Tax Policies and Informality in South Africa

Eliane El Badaoui* Riccardo Magnani†

April 2013
Preliminary version

Abstract

We develop a micro-macro simulation model for evaluating tax policies in developing countries characterized with a high informal employment. Our approach allows to analyze the effects of a reform, such as a tax policy, both at the micro level, i.e. on individual's labor supply choices (formal/informal employment, work/leisure) and at the macro level (GDP, equilibrium wages, size of the informal sector, etc.). We analyze, for the South-African case, the effects of two tax reforms, the introduction of a proportional tax system and of a lump-sum tax system, instead of the current progressive tax system. We show that both reforms allow to reduce the size of the informal sector and produce a positive effect on GDP, but they increase income inequality.

JEL Classification: J22; H24; O17; C68

Key-words: Tax Reform; Informal Sector; Labor Supply; CGE Model; Microsimulation

*EconomiX - Université de Paris Ouest Nanterre - La Défense, 200 Avenue de la République, 92001 Nanterre Cedex, France. E-mail: eliane.badaoui@u-paris10.fr.

†CEPN - Université de Paris 13, 99 Avenue Jean-Baptiste Clément, 93430 Villetaneuse, France. E-mail: riccardo.magnani@univ-paris13.fr.

1 Introduction

Unlike the advanced countries, the developing countries have a large informal sector that is difficult to tax. 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 for direct taxation efficiency. This ratio is usually higher in developed countries than in poor countries (McLaren, 1998).[1] South Africa is an interesting case study. 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 represents, over the period 1997-2002, 25% of the South African GDP, the highest among middle-income countries for which the average ratio is 15%. However taxing the informal sector remains much less successful mainly because the cost of collections is too high. Thus there are still many South Africans who remain outside the tax net. According to the 2007 Labour Force Survey, 19.5% of employed workers operate in the informal sector in South Africa.[2] In countries where the labor market is characterized by the existence of a large informal sector and where governance quality is weak, the fiscal policy plays a significant role since it affects the individual's labor supply decision and sectoral choice.

In this paper, we analyze the discrete choice of individuals who decide whether not to work, to work in the formal sector or to work in the informal sector. We assume that workers compare the costs and benefits and choose the option that maximizes their utility given their preferences and job characteristics. Importantly here the informal employment choice is endogenous. More particularly, taxes affect the choice of individuals on the labor market (Johnson *et al.*, 1998; Fugazza and Jacques, 2003) and, by consequence, the size of the informal sector in the economy. Informality was traditionally viewed as an unproductive and low-paid activity sector, an option for workers who cannot find a formal job. But recent economic analysis suggests that informal employment is voluntary. Maloney (1999) considers that workers are self-selected into the informal sector because of the various benefits and opportunities that it can offer or because the comparative advantage they have in the informal sector. Subject to job availability, workers choose the sector of activity that maximizes their utility. Obviously wages and job security are

[1]For 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) while indirect taxation represents 36.4% of total tax revenue in high-income countries.

[2]At the enterprize level, several studies consider that the high entry costs into the formal economy explain the presence of a large informal sector in developing countries. Djankov *et al.* (2002) report that becoming formal in South Africa requires the completion of 7 procedures that take 30 business days, the cost of which represents 36.7% of the 1997 GDP per capita.

important determinants of the utility. However there are non-monetary advantages, like autonomy, flexibility, distance to work and working hours, which affect utility, determine job satisfaction and make desirable the informal employment.[3] Thus, depending on their characteristics, some workers choose the informal sector (Günther and Launov, 2006). Using data from Argentine, Pratap and Quintin (2006) conclude, after controlling for individual and firm characteristics, on the absence of a formal sector wage premium and thus reject the segmentation hypothesis in the labor market. 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 view of the informal activity as an attractive option.

In fact, the important distinction between the formal and the informal sector is due to tax payment, where the informal sector is the tax free option. Informality has a direct impact on public revenues and expenditures, thus it accounts for fiscal policy analysis. In the literature, several studies attempt to analyze the link between the informal sector and taxation. According to Gutmann (1977), higher taxes increase the size of the informal economy.[4] Using a simple AK-type endogenous growth model, Loayza (1996) shows that the informal sector arises when governments with weak enforcement capacity/technology impose excessive taxes and regulations. As the AK framework does not generate the transitional dynamics of the informal sector perceived in actual economies, Ihrig and Moe (2004) propose a simple dynamic model consistent with the empirical observation. The model captures the negative and convex relationship between the size of the informal sector and the real GDP per capita. Moreover, the study shows that a reduction in tax rates plays a significant role in attracting people out of the informal sector and thus improves the standard of living in the economy. The authors show for Sri Lanka that simply raising the tax rate from 9.3% to 10%, holding enforcement constant, generates more than a 2.3% increase in the informal labor in steady state.[5] Larger changes in the tax rate have even greater effects on steady state levels of the informal sector. 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. Kugler and Kugler (2009) analyze the large increase in payroll taxes that followed the social security reform in Colombia in 1993 in

[3] See, for instance, Mulinge and Mueller (1998) and Saavedra and Chong (1999).

[4] Even though reducing tax rates is not the only policy that has the potential to discourage tax evasion, the effects of tax rates are important in the evaluation of alternative tax policies.

[5] The authors define the size of the informal sector as the fraction of total labor hours devoted to informal work.

order to measure the impact of these taxes on employment and wages. They show that a 10% increase in payroll taxes reduces formal employment by between 4% and 5%.[6]

Computable General Equilibrium (CGE) and microsimulation models are widely used in policy analysis, and combining both types allows to profit from their respective advantages.[7] Accordingly, we use in this paper the micro-macro simulation approach developed by Magnani and Mercenier (2009) to analyze the impact of fiscal policies on the labor supply in poor countries. This approach uses the exact aggregation theory of Anderson, de Palma and Thisse (AdPT, 1992) and allows the researcher to avoid the complexity we find in alternative micro-macro simulation approaches. The theoretical aggregation's result by AdPT (1992) shows that individuals, who make a choice among discrete alternatives, are aggregated in a representative individual with CES preferences. We assume that individuals face a discrete choice problem and have to choose between not working, working in the formal sector or working in the informal sector. The exact aggregation procedure allows one to aggregate individual optimal decisions into explicit labor supply functions that we introduce in a CGE model. The CGE model is used to simulate the impact of different fiscal policies and, in particular, determines the macroeconomic effects of theses policies on equilibrium wages. These results are then introduced in the microsimulation model in order to evaluate the individual behavior on the labor market and the effects on the income distribution and poverty.

The paper is organized as follows. In section 2, we specify the individual discrete choice problem and present the micro analysis results. The multisectoral CGE model used in this paper is presented in Section 3. Section 4 constitutes a short description of the South African tax system. In Section 5, we analyze the micro and macro effects of tax reforms on the labor market participation and the formal/informal employment decisions. Concluding remarks are given in Section 6.

[6]For high-income countries, Davis and Henrekson (2004) show that there is a positive relation between tax rates and the informal economy. Kuehn (2010) find that higher fines, lower tax rates and stronger enforcement reduce informality. Lemieux *et al.* (1994) assume that increases in the tax rate drive people to reallocate labor from the formal to the informal sector. Using a data set collected in Québec City in Canada, they found that this relation is significant for particular groups of the population such as social-welfare claimants. Fugazza and Jacques (2004) show that lower workers taxation stimulate participation in the formal sector.

[7]On the one hand, CGE models are subject to criticisms regarding their inherent assumption that the representative agent perfectly incorporates the optimal behavior of all individuals. On the other hand, microsimulation models are carried out on a large sample of economic agents and allow taking into account the heterogeneity present in the micro-data. However, it is subject to important disapproval due to the lack of consideration of the GE effects. An integration between the macro GE models and the microsimulation models seems fundamental to avoid the weakness of these methodologies.

2 The Individual Discrete Choice Problem

2.1 Model Specification

In this paper we estimate a discrete choice labor supply model in wich individuals are assumed to choose among three options of economic activity: (1) to work in the formal wage sector, (2) to work in the informal wage sector, and (3) not to work. We specify a nested logit model (McFadden, 1981) with two nests B_k where $k \in \{L, l\}$ denotes the leisure (L) and the labor (l) status. We assume that choices depend on annual wages, and we use the Lee's (1983) generalized selectivity model to correct for sample selection bias.[8]

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

$$u_{hj} = U_{hj} + V_{hk} + \epsilon_{hj} \tag{1}$$

where h indexes individuals and $j \in \{1, 2, 3\}$ indexes, respectively, formal/informal employment activity and not to work. V_{hk} represents the upper nest specific component, U_{hj} the alternative specific component, and ϵ_{hj} the error term that is assumed to follow a generalized extreme value (GEV) distribution.

