(0.158)	1.603	(0.425)
Diploma	0.400	(0.084)	3.392	(0.184)
Urban area	-0.086	(0.025)	0.215	(0.056)
House owner	-0.568	(0.025)	-0.614	(0.052)
Household head	1.126	(0.029)	0.526	(0.054)
Children 0-3 years	0.013	(0.016)	-0.054	(0.038)
Children 4-6 years	-0.015	(0.018)	-0.060	(0.043)
Children 7-12 years	-0.105	(0.012)	-0.023	(0.027)
Province unemployment rate	0.000	(0.005)	-0.004	(0.011)
Pension dummy	-0.065	(0.029)	-	-
Household workers number	0.157	(0.014)	-	-
Household informal workers	-	-	-4.144	(0.247)
2-4 employees	-	-	1.262	(0.090)
5-9 employees	-	-	2.687	(0.123)
10-19 employees	-	-	3.412	(0.140)
20-49 employees	-	-	4.212	(0.173)
50+ employees	-	-	0.623	(0.097)
<i>Inclusive Value</i>	0.462	(0.020)	-	-
Constant	-8.841	(0.221)	-6.041	(0.496)
R²	0.24		0.62	
Observations	52,638		24,743	

Notes: (i) The second and third columns give estimates of the binary choice models with the formal employment dummy and the work dummy as dependent variables, respectively; (ii) Standard errors are in parentheses; (iii) We include dummies for the different provinces in South Africa; (iv) Industry dummies are included in the participation choice.

Table 3: Estimates of the Wage Equations with Sample Bias Correction

	<i>Formal Sector</i>		<i>Informal Sector</i>	
Male dummy	0.242	(0.014)	0.152	(0.035)
Age	0.099	(0.005)	0.062	(0.010)
Age squared	-0.001	(0.000)	-0.001	(0.000)
White	0.435	(0.026)	0.901	(0.179)
Black	-0.147	(0.047)	0.024	(0.110)
English	0.249	(0.049)	0.045	(0.131)
Afrikaans	-0.015	(0.048)	-0.154	(0.114)
Read	0.234	(0.072)	-0.234	(0.129)
Write	0.045	(0.067)	0.248	(0.129)
Nursery	-0.144	(0.139)	-0.031	(0.283)
Primary	-0.006	(0.043)	0.109	(0.073)
Secondary	0.372	(0.044)	0.221	(0.077)
NTC	0.899	(0.083)	0.670	(0.570)
Diploma	1.248	(0.047)	0.420	(0.194)
Urban area	0.392	(0.015)	0.135	(0.033)
2-4 employees	0.154	(0.038)	-0.012	(0.040)
5-9 employees	0.319	(0.038)	-0.079	(0.075)
10-19 employees	0.390	(0.038)	-0.253	(0.098)
20-49 employees	0.454	(0.038)	-0.100	(0.131)
50+ employees	0.580	(0.035)	-0.118	(0.131)
<i>Selection Correction</i>	-0.069	(0.018)	-0.119	(0.039)
Constant	-1.741	(0.134)	-0.960	(0.234)
R²	0.47		0.17	
Observations	14,138		2,916	

Notes: (i) The dependent variables are the logarithm of the net hourly wage in the formal and informal sector, respectively; (ii) Standard errors are in parentheses; (iii) We include dummies for the different provinces in South Africa.

Table 4: Nested Logit Model of Labor Participation Including Wages

	<i>Work Dummy</i>		<i>Formal Dummy</i>	
Male dummy	0.087	(0.037)	-0.173	(0.081)
Age	0.340	(0.006)	0.102	(0.015)
Age squared	-0.004	(0.000)	-0.001	(0.000)
White	-1.380	(0.074)	0.975	(0.166)
Black	0.451	(0.104)	0.038	(0.181)
English	-0.650	(0.112)	0.612	(0.199)
Afrikaans	0.136	(0.105)	0.365	(0.185)
Single	-0.187	(0.030)	0.488	(0.053)
Married	-0.293	(0.069)	0.214	(0.125)
Read	0.368	(0.141)	-0.111	(0.252)
Write	-0.254	(0.135)	0.045	(0.241)
Nursery	0.003	(0.241)	0.417	(0.463)
Primary	-0.076	(0.077)	0.330	(0.145)
Secondary	-0.436	(0.079)	0.596	(0.150)
NTC	-1.690	(0.191)	1.236	(0.440)
Diploma	-1.648	(0.098)	2.821	(0.189)
Urban area	-0.149	(0.027)	-0.101	(0.058)
House owner	-0.234	(0.028)	-0.621	(0.053)
Household head	0.691	(0.032)	0.520	(0.056)
Children 0-3 years	0.052	(0.017)	-0.056	(0.039)
Children 4-6 years	0.012	(0.019)	-0.062	(0.044)
Children 7-12 years	-0.086	(0.013)	-0.020	(0.028)
Province unemployment rate	0.015	(0.005)	-0.007	(0.011)
Pension dummy	-0.093	(0.032)	-	-
Household workers number	0.136	(0.016)	-	-
Household informal workers	-	-	-4.235	(0.252)
2-4 employees	-	-	1.154	(0.093)
5-9 employees	-	-	2.414	(0.125)
10-19 employees	-	-	2.961	(0.143)
20-49 employees	-	-	3.830	(0.176)
50+ employees	-	-	0.040	(0.102)
$\ln(W_1/W_2)$	-	-	0.885	(0.040)
<i>Inclusive Value</i>	0.869	(0.012)	-	-
Constant	-12.981	(0.251)	-5.345	(0.512)
R²	0.37		0.64	
Observations	52,638		24,743	

Notes: (i) The second and third columns give estimates of the binary choice models with the formal employment dummy and the work dummy as dependent variables, respectively; (ii) Standard errors are in parentheses; (iii) We include dummies for the different provinces in South Africa; (iv) Industry dummies are included in the work choice.

Table 5: Estimation of the Probability of being Involuntarily Unemployed

Male dummy	-0.003	(0.056)
Age	-0.174	(0.011)
Age squared	0.002	(0.000)
White	-0.429	(0.141)
Black	0.309	(0.164)
English	-0.985	(0.187)
Afrikaans	-0.674	(0.165)
Single	-0.386	(0.040)
Married	0.027	(0.106)
Read	-0.174	(0.203)
Write	-0.022	(0.195)
Nursery	0.007	(0.352)
Primary	0.022	(0.113)
Secondary	-0.129	(0.116)
NTC	-0.945	(0.340)
Diploma	-2.402	(0.160)
Urban area	0.069	(0.040)
House owner	0.913	(0.039)
Household head	-1.513	(0.045)
Children 0-3 years	0.197	(0.027)
Children 4-6 years	0.135	(0.031)
Children 7-12 years	0.112	(0.019)
Pension dummy	0.207	(0.051)
Household workers number	-0.458	(0.025)
Province unemployment rate	0.040	(0.008)
Constant	2.449	(0.368)
R²	0.31	
Observations	24,743	

Notes: (i) Standard errors are in parentheses; (ii) We include dummies for the different provinces in South Africa.

Table 6: Macroeconomic Effects of the Reforms (in %)

		<i>Value Added</i>	<i>Units of labor</i>	<i>Units of effective labor</i>	<i>Unempl. rate</i>	<i>Real wage</i>
NIT						
$\tau = 20$	Total	-1.3	-2.6	-2.5	-0.6	
	Formal	-1.7	-6.1	-3.7		1.9
	Informal	4.2	14.4	13.8		-4.1
Flat Tax						
$\tau = 15$	Total	0.0	-0.8	0.0	0.0	
	Formal	-0.1	-1.7	-0.2		0.1
	Informal	0.9	3.2	2.9		-0.8
$\tau = 20$	Total	0.0	-0.1	-0.1	0.0	
	Formal	0.0	-0.2	-0.1		0.0
	Informal	0.1	0.4	0.4		-0.1
$\tau = 25$	Total	0.0	0.9	-0.1	-0.2	
	Formal	0.0	1.3	0.0		0.0
	Informal	-0.3	-1.4	-1.1		0.4
Expansionary Fiscal Policies						
Decrease in T	Total	0.6	2.2	1.2	-1.4	
	Formal	0.5	1.2	0.8		-0.2
	Informal	2.7	6.7	6.5		-1.4
Increase in G	Total	0.9	3.4	1.7	-2.8	
	Formal	0.6	1.0	0.8		0.4
	Informal	5.0	15.0	14.3		-3.5

Table 7: Effects on the Size of the Informal Sector (in p.p.)

	Production	Workers	Units of effective labor
INITIAL	<i>6.1</i>	<i>17.1</i>	<i>6.6</i>
NIT			
$\tau=20$	0.3	2.5	0.9
FLAT TAX			
$\tau=15$	0.1	0.5	0.2
$\tau=20$	0.0	0.1	0.0
$\tau=25$	0.0	-0.2	-0.1
DECREASE IN T	0.2	1.1	0.4
INCREASE IN G	0.3	2.6	0.9

Notes: The initial values are expressed in percentage.

Table 8: Effects on Individual Labor Participation Choice (in p.p.)

	Initial	NIT	Flat Tax		Decrease in T	Increase in G	
		$\tau=20\%$	$\tau=15\%$	$\tau=20\%$			$\tau=25\%$
All	<i>47.0</i>	-1.6	-0.4	0.0	0.3	0.0	-0.3
15-39	<i>42.5</i>	-1.7	-0.4	0.0	0.4	0.1	-0.3
> 40	<i>59.6</i>	-1.4	-0.2	-0.1	0.0	0.0	-0.2
Females	<i>42.3</i>	-1.6	-0.4	0.0	0.4	0.1	-0.4
Males	<i>52.3</i>	-1.7	-0.4	-0.1	0.2	0.0	-0.2
Under- educated	<i>51.9</i>	-2.1	-0.7	0.0	0.7	0.2	-0.4
Highly- educated	<i>45.2</i>	-1.4	-0.2	-0.1	0.1	0.0	-0.3
Black	<i>45.6</i>	-1.7	-0.5	-0.1	0.3	0.0	-0.3
Non-Black	<i>52.5</i>	-1.1	0.2	0.2	0.1	0.3	-0.1
Rural	<i>43.7</i>	-1.9	-0.6	0.1	0.8	0.3	-0.3
Urban	<i>49.4</i>	-1.4	-0.2	-0.2	-0.1	-0.1	-0.3

Notes: The initial values are expressed in percentage.

Table 9: Effects on Individual Formal Employment Choice after Reforms (in p.p.)

	Initial	NIT	Flat Tax		Decrease in T	Increase in G	
			$\tau=20\%$	$\tau=15\%$			$\tau=20\%$
All	<i>57.1</i>	-1.6	-0.5	0.0	0.4	0.6	0.9
15-39	<i>52.4</i>	-1.9	-0.6	0.1	0.8	0.8	1.0
> 40	<i>66.5</i>	-1.2	-0.3	-0.3	-0.2	0.3	0.8
Females	<i>46.5</i>	-1.8	-0.6	0.1	0.8	0.9	0.9
Males	<i>66.8</i>	-1.4	-0.5	-0.2	0.1	0.4	0.9
Under- educated	<i>49.3</i>	-2.8	-1.3	0.1	1.5	1.0	0.9
Highly- educated	<i>60.4</i>	-1.1	-0.2	-0.1	0.0	0.5	0.9
Black	<i>50.0</i>	-2.1	-0.8	-0.1	0.6	0.7	1.0
Non-Black	<i>82.0</i>	-0.3	0.2	0.1	0.1	0.5	0.6
Rural	<i>51.4</i>	-2.5	-1.0	0.3	1.6	1.2	1.0
Urban	<i>60.9</i>	-1.1	-0.2	-0.3	-0.3	0.3	0.9

Notes: The initial values are expressed in percentage.

Table 10: Effects on Unemployment Rate after Reforms (in p.p.)