For each alternative j, the alternative specific component of u_{hj} is defined as follows:

$$\begin{aligned} U_{hj} &= \ln W_{hj}\beta + y_{hj}\delta_j \qquad j = 1, 2 \\ &= 0 \qquad\qquad\qquad\quad\;\; j = 3 \end{aligned} \tag{2}$$

where W_{hj} represents the net annual wage of individual h from alternative $j = 1, 2$, and y_{hj} is a vector of characteristics of individuals and formal/informal employment options.

The upper nest specific component is defined as follows:

$$V_{hk} = z_h\gamma_k \tag{3}$$

The micro data give information on wages earned by individuals who are employed - in the formal sector or the informal sector - and obviously we do not have any information concerning non-working individuals. Moreover, since a worker is perceived in one option (employment sector) we observe at most one wage for each person. Thus we have to im-

[8] Lee (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 modelled 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; Vijverberg, 1995).

pute the unobserved wages for all non-chosen options. More precisely, we estimate wage equations for workers in the formal and the informal sectors. Then, using the estimated wage equations, predicted wages for each person in each employment sector can be obtained. The number of wage equations is equal to the number of options for those who choose a salaried job. The wage equation can be represented as follows:

$$\ln w_{hj} = x_{hj}\alpha_j + u_{hj} \qquad \text{for } j = 1, 2 \tag{4}$$

where w_{hj} is the net hourly wage determined by observable personal and job characteristics x_{hj} and a zero-mean normally distributed error term u_{hj}. Equation (4) is a censored regression as we observe w_{hj} only if individual h is employed in the formal or the informal sector, i.e., if $k = l$. In the case where u_{hj} is correlated with ϵ_{hj}, it is necessary to correct for the selection bias.

We estimate the wage equation and use the Lee's (1983) method to correct for selection bias - selection being specified as a nested logit. To do this, we first estimate the nested model without considering the wage as an explanatory variable. Then, we compute the probability of choosing alternative j in nest B_l and we compute for each individual the value of the sample selection correction variables $\frac{-\phi(\Phi^{-1}(P_{hj}))}{P_{hj}}$. Next, we write the conditional mean of the wage equation as follows (Lee, 1983):

$$\ln w_{hj} = x_{hj}\alpha_j + \sigma_{hj}\rho_j\Big(-\frac{\phi(\Phi^{-1}(P_{hj}))}{P_{hj}}\Big) \tag{5}$$

where Φ^{-1} represents the inverse of the standard normal distribution and ϕ is the standard normal pdf. We estimate Equation (5) separately for formal and informal salaried workers. We note that negative values of the coefficients of the selection correction variables imply positive selectivity, i.e., persons who choose to work in a sector obtain - *ceteris paribus* - a higher wage than the average of total population. Once we get the parameter estimates of the wage equations, we can impute 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 for those persons who do not work, we compute potential wages in both the formal and the informal sectors.

In the final step, we reestimate Equation (1) including the individual annual wage - observed for the chosen option and imputed for non-chosen options.

We denote by P_{hj} the probability of choosing alternative j in nest B_l. Since $P_{hj} = P_{hj|B_l}.P_{hB_l}$, after McFadden we can write, respectively, 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:

$$P_{hj|B_l} \quad = \quad \frac{W_{hj}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{y_{hj}\delta_j}{\lambda}\right)}{\sum_{j\in\{1,2\}} W_{hj}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{y_{hj}\delta_j}{\lambda}\right)} \tag{6}$$

$$P_{hB_l} = \frac{\exp\left(z_h\gamma_l + \lambda I_{hl}\right)}{\exp\left(z_h\gamma_L\right) + \exp\left(z_h\gamma_l + \lambda I_{hl}\right)} \tag{7}$$

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

2.2 Data and Descriptive Statistics

We apply the theoretical framework presented above and make use of the modeling strategy expanded. To this end we use a micro data set provided by the South African Labour Force Survey (LFS). At the first stage, we model the formal and informal sector participation, and we estimate the wage in the outside option for all salaried workers. The next stage is to use the results to predict changes in workers behavior in response to a tax reform. Thus, we estimate the potential labor supply to a firm in both the formal and the informal sectors, and see the labor supply responses to fiscal policy changes.

The source of our data is the September 2000 LFS. It is a biannual household survey conducted in South Africa from february 2000 to september 2007, and designed to measure the dynamics of employment and unemployment in the country. It measures a variety of issues related to the labor market, and most importantly enable us to obtain a good measure of the informal sector. Among the different available surveys, we adopt the september 2000 LFS for our study particularly because it contains information we are able to use as proxy for the person's wealth. Another constraint comes from the macro data, thus we use the input-output data for the year 2000.

The 2000 LFS includes a population base of 105,371 on a sample of 26,648 households. Since we are interested in the outcome associated to employment status (formal and informal), an important information required from our data is that concerning remuneration. All persons in paid employment are explicitly asked for the salary amount in their main job in 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. These information allow us to compute hourly wage rates.[9] Another important information needed for our study is that associated to the distinction between formal and informal employment. The LFS asks workers explicitly whether their main job is in the formal or the informal sector. We use this subjective information to define the informal employment dummy. Moreover, the survey gives further information related to the employment status that allow us to better identify the activity sector of workers. Among these information and besides the worker's answer, we control for whether the worker has a written contract, the company where he/she works is registered or not, and deduces unemployment contributions for him/her. However, only earnings from the formal sector are considered as being 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 tax rates range from 18% to 42%. We reduce our sample to paid non self-employed workers aged 16-64 if males and 16-60 if females. After controlling for missing values, we remain with a total of 47,977 observations.

Basic summary statistics show that 30.7% of men and 21.5% of women are engaged in a salaried activity. In the selected sample, 80% are Black and 56% of people live in urban area. The average number of hours worked per week in main job is 46.8 hours. In the subsample of working people, 33.9% of women have an informal activity while only 13.3% of men are informally employed. As shown in Table 1, the log hourly net wage is on average higher in the formal than the informal sector, and men are more likely to be paid a higher wage than women. For each individual we compute wages in the outside options, i.e., in both the formal and the informal salaried job.

2.3 Estimation Results

First, in order to analyze the individual's labor supply choice, we estimate a nested logit model with selectivity. We assume that individuals choose first whether to work on a full-time basis or not. Then conditional on participation, they choose between the informal sector and the formal sector employment. The estimates are obtained in two stages and reported in Table 2. The parameter estimates of the formal-informal employment choice show that men are more likely to have a formal job than women. Moreover the probability

[9] We note however that for about 2.6% of paid workers - salaried and self-employed - reported their salary in income categories rather than exact values. For these, we use the midpoints of the intervals and the minimum of the open-ended interval.

to get a job in the formal sector increases with age and education level. Blacks are less likely to obtain an employment in the formal sector and speaking English increases the likelihood to obtain a formal job. Importantly, if the household is owner of the dwelling then the worker is less likely to work and more likely to choose the informal sector. Moreover, being the head of the household affects positively the labor decision and negatively the formal employment choice. The proportion of workers living and working in the informal sector on the total number of workers in each household affects negatively formal employment choice. Similarly, the number of workers per household - others than the person himself/herself - has a positive impact on the individual labor decision. Concerning the work-leisure decision, the existence of at least one person receiving pension benefits affects negatively the participation decision.

In the next step, we estimate the wage equations for the two sectors using OLS with correction for the selection bias. Table 3 gives the estimation results. The dependent variable is the logarithm of the net hourly wage in the formal and the informal sector, respectively. The results are qualitatively similar for both regressions. In fact, wages increase with age and education. Moreover, men have on average higher wages than women and White people more than Black ones. Concerning the Lee's variable, the estimates give a negative parameter that means there exists a positive selectivity.

Given the estimates of the wage equations, we impute hourly net wages in non-chosen options. Then, we compute annual net wages for all individuals in both the formal (W_1) and informal (W_2) sectors, and we re-estimate the formal-informal employment decision taking into account the individual formal-informal relative wage.[10] In a final step, with respect to the nested specification, we estimate the leisure-work decision. Results are presented in Table 4. The main result of the estimates is that the relative wage (W_1/W_2) affects positively participation in the formal salaried sector. In other words, wages are significant determinants of sectoral choice in that the probability to choose a formal job increases (decreases) if wages in the formal (informal) sector increase. Moreover, men and more educated people are more likely to choose the formal employment sector. However, once we control for wages, the age dummy for those who are 40 years and more has a negative and significant parameter. The notable result in the second column of Table 4 is the parameter estimate of the inclusive value. This parameter is positive and significantly different from unity which is evidence that the nested model is a realistic way to present the labor market status analyzed in this paper.

[10] We note that observed wages are used for employed persons for the chosen option. For salaried formal (informal) workers, we consider potential informal (formal) wages. And for non-working persons, we further assume that total working hours per week is equal to the average weekly working hours of total salaried population.

2.4 The representative agent formulation

The population is partitioned into $s = 1, ..., S$ cells according to some characteristics such as the sex, race, age category, education, and area of residence. In what follows, we model the decision process of individuals belonging to any of these cells. For each cell, we assume there is a large enough set N_s of statistically identical and independent individuals in each sub-population cell. This implies that all individuals belonging to the same cell have the same probability to choose option $j \in \{1, 2, 3\}$. We assume that total time endowment is the same for all persons and equals one.

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

$$
\begin{aligned}
L_{sj} &= \frac{\exp\left(\frac{U_{sj}}{\lambda}\right)}{\exp\left(\frac{U_{s1}}{\lambda}\right) + \exp\left(\frac{U_{s2}}{\lambda}\right)} \cdot \frac{\exp\left(V_{sl} + \lambda I_{sl}\right)}{\exp\left(V_{sL}\right) + \exp\left(V_{sl} + \lambda I_{sl}\right)} \cdot N_s \\
&= \frac{\exp\left(\frac{U_{sj}}{\lambda}\right)}{\exp\left(\frac{U_{s1}}{\lambda}\right) + \exp\left(\frac{U_{s2}}{\lambda}\right)} \cdot \frac{\exp\left(V_{sl}\right) \cdot \left[\exp\left(\frac{U_{s1}}{\lambda}\right) + \exp\left(\frac{U_{s2}}{\lambda}\right)\right]^{\lambda}}{\exp\left(V_{sL}\right) + \exp\left(V_{sl}\right) \cdot \left[\exp\left(\frac{U_{s1}}{\lambda}\right) + \exp\left(\frac{U_{s2}}{\lambda}\right)\right]^{\lambda}} \cdot N_s \\
&= \frac{W_{sj}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{y_{sj}\delta_j}{\lambda}\right)}{\sum_{j \in \{1,2\}} W_{sj}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{y_{sj}\delta_j}{\lambda}\right)} \\
&\cdot \frac{\exp\left(z_s \gamma_l\right) \cdot \left[\sum_{j \in \{1,2\}} W_{sj}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{y_{sj}\delta_j}{\lambda}\right)\right]^{\lambda}}{\exp\left(z_s \gamma_L\right) + \exp\left(z_s \gamma_l\right) \cdot \left[\sum_{j \in \{1,2\}} W_{sj}^{\frac{\beta}{\lambda}} \cdot \exp\left(\frac{y_{sj}\delta_j}{\lambda}\right)\right]^{\lambda}} \cdot N_s
\end{aligned} \tag{8}
$$

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

Equation 8 is the aggregate labor supply in each employment sector ($j = 1, 2$) for given wages W_{sj}. Thus, we can introduce these labor supply functions in the CGE model since they aggregate the preferences of individuals belonging to the same cell.[11].

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 data-set of 2000 provided by the OECD. The input-output table, which includes 48 industries, is aggregated into 10 industries reported in Table 1 in

[11] Magnani and Mercenier (2009) show that the same aggregate labor supply functions can be derived from the resolution of the optimization problem of a representative agent.

the Appendix.[12] The construction of the SAM (*Social Accounting Matrix*), necessary to calibrate our CGE model, is completed with data from national accounts concerning the balance of payments and the government account.

3.1 Production side

Each industry $i = 1, ..., 10$ produces two different types of good: a formal ($j = 1$) and an informal good ($j = 2$). We denote by Y_{i1} the production level of the formal good in industry i. It depends on the total quantity of intermediate goods Z_i, formal labor L_{i1} and formal capital K_{i1}. In contrast, the production level Y_{i2} of the informal good depends on the quantity of informal labor L_{i2} and informal capital K_{i2}.[13]

We use a CES function as follows:

$$Y_{i1} = \left[(\alpha_i^Z)^{\frac{1}{\sigma_i}} \cdot Z_i^{\rho_i} + (\alpha_{i1}^L)^{\frac{1}{\sigma_i}} \cdot L_{i1}^{\rho_i} + (\alpha_{i1}^K)^{\frac{1}{\sigma_i}} \cdot K_{i1}^{\rho_i}\right]^{\frac{1}{\rho_i}} \tag{9}$$

$$Y_{i2} = \left[(\alpha_{i2}^L)^{\frac{1}{\sigma_i}} \cdot L_{i2}^{\rho_i} + (\alpha_{i2}^K)^{\frac{1}{\sigma_i}} \cdot K_{i2}^{\rho_i}\right]^{\frac{1}{\rho_i}} \tag{10}$$

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

$$Z_i = \alpha_i^Z \cdot \left[\frac{P_{i1}^Y \cdot (1 - \tau_i^Y)}{P_i^Z}\right]^{\sigma_i} \cdot Y_{i1} \tag{11}$$

$$L_{i1} = \alpha_{i1}^L \cdot \left[\frac{P_{i1}^Y \cdot (1 - \tau_i^Y)}{w_1}\right]^{\sigma_i} \cdot Y_{i1} \tag{12}$$

$$K_{i1} = \alpha_{i1}^K \cdot \left[\frac{P_{i1}^Y \cdot (1 - \tau_i^Y)}{r + \delta_k}\right]^{\sigma_i} \cdot Y_{i1} \tag{13}$$

$$P_{i1}^Y \cdot (1 - \tau_i^Y) \cdot Y_{i1} = P_{i1}^Z \cdot Z_i + w_1 \cdot L_{i1} + (r + \delta_k) \cdot K_{i1} \tag{14}$$

and, for industry i producing the informal good are:

$$L_{i2} = \alpha_{i2}^L \cdot \left[\frac{P_{i2}^Y}{w_2}\right]^{\sigma_i} \cdot Y_{i2} \tag{15}$$

$$K_{i2} = \alpha_{i2}^K \cdot \left[\frac{P_{i2}^Y}{r + \delta_k}\right]^{\sigma_i} \cdot Y_{i2} \tag{16}$$

$$P_{i2}^Y \cdot Y_{i2} = w_2 \cdot L_{i2} + (r + \delta_k) \cdot K_{i2} \tag{17}$$

[12] The aggregation of industries is done with respect to the ten industries we have in the micro dataset.

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

P_{ij}^{Y} is the equilibrium price of the formal or informal good produced by industry i, w_j is the equilibrium wage per unit of effective labor for sector $j = 1, 2$, and $r + \delta$ is the equilibrium gross remuneration rate of a unit of capital. The parameter τ_i^Y represents the tax rate on the production of the formal good, and P_i^Z represents the aggregate price of the intermediate goods used by industry i in the formal sector. Capital is supposed to be perfectly mobile across sectors and industries, while formal and informal labor are assumed to be perfectly mobile across industries. These assumptions imply that there exists an equilibrium wage for the formal labor, an equilibrium wage for the informal labor and a unique equilibrium rate of remuneration of capital.

Given the total quantity of the intermediate good Z_i, each industry i (producing formal goods) chooses the optimal quantity to buy from industry i'. The first order conditions allowing to minimize the total cost are:

$$Z_{i'i} = \alpha_{Z_{i'i}} \cdot \left(\frac{P_i^Z}{P_{i'1}^C}\right)^{\sigma_i^Z} \cdot Z_i \tag{18}$$

$$P_i^Z \cdot Z_i = \sum_{i'} P_{i'1}^C \cdot Z_{i'i} \tag{19}$$

where P_{i1}^C is the average purchase price of the formal good in industry i (see subsection 3.5).

3.2 The representative agent

In our model individuals are grouped according to the age class (two classes: people aged less than 40 and people aged 40 and more), sex, education (two classes: high and low skilled), race (two classes: Black or non-Black), and area of residence (two classes: urban and rural). Thus, we consider $s = 1, ..., 32$ cells.