	Initial	NIT	Flat Tax		Decrease in T	Increase in G	
			$\tau=20\%$	$\tau=15\%$			$\tau=20\%$
All	<i>31.1</i>	-0.6	0.0	0.0	-0.2	-1.4	-2.8
15-39	<i>38.0</i>	-0.6	0.1	-0.1	-0.4	-1.8	-3.3
> 40	<i>17.4</i>	-0.4	0.0	0.1	0.2	-0.7	-1.6
Females	<i>37.1</i>	-0.8	0.0	-0.1	-0.4	-1.7	-3.2
Males	<i>25.6</i>	-0.4	0.1	0.1	0.0	-1.1	-2.4
Under- educated	<i>31.0</i>	0.0	0.4	-0.1	-0.8	-1.6	-2.7
Highly- educated	<i>31.1</i>	-0.8	-0.1	0.1	0.1	-1.3	-2.8
Black	<i>36.7</i>	-0.5	0.2	0.1	-0.3	-1.6	-3.2
Non-Black	<i>11.5</i>	-0.5	-0.3	-0.1	0.0	-0.6	-1.1
Rural	<i>36.3</i>	-0.3	0.3	-0.3	-1.1	-2.0	-3.2
Urban	<i>27.6</i>	-0.8	-0.1	0.2	0.3	-1.0	-2.5

Notes: The initial values are expressed in percentage.

Table 11: Gini Inequality Index

	Initial	NIT	Flat Tax		Decrease in T	Increase in G	
		$\tau=20\%$	$\tau=15\%$	$\tau=20\%$			$\tau=25\%$
All	0.604	0.589	0.615	0.608	0.601	0.603	0.600
15-39	0.598	0.578	0.607	0.600	0.593	0.596	0.593
> 40	0.599	0.594	0.612	0.605	0.599	0.601	0.597
Females	0.623	0.605	0.632	0.626	0.619	0.621	0.619
Males	0.582	0.570	0.594	0.586	0.579	0.582	0.578
Under-educated	0.536	0.498	0.542	0.535	0.528	0.531	0.532
Highly-educated	0.603	0.591	0.612	0.607	0.601	0.603	0.598
Black	0.568	0.535	0.572	0.567	0.561	0.562	0.563
Non-Black	0.518	0.521	0.522	0.521	0.519	0.517	0.516
Rural	0.628	0.586	0.632	0.625	0.617	0.620	0.622
Urban	0.549	0.543	0.560	0.556	0.552	0.552	0.545

Table 12: The Headcount Ratio

	Initial	NIT	Flat Tax		Decrease in T	Increase in G	
		$\tau=20\%$	$\tau=15\%$	$\tau=20\%$			$\tau=25\%$
All	0.344	0.326	0.351	0.344	0.336	0.336	0.336
15-39	0.368	0.348	0.375	0.368	0.358	0.358	0.358
> 40	0.279	0.265	0.285	0.280	0.275	0.273	0.274
Females	0.368	0.349	0.375	0.368	0.359	0.359	0.359
Males	0.317	0.301	0.324	0.318	0.310	0.310	0.310
Under-educated	0.407	0.390	0.419	0.406	0.392	0.395	0.399
Highly-educated	0.321	0.303	0.326	0.322	0.315	0.314	0.312
Black	0.392	0.371	0.401	0.392	0.381	0.382	0.382
Non-Black	0.147	0.142	0.145	0.148	0.150	0.145	0.145
Rural	0.474	0.450	0.485	0.468	0.449	0.459	0.464
Urban	0.247	0.234	0.250	0.251	0.251	0.244	0.240

Table 13: Effects of Tax Reforms on Income Distribution (in Rands)

		Initial	NIT	Flat Tax		Decrease in T	Increase in G	
			$\tau=20\%$	$\tau=15\%$	$\tau=20\%$	$\tau=25\%$		
10th percentile								
All		389	887	376	388	395	400	405
Age	15-39	360	861	351	360	368	373	383
	> 40	469	974	458	468	470	487	486
Sex	Females	352	835	342	352	360	366	370
	Males	424	944	409	424	429	445	450
Education	Under- educated	289	801	276	289	300	300	297
	Highly- educated	433	941	425	431	439	450	458
Race	Black	329	840	316	331	341	342	349
	Non-Black	1202	1608	1220	1179	1107	1241	1241
Zone	Rural	275	773	269	278	293	291	288
	Urban	532	1064	528	524	524	548	572
50th percentile								
All		6170	6006	5920	6102	6270	6277	6343
Age	15-39	5501	5431	5263	5445	5648	5638	5648
	> 40	8365	7952	8054	8232	8341	8424	8469
Sex	Females	5456	5395	5212	5402	5585	5574	5568
	Males	7039	6701	6730	6963	7120	7196	7223
Education	Under- educated	4413	4379	4197	4424	4618	4584	4566
	Highly- educated	7172	7002	6960	7085	7246	7289	7392
Race	Black	4832	4885	4623	4803	4990	4963	4982
	Non-Black	16778	16649	17149	16890	16841	17158	16982
Zone	Rural	3416	3595	3280	3455	3707	3567	3514
	Urban	9456	9172	9289	9232	9254	9463	9693
90th percentile								
All		24329	23967	24790	24472	24133	24669	24446
Age	15-39	21781	21216	21879	21827	21594	22052	21894
	> 40	31770	31878	33382	32362	31369	32677	32030
Sex	Females	22794	22444	23144	22843	22440	23036	22962
	Males	25931	25693	26501	26070	25649	26415	26202
Education	Under- educated	14329	13498	13848	14199	14628	14552	14459
	Highly- educated	28120	27932	29028	28426	27892	28741	28331
Race	Black	17492	16646	17033	17102	17224	17362	17729
	Non-Black	46414	47725	49817	48062	46855	48251	46705
Zone	Rural	13461	12592	12977	13497	13826	13752	13587
	Urban	30963	30936	32108	31199	30622	31683	31255

Table 14: Winners and Losers (in %)