Each cell earns a net labor income $W_{sj} = (1-\tau_{sj}) \cdot \widehat{w}_j \cdot A_{sj}$, where $\widehat{w}_j$ is the (equilibrium) wage per unit of effective labor in sector $j = 1, 2$, τ is the average tax rate defined later (with $\tau_{sj} = 0$ for $j = 2$, and A_{sj} is the productivity of the cell s in sector j.[14] Each cell s supplies formal and informal labor according to the following labor supply functions:

[14] $(1 - \tau_{sj})\widehat{w}_j \cdot A_{sj}$ represents the average net wage that individuals belonging to the cell s earn in sector j. The value of $w_j \cdot A_{sj}$ is estimated from a wage equation using the age class, sex, education class, race, and area of residence as explanatory variables.

$$
\begin{aligned}
L_{sj} \quad &= \quad \frac{\delta_{sj} \cdot [(1-\tau_{sj}) \cdot \widehat{w}_j \cdot A_{sj}]^{\frac{\sigma_A}{\sigma_B}}}{\sum_{j'} \delta_{sj'} \cdot \left[(1-\tau_{sj'}) \cdot \widehat{w}_{j'} \cdot A_{sj'}\right]^{\frac{\sigma_A}{\sigma_B}}} \\
&\cdot \quad \frac{\phi^S \cdot \left[\sum_{j'} \delta_{sj'} \cdot \left[(1-\tau_{sj'}) \cdot \widehat{w}_{j'} \cdot A_{sj'}\right]^{\frac{\sigma_A}{\sigma_B}}\right]^{\sigma_B}}{1 + \phi^S \cdot \left[\sum_{j'} \delta_{sj'} \cdot \left[(1-\tau_{sj'}) \cdot \widehat{w}_{j'} \cdot A_{sj'}\right]^{\frac{\sigma_A}{\sigma_B}}\right]^{\sigma_B}} \cdot N_s
\end{aligned}
\tag{20}
$$

where L_{sj} is the number of individuals of cell s who work in sector j, while N_s is the total number of individuals belonging to the cell s. Note that the labor supply functions aggregate the preferences of individuals belonging to each cell s.[15] σ_A and σ_B are elasticities that come from the estimation of the discrete-choice labor supply.

Income taxation is progressive in South-Africa. In order to take into account the progressivity of income taxation in the CGE model, we estimate, for each cell, a *taxation equation*, i.e. a non-linear expression that links the income tax with the income. In particular, we use the following expression:

$$
tax_{sj} = \alpha_s^0 + \alpha_s^1 \cdot w_j \cdot A_{sj} + \alpha_s^2 \cdot (w_j \cdot A_{sj})^2 + \alpha_s^3 \cdot (w_j \cdot A_{sj})^3 + \alpha_s^4 \cdot \ln(w_j \cdot A_{sj}) \tag{21}
$$

where $j = 1$. Of course, the informal labor incomes are not taxed, implying $tax_{s2} = 0$. The coefficients are estimated separately for each cell s from the micro data-set in order to minimize the sum of the squared errors and by imposing that the total taxation predicted by the estimation coincides with the observed value for each cell.

The income tax rate is then equal to:

$$
\tau_{sj} = \frac{tax_{sj}}{w_j \cdot A_{sj}} \tag{22}
$$

Then, we consider one representative agent who perceives the net labor incomes (formal and informal) earned by the 32 cells. This representative agent owns an exogenous wealth W remunerated at the rate r and receives exogenous transfers Γ from the government. The net income earned by the representative agent is given by:

$$
Y_{net} = \sum_{sj} (1-\tau_{sj}) \cdot w_j \cdot A_{sj} \cdot L_{sj} + r \cdot W + \Gamma \tag{23}
$$

[15] Equations 20 and 8 coincide with $\delta_{sj} = \exp\left(\frac{y_{sj}\delta_j}{\lambda}\right)$, $(1-\tau_{sj}) \cdot \widehat{w}_j \cdot A_{sj} = W_{sj}$, $\sigma_A = \beta$, $\sigma_B = \lambda$, $\phi^S = \exp(z_s \gamma_l)$, and $1 = \exp(z_s \gamma_L)$.

An exogenous fraction s_{rate} of the net income is saved and the complementary fraction is consumed. The budget constraint is then:

$$P^C \cdot C = (1 - s_{rate}) \cdot Y_{net} \tag{24}$$

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

The choice of the quantity of good i of sector j, C_{ij}, is made in order to maximize a CES utility function. The first order conditions are:

$$C_{ij} = \alpha_{ij}^C \cdot \left(\frac{P^C}{P_{ij}^C}\right)^{\sigma^C} \cdot C \tag{25}$$

$$P^C \cdot C = \sum_{ij} P_{ij}^C \cdot C_{ij} \tag{26}$$

where P_{ij}^C is the consumer price of good i of sector j.

3.3 Government

Concerning the government, revenues are given by indirect taxes on the production of formal goods and direct taxation on labor incomes, while expenditures are given by the public consumption of goods and services,[16] interests on the public debt (B) and transfers to families Γ. The difference determines the public surplus (or deficit) S_G. In particular, we assume that the ratio between the public surplus (or deficit) and the GDP ratio is kept constant, and the total public expenditure is endogenously determined by the following budget constraint:

$$S_G = \left[\sum_i \tau_i^Y \cdot P_{i1}^Y \cdot Y_{i1} + \sum_s tax_{s1}\right] - \left[\sum_i P_{i1}^C \cdot G_i + r \cdot B + \Gamma\right] \tag{27}$$

3.4 Investment

The aggregate investments, in a pure neoclassical model, are determined by aggregate savings, i.e. the sum of private savings, public savings, and with respect to the rest of the world:

$$I = s_{rate} \cdot Y_{net} + S_G + S_{Row} \tag{28}$$

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

[16]National accounts indicate that the public consumption concerns only the industry 9 "Public administration". The public consumption of the other industries is then fixed at zero in the model.

The aggregate investment I must be the split out between different industries. The first order conditions in order to minimize the total investment cost are:

$$
\begin{aligned}
I_i &= \alpha_i^I \cdot \left(\frac{P^I}{P_{i1}^C}\right)^{\sigma^I} \cdot I && (29)\\
P^I \cdot I &= \sum_i P_{i1}^C \cdot I_i && (30)
\end{aligned}
$$

where P^I is the average investment price of industry i in the formal sector.

3.5 International trade and the balance of payments

The formal production for each industry i can be sold in the domestic market or exported. Goods that are exported are supposed to be identical to those sold in the domestic market, implying that the selling price is the same. Exports are defined by a demand function that negatively depends on the relative price i.e. the ratio between the domestic price and the foreign price expressed in domestic currency:

$$
E_i = \alpha_i^E \cdot \left(\frac{\bar{P}_i \cdot \varepsilon}{P_{i1}^Y}\right)^{\sigma_i^E} \tag{31}
$$

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

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

The total demand for each industry i of the formal good is given by $\sum_{i'} Z_{ii'} + C_{i1} + G_i + I_i$. This total demand can be satisfied by domestic production or by imports.[17] We use an Armington formulation implying that domestic (X) and foreign (M) productions are supposed to be not perfectly substitutes due to the different origin of the products. The first order conditions in order to minimize the total cost of the demand are:

$$
\begin{aligned}
X_i &= \alpha_i^X \cdot \left(\frac{P_{i1}^C}{P_{i1}^Y}\right)^{\sigma^M} \cdot \left[\sum_{i'} Z_{ii'} + C_{i1} + G_i + I_i\right] && (32)\\
M_i &= \alpha_i^M \cdot \left(\frac{P_{i1}^C}{\bar{P}_i \cdot \varepsilon}\right)^{\sigma^M} \cdot \left[\sum_{i'} Z_{ii'} + C_{i1} + G_i + I_i\right] && (33)\\
P_{i1}^Y \cdot X_i + \bar{P}_i \cdot \varepsilon \cdot M_i &= P_{i1}^C \cdot \left[\sum_{i'} Z_{ii'} + C_{i1} + G_i + I_i\right] && (34)
\end{aligned}
$$

[17] Instead, we suppose that the total demand for each industry i of the informal good is satisfied only by domestic production.

For the formal goods, the consumer price of industry i (P^C_{i1}) is then equal to a weighted average between the domestic price P^Y_{i1} and the foreign price $\bar{P}_i \cdot \varepsilon$.

Concerning the balance of payments, the sum between the current account and the capital account must be zero. The capital account surplus S_{Row} is fixed at the initial level and the exchange rate ε is endogenously determined to equilibrate the balance of payments:

$$S_{Row} = \sum_i P^Y_{i1} \cdot E_i - \sum_i \bar{P}_i \cdot \varepsilon \cdot M_i \tag{35}$$

3.6 Equilibrium conditions

We assume that the domestic prices are perfectly flexible and guarantee the equilibrium in each market, i.e. the markets of goods and services of the 10 industries, the labor market and the capital market.

Concerning the industry i, the production level of the formal good must coincide with the domestic and foreign demand, while the production level of the informal good must coincide with private consumption:

$$Y_{i1} = X_i + E_i \tag{36}$$

$$Y_{i2} = C_{i2} \tag{37}$$

These equations determine the domestic equilibrium price for the industry i in the sector j, P^Y_{ij}.