	Fraction of winners	Average gain	Fraction of losers	Average loss
NIT				
$\tau=20$	64.1	7.4	35.9	-7.0
FLAT TAX				
$\tau=15$	13.1	5.6	45.8	-3.8
$\tau=20$	16.7	3.3	42.3	-2.2
$\tau=25$	32.1	3.7	27.6	-2.5
DECREASE IN T	24.3	5.0	35.5	-1.2
INCREASE IN G	48.5	2.1	11.9	-3.7

Appendix A

Table A.1: List of Industries of the CGE Model

	(1)	(2)	(3)	(4)	(5)
1 Agriculture, hunting, forestry and fishing	4.0%	8.5%	4.5%	12.6%	13.9%
2 Mining and quarrying	6.2%	81.5%	25.4%	0.3%	0.3%
3 Manufacturing	33.8%	16.6%	25.5%	1.7%	2.0%
4 Electricity, gas and water supply	2.6%	2.0%	0.0%	0.5%	0.5%
5 Construction	4.7%	0.2%	0.8%	23.5%	24.9%
6 Wholesale and retail trade	11.6%	11.0%	2.1%	6.0%	6.5%
7 Transport and communication	8.1%	12.2%	15.8%	8.1%	10.2%
8 Financial intermediation, insurance, real estate	13.4%	3.3%	2.4%	0.6%	0.6%
9 Public administration	14.0%	0.5%	0.8%	0.6%	0.7%
10 Private households	1.7%	0.0%	0.0%	85.1%	85.1%
Total	100%			6.1%	6.6%

Notes: Column (1) gives the industry weight in the total valued added produced in the economy; Columns (2) and (3) indicate for each industry the ratio between the exports and the production, and the ratio between the imports and the production, respectively; In columns (4) and (5), we present for each industry the size of the informal sector measured respectively in terms of the value added and labor remuneration.

Table A.2: Tax Rates for Individuals - 2000/2001

Taxable Income	Rates of Tax	Proportion
R1 - R35 000	18% of each R1	74.6%
R35 001 - R45 000	R6 300 + 26% of the amount above R35 000	7.1%
R45 001 - R60 000	R8 900 + 32% of the amount above R45 000	6.0%
R60 001 - R70 000	R13 700 + 37% of the amount above R60 000	4.0%
R70 001 - R200 000	R17 400 + 40% of the amount above R70 000	7.7%
R200 001 - and above	R69 400 + 42% of the amount above R200 000	0.7%
Individual Primary Rebates	R3 800	
Tax Threshold under 65 years	R21 111	

Source: South African Revenue Service (*SARS*).

Among formal workers, 55.7% does not pay taxes due to rebates and tax threshold.

Table A.3: Effects on Individual Labor Participation Choice (in p.p.)

RACE	EDUCATION	AREA OF RESIDENCE	SEX	AGE	Initial	NIT				Decrease in T	Increase in G
						$\tau=20\%$	$\tau=15\%$	Flat Tax			
						$\tau=20\%$	$\tau=25\%$	$\tau=20\%$	$\tau=25\%$		
Black	Under-educated	Rural	Females	15-39	40.7	-1.7	-0.3	0.7	1.6	0.8	-0.5
				>40	43.5	-2.0	-0.6	0.3	1.1	0.4	-0.4
			Males	15-39	50.5	-2.6	-0.8	0.4	1.6	0.7	-0.3
		>40		65.7	-2.6	-1.1	-0.5	0.1	-0.3	-0.2	
		Urban	Females	15-39	49.2	-1.4	-0.4	0.0	0.4	-0.1	-0.6
				>40	54.5	-1.7	-0.6	-0.2	0.2	-0.2	-0.5
	Males		15-39	49.9	-2.0	-0.7	-0.2	0.3	-0.1	-0.4	
		>40	66.6	-2.3	-1.0	-0.9	-0.7	-0.8	-0.2		
	Highly-educated	Rural	Females	15-39	30.1	-1.4	-0.4	0.2	0.7	0.2	-0.4
				>40	56.2	-1.4	-0.2	-0.1	0.0	-0.1	-0.3
			Males	15-39	37.1	-2.0	-0.7	-0.1	0.5	0.0	-0.3
		>40		72.4	-1.5	-0.5	-0.5	-0.6	-0.5	-0.1	
		Urban	Females	15-39	40.2	-1.7	-0.5	-0.2	0.1	-0.2	-0.4
				>40	66.0	-1.3	-0.2	-0.3	-0.4	-0.3	-0.2
	Males		15-39	46.4	-1.7	-0.4	-0.3	-0.2	-0.2	-0.2	
		>40	71.5	-0.8	0.2	-0.1	-0.4	-0.1	-0.1		
Non-Black	Under-educated	Rural	Females	15-39	60.6	-3.2	-0.9	1.2	3.0	1.7	-0.2
				>40	53.0	-3.1	-1.2	0.1	1.4	0.5	-0.2
			Males	15-39	83.9	-2.3	-0.9	0.0	0.9	0.3	0.0
		>40		86.3	-1.9	-0.9	-0.3	0.3	0.0	0.0	
		Urban	Females	15-39	45.3	-1.9	-0.7	-0.1	0.5	-0.1	-0.5
				>40	30.7	-1.8	-0.7	-0.4	0.0	-0.3	-0.3
	Males		15-39	48.6	-2.7	-1.2	-0.6	0.0	-0.4	-0.2	
		>40	45.6	-2.4	-1.0	-0.8	-0.6	-0.6	-0.1		
	Highly-educated	Rural	Females	15-39	48.4	-1.5	0.1	0.5	0.8	0.7	-0.1
				>40	44.8	-1.5	-0.1	-0.3	-0.4	-0.1	0.0
			Males	15-39	66.9	-1.8	-0.3	-0.1	0.2	0.2	0.0
		>40		77.9	-0.3	0.6	0.2	-0.1	0.3	0.0	
		Urban	Females	15-39	45.3	-1.6	-0.2	-0.3	-0.4	-0.2	-0.1
				>40	45.3	-0.4	0.8	0.3	-0.3	0.4	0.0
	Males		15-39	53.6	-0.4	0.8	0.4	0.0	0.5	0.0	
		>40	66.2	0.9	1.9	1.1	0.4	1.2	0.0		

Notes: The initial values are expressed in percentage.

Table A.4: Effects on Individual Formal Employment Choice (in p.p.)

RACE	EDUCATION	AREA OF RESIDENCE	SEX	AGE	Initial	Flat Tax				Decrease in T	Increase in G
						NIT	Flat Tax				
						$\tau=20\%$	$\tau=15\%$	$\tau=20\%$	$\tau=25\%$		
Black	Under-educated	Rural	Females	15-39	27.1	-2.8	-0.5	2.4	5.1	3.6	0.8
				>40	37.5	-3.2	-1.2	0.9	2.9	2.0	1.0
			Males	15-39	52.3	-3.8	-1.5	1.0	3.2	2.1	1.0
		>40		67.8	-2.9	-1.5	-0.8	0.0	-2.1	0.9	
		Urban	Females	15-39	17.6	-1.8	-0.7	0.3	1.3	0.0	0.6
				>40	30.1	-2.6	-1.4	-0.5	0.3	0.2	0.9
	Males		15-39	45.4	-2.8	-1.3	-0.4	0.6	0.5	1.0	
		>40	69.3	-2.3	-1.4	-1.4	-1.3	-0.8	0.9		
	Highly-educated	Rural	Females	15-39	28.8	-2.6	-0.9	0.9	2.6	1.8	0.8
				>40	55.6	-1.0	0.0	-0.1	-0.2	0.5	1.0
			Males	15-39	48.9	-3.1	-1.4	-0.1	1.1	0.8	1.0
		>40		70.5	-1.4	-0.6	-0.9	-1.1	-0.4	0.8	
		Urban	Females	15-39	41.2	-2.2	-1.0	-0.5	0.1	0.3	1.0
				>40	61.7	-1.0	-0.2	-0.6	-0.9	0.0	1.0
	Males		15-39	62.5	-1.4	-0.5	-0.5	-0.5	0.1	1.0	
		>40	76.8	0.0	0.4	-0.1	-0.7	0.2	0.7		
Non-Black	Under-educated	Rural	Females	15-39	69.7	-3.6	-1.1	1.7	4.0	2.8	0.9
				>40	65.7	-3.7	-1.7	0.2	2.0	1.2	0.9
			Males	15-39	90.7	-1.4	-0.7	0.0	0.6	0.4	0.3
		>40		92.9	-1.0	-0.5	-0.2	0.2	0.1	0.3	
		Urban	Females	15-39	35.7	-3.2	-1.6	-0.2	1.1	0.7	0.9
				>40	45.5	-3.1	-1.7	-1.0	-0.2	-0.2	1.0
	Males		15-39	62.3	-3.3	-1.9	-1.0	-0.2	-0.2	1.0	
		>40	73.7	-1.9	-1.0	-0.9	-0.8	-0.4	0.8		
	Highly-educated	Rural	Females	15-39	79.1	-0.6	0.2	0.5	0.7	0.9	0.7
				>40	90.0	-0.3	0.0	-0.1	-0.2	0.1	0.4
			Males	15-39	89.1	-0.5	-0.1	0.0	0.1	0.2	0.4
		>40		94.8	0.1	0.2	0.1	0.0	0.2	0.2	
		Urban	Females	15-39	77.9	-0.7	-0.1	-0.3	-0.5	0.1	0.7
				>40	87.1	0.3	0.6	0.2	-0.2	0.4	0.5
	Males		15-39	85.3	0.4	0.7	0.3	-0.1	0.5	0.5	
		>40	91.2	0.7	0.9	0.5	0.1	0.6	0.3		

Notes: The initial values are expressed in percentage.

Table A.5: Effects on Unemployment Rate (in p.p.)

RACE	EDUCATION	AREA OF RESIDENCE	SEX	AGE	Initial	NIT	Flat Tax		Decrease in T	Increase in G	
						$\tau=20\%$	$\tau=15\%$	$\tau=20\%$			$\tau=25\%$
Black	Under-educated	Rural	Females	15-39	56.0	-0.9	-0.1	-1.9	-3.7	-4.4	-4.5
				>40	31.1	-0.1	0.3	-0.5	-1.3	-1.9	-2.6
			Males	15-39	36.5	0.8	0.8	-0.8	-2.3	-2.7	-3.2
		>40		18.8	0.6	0.7	0.4	0.1	-0.6	-1.8	
		Urban	Females	15-39	55.6	-1.8	-0.1	-0.2	-0.7	-2.3	-4.2
				>40	22.4	-0.4	0.2	0.1	0.0	-0.8	-1.8
	Males		15-39	42.0	-0.2	0.6	0.2	-0.3	-1.7	-3.6	
		>40	17.5	0.3	0.6	0.8	0.8	-0.1	-1.7		
	Highly-educated	Rural	Females	15-39	62.2	-1.1	0.2	-0.8	-2.0	-3.5	-5.0
				>40	25.5	-0.8	-0.2	0.1	0.2	-1.1	-2.3
			Males	15-39	43.5	0.2	0.7	0.1	-0.8	-2.0	-3.8
		>40		20.6	-0.2	0.2	0.6	0.9	-0.4	-2.0	
		Urban	Females	15-39	45.3	-0.7	0.3	0.3	0.1	-1.6	-3.9
				>40	18.8	-0.5	-0.1	0.3	0.5	-0.6	-1.7
	Males		15-39	31.8	-0.5	0.1	0.4	0.6	-1.1	-3.0	
		>40	17.7	-0.9	-0.5	0.1	0.6	-0.7	-1.7		
Non-Black	Under-educated	Rural	Females	15-39	13.8	0.8	0.4	-0.8	-1.8	-1.7	-1.3
				>40	5.7	0.3	0.2	0.0	-0.3	-0.4	-0.5
			Males	15-39	3.7	0.3	0.2	0.0	-0.2	-0.3	-0.4
		>40		1.5	0.1	0.1	0.0	0.0	-0.1	-0.2	
		Urban	Females	15-39	40.0	-0.2	0.6	0.1	-0.5	-1.6	-3.3
				>40	10.0	0.0	0.2	0.2	0.1	-0.3	-0.9
	Males		15-39	27.4	0.8	1.1	0.7	0.2	-0.7	-2.5	
		>40	11.1	0.2	0.3	0.4	0.4	-0.2	34.4		
	Highly-educated	Rural	Females	15-39	11.5	-0.3	-0.2	-0.3	-0.4	-0.9	-1.1
				>40	2.2	-0.1	0.0	0.0	0.1	-0.1	-0.2
			Males	15-39	5.9	0.0	0.0	0.0	0.0	-0.3	-0.6
		>40		0.9	-0.1	0.0	0.0	0.0	-0.1	-0.1	
		Urban	Females	15-39	18.1	-0.4	-0.1	0.2	0.5	-0.6	-1.8
				>40	8.2	-0.6	-0.5	-0.1	0.2	-0.5	-0.8
	Males		15-39	11.6	-0.9	-0.7	-0.2	0.1	-0.8	-1.2	
		>40	6.1	-0.8	-0.7	-0.4	-0.1	-0.6	-0.6		