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

$$\sum_i L_{i1} = \sum_s L_{s1} \cdot A_{s1} \tag{38}$$

$$\sum_i L_{i2} = \sum_s L_{s2} \cdot A_{s2} \tag{39}$$

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

Concerning the capital market, the (exogenous) total quantity of capital supplied by individuals W must finance the total formal and informal capital demanded by firms and the public debt B:

$$\sum_{ij} K_{ij} + B = W \tag{40}$$

This equation determines the domestic equilibrium rate of remuneration of capital.

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

3.7 Data

The quantity of formal labor and capital used by each sector comes from the input-output data-set.[18] In order to determine the quantity of informal labor used by each sector we take into account the weight of the informality (in terms of wages) that comes from the micro data-set (see Table 2). The remuneration of informal capital by sector is determined by applying the same weights, i.e. by assuming that the intensity of the informality is the same for labor and capital in each sector.

Given that the GDP is computed by taking into account only the remuneration of formal factors, the fact that in our model we also include informal remunerations, we have to increase one (or more) components of the aggregate demand (i.e. private and public consumption, investment, and current account). We assume that all the informal remunerations of labor and capital are perceived by the representative agent and consumed.

4 The South African Tax System

For the purpose of our study, we focus our analysis on the personal income taxation in South Africa. Here we summarize its most important features. Given the year in which the sample used is collected, we apply the 2000/2001 tax rates.[19] The income tax consists of 6 brackets ranging from a rate of 18% applied to incomes less than or equal to R35,000 to 42% for incomes above R200,000 (Table A.2). The primary tax rebate, for which all taxpayers are eligible, increased from R3,710 in the 1999/2000 to R3,800 in the 2000/01 tax year, and the tax threshold for persons under 65 years increased from R19,526 to R21,111 during the same period.

The South African tax system is progressive, and the taxable unit is not the family, but the individual. In other words, the individual income tax rate structure is applicable

[18] Concerning the sector 10 "Private households", no information are available about the labor and capital remuneration in the input-output data-set. We determine the factors' remunerations by taking into account that from the micro data-set the labor remuneration in sector 10 represents 0.56% of the total formal labor remuneration.

[19] The South African fiscal year goes from the first of March to the end of February.

to all persons irrespective of marital status.[20] However, this method implies that in the presence of a progressive tax system, a fiscal discrimination remains between a one income earner family and a family with two members earning together the same taxable income. Obviously, because of the tax progressivity, this discrimination increases with the total income of the family. The South African tax legislation aims to reduce the tax gap and increase tax collections. The child rebate was removed to reduce fraud, but 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 a formal employment activity. In 2000/2001, the personal income tax accounts for R86,478 million (9.1% of GDP) and represents 39.2% of total tax revenue.

5 Tax Reforms: Macro and Micro Effects

In this section we analyze the effects of two tax reforms: in the first, noted *Reform 1*, the progressive tax system is replaced by a proportional tax system; in the second, noted *Reform 2*, it is replaced by a lump-sum tax system. In the first case, only workers in the formal sector pay taxes according to a rate equal to 12% computed in a way that guarantees that the total amount of taxes remains unchanged. In the second case, all individuals, independently on the labor choice and the income earned, pay the same amount of taxes (485 Rand) that is computed in a way that guarantees that the total amount of taxes remains unchanged.

We first analyze the macroeconomic and then the effects at the individual level.

5.1 Macroeconomic Effects on the Whole Economy

Tax policies produce important effects in the labor market and, consequently, in the whole economy. First, tax policies produce changes in the labor supply since they modify the opportunity-cost of working and of choosing the employment sector. Second, a general equilibrium effect on wages is produced. In fact, the change in the labor supply may be absorbed by the demand of firms only via the adjustment in wages.

[20] The 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 income tax rate for all individuals, the adjustments of the tax rates and income brackets, and the reduction of the number of income brackets.

Labor choices, both at the individual level and at the macro level, depend on the yearly net (of direct taxes) wage. In particular, the amount of direct taxes is modified with *Reform 1* which considers a proportional tax system implying that poor cells pay on the basis of a greater tax rate while rich cells pay on the basis of a lower tax rate. However, the amount of direct taxes becomes zero with *Reform 2*.

Concerning choices at the national level, the fraction of people who choose to work in the formal sector increases in both scenarios from 77.3% to 78.2% and 78.0%, respectively with *Reform 1* and 2. The fraction of people who choose to work decreases with *Reform 1* from 25.7% to 25.1%, while it increases with *Reform 2* to 26.1%. The same results are obtained at the microeconomic level.

Table 5 shows the main macroeconomic results. The number of workers in the formal sector decreases with *Reform 1* by 1.1% and increases with *Reform 2* by 2.8%; the number of workers in the informal sector decreases in both scenarios respectively by 6.3% and 1.4%. The total number of workers decreases by 2.3% with *Reform 1* and increases by 1.8% with *Reform 2*. However, the labor supply depends not only on the number of workers but also on their productivity. The number of units of effective labor in the formal sector increases by 0.7% with *Reform 1* and by 4.5% with *Reform 2*, while it decreases in the informal sector by 7.8% with *Reform 1* and by 3% with *Reform 2*. These effects represent the change in the labor supply at the macroeconomic level.

Wages, in the formal and informal sectors, adjust in order to guarantee the equilibrium in the labor markets. In particular, the equilibrium wage in the formal sector decreases in both scenarios, respectively by 0.4% and 1.6%, while the equilibrium wage in the informal sector increases in both scenarios, respectively by 2.5% and 1.9%.

The change in the labor demand affects the demand of capital which decreases in the formal sector and increases in the informal sector. Importantly, the level of production at the national level in the formal sector, which represents the real GDP, increases with both reforms (by 0.2% with *Reform 1* and by 2.2% with *Reform 2*) while the production in the informal sector decreases (-1.2% and -0.3% with *Reform 2*). Total production remains essentially unchanged with *Reform 1* and increases by 2% with *Reform 2*. Thus, the introduction of a lump-sum tax, implying that no distortions are produced in the labor market, induces a positive macroeconomic effect on real GDP and total (formal and informal) production.

The size of the informal sector is measured in terms of production, number of workers and of units of effective labor. Table 6 shows that in all these cases, the tax reforms provoke a (small) reduction in the size of the informal sector, especially with the *Reform*

1. In terms of production, the size of the informal sector decreases from 8.1% to 7.9% with both tax reforms, while, in terms of number of workers, the size of the informal sector decreases from 22.7% to 21.8% with *Reform 1* and to 22% with *Reform 2*.

In Table A.5, we show the detailed effects at the industry level, in the formal and informal sectors, concerning the production, the labor and capital demand.

5.2 Microeconomic Effects

We estimate the micro effects of different tax reforms on discrete labor supply choices. As shown in Table 7, we study the effect of replacing the actual 2000 income tax with a proportional tax equivalent to 12% (*Reform 1*) and a lump-sum tax (*Reform 2*) for the different cells.[21] We want to emphasize that general equilibrium effects related to simulated tax changes are considered.[22] We present in Table 7 the effects on the formal/informal employment and the work/leisure choices respectively, subsequent to tax reforms we simulate. The effect of the proportional tax on the total sample study is an decrease in the participation decision and an increase in the formal employment choice. The lump-sum tax increases the net wage in the formal sector with respect to the informal sector, thus we observe an increase in both the participation decision and the formal employment.

The main observation to make from Table 7 is that a proportional tax tends to reduce labor market participation of those categories who are poor or less productive. In fact, we observe a decrease in the decision to work for females, low-educated persons, Blacks and persons living in rural areas. The reason of these changes is that those people who are less productive face a higher taxation with *Reform 1* (12% *vs.* very low tax rates in the actual South African income taxation) and thus adjust their choices by reducing their participation. On the contrary, for males, non-Blacks and more-educated persons the effect of *Reform 1* is an increase in their labor market participation. Concerning the lump-sum tax (*Reform 2*), the direct taxation becomes nil. Thus individuals are induced to participate more, except in those cells where the income level is very low, i.e., females, low-educated and Blacks, implying that the elimination of the direct taxation has no effect on net labor incomes.

Concerning the formal/informal employment decision, the increase in the share of the formal employment is expected to be greater with *Reform 2* than *Reform 1*. This is not the case since the sector choice is conditional to the decision to participate. For those who were initially and still employed, the probability to choose a formal job increases. The main concern is that of people who modify their participation decision. In particular, those

[21]For more detailed results across cells, see Table A.6.

[22]The results presented do not concern self-employed workers who could fit in the category of taxpayers.

who loose their jobs are those who are more likely to have an informal job before the tax reform. Similarly, workers who increase their participation subsequently to tax reforms have a higher probability to choose an informal job. For low-productive individuals, given that they reduce their participation with *Reform 1*, but not with *Reform 2*, we observe a reduction in the informal employment implying that the proportion of formal employment increases more with *Reform 1*.

The analysis in terms of inequality concerns the household annual net wages while taking into consideration equivalence scales. In Table 8, we present the Gini index and the Theil index for the total sample as well as for the different cells in the initial case and after the tax reforms. We notice an overall increase of inequality with the proportional tax and the lump-sum tax. This increase in inequality is observed for all the cells.

6 Conclusion

In this paper, we present a micro-macro simulation approach to analyze the impact of different tax policies on the labor supply in developing countries characterized by a high level informality and a low level of participation. Our micro-macro model is built using the approach developed by Magnani and Mercenier (2009) which is based on the aggregation theory of Anderson, de Palma and Thisse (1992) and permits to aggregate the individual choices when individuals make discrete choices.

We analyze, for the South-African case, the discrete choice problem of individuals who have to decide whether not to work, to work in the formal sector or to work in the informal sector, and how tax policies affect these choices. In particular, we analyzed two types of tax reform. The first in which the current progressive tax system is replaced by a proportional tax system; the second in which the it is replaced by a lump-sum tax system.

We quantify the effects of these reforms at the individual level, i.e. on the labor market choice of each individual, and at the macro level. The main results are that both reforms produce a positive effect at the macroeconomic level in terms of GDP and total production, i.e. including informal activities, and in terms of reduction of the size of the informal sector, but a negative effect in terms of income inequality.

References

[1] Anderson S., A. de Palma and J-F. Thisse, 1992. Discrete Choice Theory of Product Differentiation. Cambridge: MIT Press.

[2] Blomquist, N. S. and U. Hansson-Brusewitz, 1990. "The Effect of Taxes on Male and Female Labor Supply in Sweden", *Journal of Human Resources*, Vol. 25, No. 3, pp. 317-357.

[3] Davis, S. and M. Henrekson, 2004. "Tax Effects on Work Activity, Industry Mix and Shadow Economy Size: Evidence from Rich-Country Comparisons", Ratio Working Papers 57, The Ratio Institute.

[4] El Badaoui, E., E. Strobl and F. Walsh, 2008. "Is There an Informal Employment Wage Penalty? Evidence from South Africa", *Economic Development and Cultural Change*, Vol. 56, pp. 683-710.

[5] Fugazza, M. and J-F. Jacques, 2004. "Labor market institutions, taxation and the underground economy", *Journal of Public Economics*, Vol. 88, Issues 1-2, pp. 395-418.

[6] Gong, X. and A. van Soest, 2002. "Wage differentials and mobility in the urban labour market: A panel data analysis for Mexico", *Labour Economics*, vol. 9, No. 4, pp. 513-529.

[7] Gong, X., A. van Soest and E. Villagomez, 2004. "Mobility in the urban labor market: A panel data analysis for Mexico", *Economic Development and Cultural Change*, Vol. 53, No. 1, pp. 1-36.

[8] Günther, I. and A. Launov, 2006. "Competitive and Segmented Informal Labour Markets", IZA Discussion Paper 2349.

[9] Gutmann, P.M., 1977. "The subterranean economy", *Financial Analysts Journal*, Vol. 34, pp. 26-27.

[10] Ihrig, J. and K. Moe, 2004. "Lurking in the shadows: the informal sector and government policy", *Journal of Development Economics*, Vol. 73, No. 2, pp. 541-557.

[11] Johnson, S., D. Kaufmann, and P. Zoido-Lobaton. 1998. "Regulatory Discretion and the Unofficial Economy", *American Economic Review*, Vol. 88, pp. 387-392.

[12] Kuehn, Z., 2010. "Labor Market Effects of Payroll Taxes in Developing Countries: Evidence from Colombia", Imdea Working Paper Series in Economics and Social Sciences, 2010/07.

[13] Kugler, A. and M. Kugler, 2009. "Tax Rates, Governance, and the Informal Economy in High-Income Countries", *Economic Development and Cultural Change*, Vol. 57, No. 2, pp. 335-358.

[14] Lemieux, T., B. Fortin and P. Fréchette, 1994. "The effect of Taxes on Labor

Supply in the Underground Economy", *The American Economic Review*, Vol. 84, No. 1, pp. 231-254.

[15] Loayza, N., 1996. "The economics of the informal sector: a simple model and some empirical evidence from Latin America", *Carnegie-Rochester Conference Series on Public Policy*, Vol. 45, No. 1, pp. 129-162.

[16] McLaren, J., 1998. "Black Markets and Optimal Evadable Taxation", *The Economic Journa*l, Vol. 108, No. 448, pp. 665-679.

[17] Magnani, R. and J. Mercenier, 2009. "On linking microsimulation and computable general equilibrium models using exact aggregation of heterogeneous discrete-choice making agents", *Economic Modelling*, Vol. 26, No. 3, pp. 560-570.

[18] Maloney, W., 1999. "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico", *World Bank Economic Review*, Vol. 13, No. 2, pp. 275-302.

[19] Maloney, W., 2004. "Informality Revisited", *World Development*, Vol. 32, No. 7, pp. 1159-1178.

[20] Mulinge, M. and C. W. Mueller, 1998. "Employee Satisfaction in Developing Countries: the case of Kenya", *World Development*, Vol. 26, No. 12, pp. 2181-2199.

[21] Saavedra, J. and A. Chong, 1999. "Structural Reform, Institutions and Earnings: Evidence from the Formal and Informal Sectors in Urban Peru", *The Journal of Development Studies*, Vol. 35, No. 4, pp.

[22] Saracoglu, D. S., 2008. "The informal sector and tax on employment: A dynamic general equilibrium investigation", *Journal of Economic Dynamics and Control*, Vol. 32, No. 2, pp. 529-549.

[23] Tanzi, V., 1987. Quantitative characteristics of the tax systems of developing countries. In Newbery and Stern.

[24] Tanzi, V. and H. Zee, 2000. "Tax Policy for Emerging Markets - Developing Countries", IMF Working Papers, No. 00/35.

Table 1: Summary Statistics

	Men		Women	
	Mean	***SD***	***Mean***	***SD***
Age	30.396	(13.153)	30.613	(12.132)
White	0.050	(0.218)	0.050	(0.218)
Black	0.806	(0.395)	0.804	(0.397)
Asian	0.021	(0.142)	0.020	(0.140)
Coloured	0.112	(0.316)	0.112	(0.315)
English	0.055	(0.228)	0.054	(0.226)
Afrikaans	0.146	(0.353)	0.149	(0.356)
Single	0.296	(0.458)	0.339	(0.473)
Married	0.012	(0.110)	0.043	(0.203)
No education	0.073	(0.260)	0.079	(0.270)
Nursery	0.002	(0.046)	0.002	(0.044)
Primary	0.213	(0.409)	0.188	(0.391)
Secondary	0.663	(0.473)	0.689	(0.463)
NTC	0.008	(0.091)	0.003	(0.055)
Diploma	0.041	(0.198)	0.039	(0.193)
Urban area	0.563	(0.496)	0.562	(0.496)
Employment dummy	0.307	(0.461)	0.215	(0.411)
Informal dummy	0.133	(0.339)	0.339	(0.473)
Log hourly gross wage	1.704	(1.108)	1.342	(1.045)
Log hourly net wage	1.667	(1.048)	1.326	(1.011)
Formal (log) hourly wage	1.781	(1.021)	1.574	(1.006)
Informal (log) hourly wage	0.927	(0.910)	0.842	(0.832)
Weekly working hours	48.881	(14.519)	44.411	(15.922)

Table 2: Formal-Informal Employment and Work-Leisure Choices

	Formal Dummy		*Work Dummy*	
Male dummy	1.270	(0.071)	0.144	(0.058)
Age	0.056	(0.016)	0.362	(0.007)
Age squared	-0.001	(0.000)	-0.004	(0.000)
White	1.360	(0.250)	-0.196	(0.089)
Black	-0.733	(0.173)	-0.069	(0.104)
English	0.805	(0.216)	0.199	(0.113)
Afrikaans	-0.326	(0.172)	0.228	(0.100)
Single	0.218	(0.055)	0.183	(0.031)
Married	0.046	(0.116)	-0.281	(0.071)
Nursery	0.576	(0.445)	0.163	(0.255)
Primary	0.120	(0.080)	0.053	(0.049)
Secondary	0.808	(0.081)	-0.097	(0.055)
NTC	2.707	(1.022)	-0.502	(0.211)
Diploma	2.409	(0.254)	0.001	(0.125)
Region	0.175	(0.069)	0.165	(0.037)
Urban area	-0.317	(0.055)	-0.196	(0.031)
House owner	-0.386	(0.053)	-0.857	(0.031)
Children 0-3 years	0.092	(0.042)	-0.149	(0.021)
Children 4-6 years	0.058	(0.048)	-0.116	(0.024)
Children 7-12 years	0.083	(0.031)	-0.167	(0.016)
Household head	-0.322	(0.059)	1.554	(0.034)
Province unemployment rate	-0.037	(0.008)	-0.024	(0.004)
Household informal workers	-4.135	(0.196)	–	–
Household workers number	–	–	0.406	(0.017)
Pension dummy	–	–	-0.239	(0.042)
Inclusive Value	–	–	0.078	(0.048)
Constant	1.157	(0.412)	-6.929	(0.218)
Observations	**12,309**		**47,948**	

(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) The "household informal workers" variable indicates the proportion of informal workers among employed persons within the household.