Notes: The initial values are expressed in percentage.

Appendix B: Technical Description of the CGE Model

B.1 The Production Side

Each industry $i = 1, \dots, 10$ produces two different types of goods: a formal good ($j = 1$) and an informal good ($j = 2$). We denote by Y_{i1} the production level of the formal good in industry i , which depends on the total quantity of intermediate goods Int_i , formal labor (expressed in efficiency units) L_{i1} , and formal capital K_{i1} . In contrast, the production level of the informal good Y_{i2} depends on the quantity of informal labor (expressed in efficiency units) L_{i2} and informal capital K_{i2} .³⁴

We use a CES production function as follows:

$$Y_{i1} = \left[(\alpha_i^{Int})^{\frac{1}{\sigma_i}} \cdot Int_i^{\frac{\sigma_i-1}{\sigma_i}} + (\alpha_{i1}^L)^{\frac{1}{\sigma_i}} \cdot L_{i1}^{\frac{\sigma_i-1}{\sigma_i}} + (\alpha_{i1}^K)^{\frac{1}{\sigma_i}} \cdot K_{i1}^{\frac{\sigma_i-1}{\sigma_i}} \right]^{\frac{\sigma_i}{\sigma_i-1}} \quad (\text{B.1})$$

$$Y_{i2} = \left[(\alpha_{i2}^L)^{\frac{1}{\sigma_i}} \cdot L_{i2}^{\frac{\sigma_i-1}{\sigma_i}} + (\alpha_{i2}^K)^{\frac{1}{\sigma_i}} \cdot K_{i2}^{\frac{\sigma_i-1}{\sigma_i}} \right]^{\frac{\sigma_i}{\sigma_i-1}} \quad (\text{B.2})$$

Each industry i produces the quantity of formal and informal goods by choosing the optimal level of production factors that maximizes its profit given the technological constraint. The first order conditions for industry i that produces the formal good are:

$$Int_i = \alpha_i^{Int} \cdot \left[\frac{P_{i1}^Y \cdot (1 - \tau_i^Y)}{P_i^{Int}} \right]^{\sigma_i} \cdot Y_{i1} \quad (\text{B.3})$$

$$L_{i1} = \alpha_{i1}^L \cdot \left[\frac{P_{i1}^Y \cdot (1 - \tau_i^Y)}{w_1} \right]^{\sigma_i} \cdot Y_{i1} \quad (\text{B.4})$$

$$K_{i1} = \alpha_{i1}^K \cdot \left[\frac{P_{i1}^Y \cdot (1 - \tau_i^Y)}{r + \delta_K} \right]^{\sigma_i} \cdot Y_{i1} \quad (\text{B.5})$$

$$P_{i1}^Y \cdot (1 - \tau_i^Y) \cdot Y_{i1} = P_i^{Int} \cdot Int_i + w_1 \cdot L_{i1} + (r + \delta_K) \cdot K_{i1} \quad (\text{B.6})$$

³⁴Given the lack of data, we suppose that informal goods are produced without using intermediate inputs.

and, for industry i that produces the informal good are:

$$L_{i2} = \alpha_{i2}^L \cdot \left[\frac{P_{i2}^Y}{w_2} \right]^{\sigma_i} \cdot Y_{i2} \quad (\text{B.7})$$

$$K_{i2} = \alpha_{i2}^K \cdot \left[\frac{P_{i2}^Y}{r + \delta_K} \right]^{\sigma_i} \cdot Y_{i2} \quad (\text{B.8})$$

$$P_{i2}^Y \cdot Y_{i2} = w_2 \cdot L_{i2} + (r + \delta_K) \cdot K_{i2} \quad (\text{B.9})$$

P_{ij}^Y is the equilibrium price of the formal good ($j = 1$) and the informal good ($j = 2$) that are produced by industry i , w_j is the equilibrium wage per unit of effective labor for sector $j = 1, 2$, and $(r + \delta_K)$ is the equilibrium gross remuneration rate of a unit of capital, where δ_K is the depreciation rate. The parameter τ_i^Y represents the tax rate on the production of the formal good and P_i^{Int} represents the aggregate price of intermediate goods used by industry i in the formal sector. The capital is considered to be perfectly mobile across sectors and industries, whereas formal and informal labor are assumed to be perfectly mobile across industries. These assumptions imply the existence of an equilibrium wage for formal labor, an equilibrium wage for informal labor, and a unique equilibrium rate for capital remuneration.

Given the total quantity of intermediate goods Int_i , each industry i that produces a formal good chooses the optimal quantity of goods to buy from industry i' . The first order conditions that allow minimization of the total cost are:

$$Int_{i'i} = \alpha_{Int_{i'i}} \cdot \left(\frac{P_i^{Int}}{P_{i'1}^C} \right)^{\sigma_i^{Int}} \cdot Int_i \quad (\text{B.10})$$

$$P_i^{Int} \cdot Int_i = \sum_{i'} P_{i'1}^C \cdot Int_{i'i} \quad (\text{B.11})$$

where P_{i1}^C is the average purchase price of the formal good in industry i , defined in Equation B.27 hereafter.

B.2 The Representative Agent

In our model, individuals are grouped according to five characteristics: age category, sex, education, race, and area of residence. In particular, we consider two age groups (people aged less than 40 and people aged 40 and more), two education groups (highly-educated and under-educated), two race groups (Black and non-Black), and two areas of residence (urban and rural). Thus, we consider $s = 1, \dots, S$ with $S = 32$ cells.

The net wage earned by each cell is equal to $[(1 - \tau_{sj}) \cdot w_j \cdot A_{sj}]$, where (i) w_j is the (equilibrium) wage per unit of effective labor in sector $j = 1, 2$; (ii) A_{sj} is the productivity of cell s in sector j ; and (iii) τ_{sj} is the average tax rate paid by workers belonging to cell s (see Equations B.13 and B.14), with $\tau_{sj} = 0$ for $j = 2$ because informal workers do not pay direct taxes. Thus, $(w_j \cdot A_{sj})$ represents the average wage earned in sector j by individuals who belong to cell s . This average wage is obtained from the estimation of a wage equation using the micro dataset and with age category, sex, education, race, and area of residence as explanatory variables.

Each cell s supplies formal and informal labor according to the following labor supply functions:

$$L_{sj} = \frac{[(1 - \tau_{sj}) \cdot w_j \cdot A_{sj}]^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{sj}}{\lambda}\right)}{\sum_{j'=1}^2 [(1 - \tau_{sj'}) \cdot w_{j'} \cdot A_{sj'}]^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{sj'}}{\lambda}\right)} \quad (\text{B.12})$$

$$\cdot \frac{\exp(\gamma_{sl}) \cdot \left[\sum_{j'=1}^2 [(1 - \tau_{sj'}) \cdot w_{j'} \cdot A_{sj'}]^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{sj'}}{\lambda}\right) \right]^{\lambda}}{\exp(\gamma_{sL}) + \exp(\gamma_{sl}) \cdot \left[\sum_{j'=1}^2 [(1 - \tau_{sj'}) \cdot w_{j'} \cdot A_{sj'}]^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{\delta_{sj'}}{\lambda}\right) \right]^{\lambda}} \cdot N_s \quad \text{if } j = 1, 2$$

L_{sj} is the number of individuals of cell s who want to work in sector j , whereas N_s is the total number of individuals who belong to cell s . Note that the labor supply functions aggregate the preferences of individuals who belong to each cell s , given that Equations B.12 and 10 coincide.

Income taxation is progressive in South Africa. To account for the progressivity of income taxation in the CGE model, we estimate a *taxation equation* for each cell as a nonlinear relationship between income tax and labor income, as follows:

$$tax_{sj} = \omega_s^0 + \omega_s^1 \cdot w_j \cdot A_{sj} + \omega_s^2 \cdot (w_j \cdot A_{sj})^2 + \omega_s^3 \cdot (w_j \cdot A_{sj})^3 + \omega_s^4 \cdot \ln(w_j \cdot A_{sj}) \quad (\text{B.13})$$

where $j = 1$. The informal labor incomes are not taxed, thus $tax_{s2} = 0$. The coefficients are estimated separately for each cell s using the micro dataset. More concretely, we minimize the sum of squared errors subject to the constraint that the predicted total taxation coincides with the observed total taxation for each cell.

The income tax rate for each cell s is computed as follows:

$$\tau_{sj} = \frac{tax_{sj}}{w_j \cdot A_{sj}} \quad (\text{B.14})$$

In our CGE model, we consider one representative household who perceives the net labor income (formal and informal) earned by the 32 cells. This representative household owns an exogenous wealth Ψ remunerated at rate r and receives exogenous transfers Γ from the government. The disposable income of the representative household is given by:

$$Y_{net} = \sum_{sj} (1 - \tau_{sj}) \cdot w_j \cdot A_{sj} \cdot L_{sj} \cdot (1 - u_{sj}) + r \cdot \Psi + \Gamma \quad (\text{B.15})$$

where u_{sj} is the unemployment rate of cell s . We assume that people who want to work in the formal sector find a job (which implies that $u_{s1} = 0$), whereas people who want to work in the informal sector could be unemployed because of the presence of barriers to entry in the informal sector. The unemployment rate at the national level is endogenously determined in order to respect the macroeconomic equilibrium condition between investments and aggregate savings (see Section B.2.2). The unemployment rate for each cell is determined by assuming that the

share of the unemployed in each cell with respect to the total number of unemployed remains unchanged, i.e., $\frac{u_{s2} \cdot L_{s2}}{\sum_{s'} u_{s'2} \cdot L_{s'2}} = \vartheta_{s2}$.

We assume that the representative household saves an exogenous fraction s_{rate} of his disposable income and consumes the complementary fraction. The budget constraint is represented as follows:

$$P^C \cdot C = (1 - s_{rate}) \cdot Y_{net} \quad (\text{B.16})$$

where C represents the aggregate consumption and P^C the consumer price index.

The consumption level of formal and informal goods C_{ij} produced by industry i is chosen to maximize a CES utility function. The first order conditions are:

$$C_{ij} = \alpha_{ij}^C \cdot \left(\frac{P^C}{P_{ij}^C} \right)^{\sigma^C} \cdot C \quad (\text{B.17})$$

$$P^C \cdot C = \sum_{ij} P_{ij}^C \cdot C_{ij} \quad (\text{B.18})$$

where P_{ij}^C is the consumer price of good i in sector j . The consumer price index is subsequently equal to the weighted average of the consumer prices of formal and informal goods that are produced by each industry i .