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

	Formal Sector		***Informal Sector***	
Male dummy	0.279	(0.019)	0.120	(0.034)
Age	0.078	(0.006)	0.064	(0.010)
Age squared	-0.001	(0.000)	-0.001	(0.000)
White	0.657	(0.031)	0.922	(0.199)
Black	-0.234	(0.054)	-0.138	(0.109)
English	0.300	(0.057)	0.158	(0.134)
Afrikaans	-0.126	(0.054)	-0.168	(0.107)
Nursery	0.101	(0.154)	-0.014	(0.284)
Primary	0.198	(0.031)	0.077	(0.046)
Secondary	0.552	(0.030)	0.173	(0.048)
NTC	1.004	(0.093)	0.643	(0.800)
Diploma	1.128	(0.043)	0.369	(0.200)
Region	0.237	(0.018)	0.533	(0.039)
Urban area	0.549	(0.018)	0.117	(0.033)
Selection Correction	-0.033	(0.023)	-0.132	(0.042)
Constant	-1.019	(0.143)	-0.900	(0.259)
F	**546.98**		**30.85**	
$\mathbf{R}^2$	**0.46**		**0.14**	
Observations	**9,516**		**2,793**	

(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.

Table 4: Labor Market Choices Depending on Wages

	Formal Dummy		***Work Dummy***	
Male dummy	1.242	(0.070)	-0.354	(0.043)
Age dummy	-0.229	(0.056)	-0.011	(0.032)
Education dummy	0.664	(0.056)	-0.408	(0.034)
Black	-0.918	(0.175)	0.299	(0.098)
Urban area	-0.495	(0.057)	-0.046	(0.028)
English	0.941	(0.218)	-0.091	(0.105)
Afrikaans	-0.304	(0.175)	0.297	(0.096)
Single	0.264	(0.053)	0.640	(0.029)
Married	0.030	(0.117)	-0.145	(0.068)
Region	0.335	(0.069)	0.140	(0.097)
House owner	-0.369	(0.053)	-0.800	(0.028)
Children 0-3 years	0.069	(0.042)	-0.169	(0.020)
Children 4-6 years	0.048	(0.049)	-0.099	(0.023)
Children 7-12 years	0.086	(0.031)	-0.157	(0.015)
Household head	-0.313	(0.057)	1.902	(0.031)
Province unemployment rate	-0.023	(0.008)	-0.022	(0.004)
Household informal workers	-4.358	(0.199)	–	–
$Ln(W_1/W_2)$	0.517	(0.032)	–	–
Household workers number	–	–	0.371	(0.016)
Pension dummy	–	–	-0.052	(0.041)
Inclusive Value	–	–	0.298	(0.025)
Constant	1.604	(0.472)	-2.086	(0.284)
Observations	**12,309**		**47,948**	

(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) Age dummy equals one for those people aged 40 and more; (iv) The education dummy indicates whether people are high educated (secondary education and more).

Table 5: Main Macroeconomic Effects of Tax Reforms

	Formal Sector			Informal Sector			Total Economy		
	Initial	***Reform 1***	***Reform 2***	***Initial***	***Reform 1***	***Reform 2***	***Initial***	***Reform 1***	***Reform 2***
Production	922,147	0.2%	2.2%	81,039	-2.5%	-0.3%	1,003,186	0.0%	2.0%
Units of labor	22.45	-1.1%	2.8%	6.59	-6.3%	-1.4%	29.04	-2.3%	1.8%
Units of effective labor	4,201	0.7%	4.5%	932	-7.8%	-3.0%	5133	-0.8%	3.1%
Capital	2,496,986	-0.3%	-0.3%	255,600	3.3%	2.6%	2,752,586	0.0%	0.0%
Real wage	105.3	-0.4%	-1.6%	45.7	2.5%	1.9%			
Gross capital remuneration rate							15.0%	0.0%	0.1%
Deficit / GDP							2.10%	0.0%	0.1%
Net labor incomes	387,529	0.4%	17.4%	42,578	-5.5%	-1.1%	430,107	0.2%	2.8%
Total net incomes							723,078	0.0%	2.1%

Table 6: Weight of the Informal Sector

	Initial Value	***Reform 1***	***Reform 2***
% of production	8.1%	7.9%	7.9%
% of workers	22.7%	21.8%	22.0%
% of units of effective labor	8.8%	8.1%	8.2%

Table 7: Effects of Tax Reforms on Individual Labor Market Choices

		% Formal Employment			% Work		
		Initial	***Reform 1***	***Reform 2***	***Initial***	***Reform 1***	***Reform 2***
All		0.773	0.782	0.780	0.257	0.251	0.261
Age	15-39	0.802	0.802	0.806	0.209	0.206	0.214
	> 40	0.727	0.751	0.738	0.402	0.388	0.405
Sex	Females	0.661	0.680	0.673	0.215	0.196	0.215
	Males	0.867	0.868	0.870	0.307	0.316	0.317
Education	Low education	0.665	0.687	0.674	0.318	0.292	0.315
	High education	0.830	0.831	0.835	0.233	0.235	0.241
Race	Non black	0.904	0.908	0.908	0.398	0.413	0.410
	Black	0.717	0.729	0.725	0.223	0.211	0.225
Zone	Rural	0.759	0.764	0.763	0.229	0.218	0.231
	Urban	0.782	0.794	0.791	0.279	0.276	0.285

Table 8: Inequality Indexes

		Gini Index			Theil Index		
		Initial	***Reform 1***	***Reform 2***	***Initial***	***Reform 1***	***Reform 2***
All		0.877	0.892	0.893	0.491	0.635	0.677
Age	15-39	0.897	0.909	0.909	0.467	0.610	0.649
	> 40	0.813	0.839	0.840	0.520	0.663	0.710
Sex	Females	0.896	0.914	0.911	0.492	0.629	0.679
	Males	0.847	0.862	0.866	0.454	0.609	0.637
Education	Low education	0.822	0.841	0.837	0.340	0.394	0.434
	High education	0.882	0.895	0.896	0.438	0.585	0.614
Race	Non black	0.804	0.823	0.828	0.448	0.595	0.615
	Black	0.881	0.895	0.893	0.386	0.486	0.525
Zone	Rural	0.875	0.887	0.886	0.381	0.491	0.518
	Urban	0.858	0.874	0.875	0.428	0.560	0.595

Appendix

Table A.1: Proportion of Informal Sector Remuneration by Industry

1	Agriculture, hunting, forestry and fishing	4.02%
2	Mining and quarrying	0.07%
3	Manufacturing	0.59%
4	Electricity, gas and water supply	0.02%
5	Construction	13.15%
6	Wholesale and retail trade	1.47%
7	Transport and communication	4.43%
8	Financial intermediation, insurance, real estate	0.12%
9	Public administration	0.19%
10	Private households	71.91%

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

Taxable Income	Rates of Tax	Proportion
R1 - R35 000	18% of each R1	78.56%
R35 001 - R45 000	R6 300 + 26% of the amount above R35 000	5.99%
R45 001 - R60 000	R8 900 + 32% of the amount above R45 000	6.30%
R60 001 - R70 000	R13 700 + 37% of the amount above R60 000	2.00%
R70 001 - R200 000	R17 400 + 40% of the amount above R70 000	6.47%
R200 001 - and above	R69 400 + 42% of the amount above R200 000	0.71%
Individual Primary Rebates	R3 800	
Tax Threshold under 65 years	R21 111	

Source: South African Revenue Service (*SARS*).
Among the 78.56%, 62.59% does not pay taxes due to rebates and tax threshold.
We note that the tax brackets and tax thresholds are adjusted on an annual basis.