B.2.1 Government

With respect to the government, revenues are given by indirect taxes on the production of formal goods and direct taxes on formal labor incomes, whereas expenditures are indicated by the public consumption of goods and services,³⁵ the interest on the public debt B and transfers to families Γ . The difference determines public savings S_G . In the model, we assume that the ratio between public savings and GDP is constant, and the total public expenditure is

³⁵According to national accounts, the public consumption only concerns industry 9 “Public administration”. The public consumption of all the other industries is subsequently fixed to zero in the model.

endogenously determined to respect the following government budget constraint:

$$S_G = \left[\sum_i \tau_i^Y \cdot P_{i1}^Y \cdot Y_{i1} + \sum_s \tau_{s1} \cdot w_1 \cdot A_{s1} \cdot L_{s1} \right] - \left[\sum_i P_{i1}^C \cdot G_i + r \cdot B + \Gamma \right] \quad (\text{B.19})$$

B.2.2 Investment

The macroeconomic equilibrium condition between aggregate investments and aggregate savings (private savings, public savings, and savings with respect to the rest of the world) is:

$$Inv = s_{rate} \cdot Y_{net} + S_G + S_{Row} \quad (\text{B.20})$$

where savings with respect to the rest of the world S_{Row} are defined in the following section.

In this paper, we adopt the Keynesian macroeconomic closure, which implies that aggregate investments are fixed at a given level:

$$Inv = \overline{Inv} \quad (\text{B.21})$$

Consequently, the macroeconomic equilibrium condition determines the equilibrium level of the unemployment rate at the national level u . In such a framework, the Keynesian closure implies that the unemployment rate, and thus, aggregate savings, adjust to respect the macroeconomic equilibrium condition.

The aggregate investment I is split out between different industries operating in the formal sector to minimize the total investment cost. The first order conditions are:

$$Inv_i = \alpha_i^{Inv} \cdot \left(\frac{P^{Inv}}{P_{i1}^C} \right)^{\sigma^{Inv}} \cdot Inv \quad (\text{B.22})$$

$$P^{Inv} \cdot Inv = \sum_i P_{i1}^C \cdot Inv_i \quad (\text{B.23})$$

where P^{Inv} is the average investment price of industry i in the formal sector.

B.2.3 International Trade and the Balance of Payments

The production of the formal good in each industry i is sold in the domestic market and exported. We assume that exported goods are identical to goods sold in the domestic market, which implies a unique selling price. The exports are defined by a demand function that negatively depends on the relative price, i.e., the ratio between the foreign price expressed in domestic currency and the domestic price:

$$E_i = \alpha_i^E \cdot \left(\frac{\bar{P}_i \cdot EXR}{P_{i1}^Y} \right)^{\sigma_i^E} \quad (\text{B.24})$$

where \bar{P}_i and EXR are the (exogenous) foreign price in industry i expressed in foreign currency and the (endogenous) exchange rate, respectively.

In contrast, the informal production is supposed to be sold only in the domestic market.

For each industry i , the total demand of the formal good is given by $\sum_{i'} Z_{ii'} + C_{i1} + G_i + Inv_i$. This total demand is satisfied by the domestic production and imports. In contrast, we assume that the total demand of informal goods is satisfied only by the domestic production. We use an Armington formulation that implies imperfect substitution between domestic D_i and foreign M_i productions due to the different origin of the products. The first order conditions minimizing the total cost are:

$$D_i = \alpha_i^D \cdot \left(\frac{P_{i1}^C}{P_{i1}^Y} \right)^{\sigma^M} \cdot \left[\sum_{i'} Z_{ii'} + C_{i1} + G_i + Inv_i \right] \quad (\text{B.25})$$

$$M_i = \alpha_i^M \cdot \left(\frac{P_{i1}^C}{\bar{P}_i \cdot EXR} \right)^{\sigma^M} \cdot \left[\sum_{i'} Z_{ii'} + C_{i1} + G_i + Inv_i \right] \quad (\text{B.26})$$

$$P_{i1}^C \cdot \left[\sum_{i'} Z_{ii'} + C_{i1} + G_i + Inv_i \right] = P_{i1}^Y \cdot D_i + \bar{P}_i \cdot EXR \cdot M_i \quad (\text{B.27})$$

For the formal goods, the consumer price P_{i1}^C of industry i is thus equal to a weighted average between the domestic price P_{i1}^Y and the foreign price ($\bar{P}_i \cdot EXR$).

The foreign capital flows S_{Row} are fixed at the initial level, whereas the exchange rate EXR

is endogenously determined to equilibrate the balance of payments:

$$S_{Row} = \sum_i P_{i1}^Y \cdot E_i - \sum_i \bar{P}_i \cdot EXR \cdot M_i \quad (\text{B.28})$$

B.2.4 Equilibrium Conditions

We assume that domestic prices are perfectly flexible and guarantee the equilibrium in each market, i.e., the markets of formal and informal goods and services of the ten industries, the two labor markets, and the capital market.

For each industry i , the production level of the formal good must coincide with the domestic and foreign demands, whereas the production level of the informal good must coincide with the private consumption:

$$Y_{i1} = D_i + E_i \quad (\text{B.29})$$

$$Y_{i2} = C_{i2} \quad (\text{B.30})$$

These equations determine the domestic equilibrium price j P_{ij}^Y for industry i in sector.

Concerning the two labor markets, the total quantity of effective units of formal and informal labor demanded by firms must coincide with the quantity supplied by individuals:

$$\sum_i L_{i1} = \sum_s L_{s1} \cdot A_{s1} \quad (\text{B.31})$$

$$\sum_i L_{i2} = \sum_s L_{s2} \cdot A_{s2} \cdot (1 - u_{s2}) \quad (\text{B.32})$$

These equations determine the equilibrium wage per unit of effective labor in the formal and informal sectors, w_1 and w_2 .

For the capital market, the (exogenous) wealth owned by the representative household Ψ is

used to finance the formal and informal capital demanded by all firms and the public debt B :

$$\sum_{ij} K_{ij} + B = \Psi \tag{B.33}$$

This equation determines the equilibrium rate of capital remuneration.

Finally, the *numéraire* chosen is the domestic producer price index, which is computed as the weighted average of domestic prices of industry i and sector j .