Table A.3: Personal Income Tax Relative to Total Tax Revenue

Tax Year	PIT Revenue	% of Total	Total Tax Revenue
1994/95	44 972.8	39.5%	113 774.5
1995/96	51 179.3	40.2%	127 278.0
1996/97	59 519.8	40.4%	147 332.3
1997/98	68 342.4	41.3%	165 327.4
1998/99	77 733.9	42.1%	184 843.6
1999/00	85 883.8	42.6%	201 386.0
2000/01	86 478.0	39.2%	220 334.2
2001/02	90 389.5	35.8%	252 298.3
2002/03	94 336.7	33.4%	282 209.7
2003/04	98 495.1	32.6%	302 507.5
2004/05	110 981.9	31.3%	354 980.3
2005/06	125 645.3	30.1%	417 334.0
2006/07	140 578.3	28.4%	495 515.1
2007/08	168 774.4	29.5%	572 870.6

Note: (i) PIT denotes *Personal Income Tax*; (ii) Revenues are in million of Rands.

Table A.4: Total Tax Revenue and Cost of Collection

Tax Year	Tax Revenue Collected	Operating Cost	Operating Cost as a % of Tax
1998/99	184 843.6	1 663.2	0.90%
1999/00	201 386.0	1 906.3	0.95%
2000/01	220 334.2	1 922.5	0.87%
2001/02	252 298.3	2 332.6	0.92%
2002/03	282 209.7	2 878.5	1.02%
2003/04	302 507.5	3 562.8	1.18%
2004/05	354 980.3	4 311.7	1.21%
2005/06	417 334.0	5 135.5	1.23%
2006/07	495 515.1	5 156.1	1.04%
2007/08	572 870.6	5 607.9	0.98%

Note: (i) Operating cost includes revenue transferred from National Treasury as well as other operating revenue; (ii) Revenues are in million of Rands.

Table A.5: Macroeconomic Effects of Tax Reforms by Industries

	Formal Sector			Informal Sector		
	Initial Value	***Reform 1***	***Reform 2***	***Initial Value***	***Reform 1***	***Reform 2***
PRODUCTION						
Agriculture, hunting, forestry and fishing	39,049	0.1%	1.2%	4,818	-1.7%	0.1%
Mining and quarrying	63,193	0.2%	2.1%	148	-2.3%	-0.2%
Manufacturing	238,431	0.2%	2.0%	4,173	-2.6%	-0.4%
Electricity, gas and water supply	27,750	0.1%	1.5%	22	-1.7%	0.1%
Construction	24,335	0.3%	1.7%	12,522	-2.9%	-0.5%
Wholesale and retail trade	120,620	0.2%	2.2%	5,759	-2.4%	-0.3%
Transport and communication	76,942	0.1%	1.8%	10,833	-2.2%	-0.1%
Financial intermediation, insurance, real estate	154,491	0.1%	1.7%	624	-1.8%	0.1%
Public administration	172,743	0.4%	3.4%	897	-4.0%	-1.1%
Private households	4,592	0.4%	3.5%	41,244	-2.6%	-0.4%
Total	922,147	0.2%	2.2%	81,039	-2.5%	-0.3%
UNITS OF EFFECTIVE LABOR						
Agriculture, hunting, forestry and fishing	115	0.8%	4.5%	37	-9.0%	-3.6%
Mining and quarrying	267	0.8%	4.7%	2	-8.1%	-3.2%
Manufacturing	1,023	0.8%	4.4%	49	-7.8%	-3.0%
Electricity, gas and water supply	87	0.8%	4.7%	0	-9.0%	-3.6%
Construction	128	0.7%	3.8%	165	-7.4%	-2.8%
Wholesale and retail trade	523	0.8%	4.7%	63	-8.0%	-3.1%
Transport and communication	250	0.8%	4.6%	105	-8.4%	-3.3%
Financial intermediation, insurance, real estate	486	0.8%	4.8%	5	-8.9%	-3.6%
Public administration	1,298	0.6%	4.2%	17	-5.7%	-1.9%
Private households	24	0.9%	5.8%	489	-7.8%	-3.0%
Total	4,201	0.7%	4.5%	932	-7.8%	-3.0%

Table A.5: Macroeconomic Effects of Tax Reforms by Industries (*cont.*)

	Formal Sector			Informal Sector		
	Initial Value	***Reform 1***	***Reform 2***	***Initial Value***	***Reform 1***	***Reform 2***
Capital						
Agriculture, hunting, forestry and fishing	151,284	-0.3%	-0.4%	20,873	2.1%	2.0%
Mining and quarrying	201,804	-0.3%	-0.2%	510	3.1%	2.5%
Manufacturing	607,622	-0.4%	-0.4%	12,732	3.5%	2.7%
Electricity, gas and water supply	113,118	-0.3%	-0.2%	95	2.1%	2.1%
Construction	59,453	-0.4%	-1.1%	33,149	3.9%	2.9%
Wholesale and retail trade	362,959	-0.3%	-0.2%	19,061	3.2%	2.6%
Transport and communication	220,399	-0.3%	-0.3%	40,115	2.8%	2.4%
Financial intermediation, insurance, real estate	608,398	-0.3%	-0.1%	2,659	2.2%	2.1%
Public administration	157,972	-0.5%	-0.6%	883	5.8%	3.8%
Private households	13,977	-0.2%	0.9%	125,523	3.5%	2.7%
Total	2,496,986	-0.3%	-0.3%	255,600	3.3%	2.6%

Table A.6: Effects of Tax Reforms on Labor Market Choices

					% Formal Employment			% Work		
Age	***Sex***	***Education***	***Race***	***Zone***	***Initial***	***Reform 1***	***Reform 2***	***Initial***	***Reform 1***	***Reform 2***
15-39	Females	Low education	Non black	Rural	0.806	0.820	0.810	0.518	0.520	0.524
15-39	Females	Low education	Non black	Urban	0.556	0.586	0.569	0.254	0.216	0.250
15-39	Females	Low education	Black	Rural	0.608	0.623	0.617	0.176	0.146	0.172
15-39	Females	Low education	Black	Urban	0.367	0.395	0.387	0.208	0.149	0.188
15-39	Females	High education	Non black	Rural	0.880	0.883	0.883	0.401	0.416	0.412
15-39	Females	High education	Non black	Urban	0.933	0.927	0.933	0.295	0.317	0.310
15-39	Females	High education	Black	Rural	0.702	0.684	0.702	0.090	0.079	0.092
15-39	Females	High education	Black	Urban	0.664	0.679	0.678	0.166	0.148	0.167
15-39	Males	Low education	Non black	Rural	0.947	0.952	0.949	0.804	0.812	0.809
15-39	Males	Low education	Non black	Urban	0.810	0.820	0.815	0.309	0.309	0.316
15-39	Males	Low education	Black	Rural	0.788	0.794	0.790	0.280	0.276	0.285
15-39	Males	Low education	Black	Urban	0.734	0.748	0.743	0.236	0.223	0.242
15-39	Males	High education	Non black	Rural	0.942	0.944	0.944	0.592	0.607	0.601
15-39	Males	High education	Non black	Urban	0.956	0.953	0.957	0.403	0.427	0.418
15-39	Males	High education	Black	Rural	0.820	0.803	0.817	0.144	0.146	0.151
15-39	Males	High education	Black	Urban	0.896	0.886	0.898	0.229	0.247	0.241
> 40	Females	Low education	Non black	Rural	0.691	0.718	0.698	0.492	0.475	0.495
> 40	Females	Low education	Non black	Urban	0.432	0.471	0.456	0.254	0.201	0.240
> 40	Females	Low education	Black	Rural	0.517	0.552	0.530	0.285	0.237	0.276
> 40	Females	Low education	Black	Urban	0.332	0.382	0.352	0.400	0.325	0.381
> 40	Females	High education	Non black	Rural	0.877	0.889	0.888	0.335	0.352	0.349
> 40	Females	High education	Non black	Urban	0.916	0.919	0.922	0.295	0.322	0.315
> 40	Females	High education	Black	Rural	0.568	0.599	0.582	0.291	0.251	0.286
> 40	Females	High education	Black	Urban	0.606	0.645	0.625	0.412	0.383	0.412
> 40	Males	Low education	Non black	Rural	0.947	0.952	0.948	0.846	0.853	0.850
> 40	Males	Low education	Non black	Urban	0.754	0.773	0.766	0.348	0.344	0.357
> 40	Males	Low education	Black	Rural	0.780	0.797	0.787	0.454	0.452	0.460
> 40	Males	Low education	Black	Urban	0.761	0.785	0.772	0.442	0.439	0.449
> 40	Males	High education	Non black	Rural	0.952	0.957	0.954	0.709	0.723	0.718
> 40	Males	High education	Non black	Urban	0.958	0.963	0.964	0.529	0.557	0.550
> 40	Males	High education	Black	Rural	0.800	0.814	0.808	0.385	0.386	0.392
> 40	Males	High education	Black	Urban	0.915	0.921	0.921	0.472	0.492	0.486