Dualistic Sector Choice and Female Labour Supply: Evidence from Formal and Informal Sectors in Cameroon

Gauthier Lanot and Christophe Muller

WPS/97-9

June 1997
(Version 4.7)

Centre for the Study of African Economies
Institute of Economics and Statistics
University of Oxford
St Cross Building
Manor Road
Oxford OX1 3UL

Gauthier Lanot is at Keele University, UK. Christophe Muller is at CSAE, University of Oxford; his e-mail address is christophe.muller@economics.ox.ac.uk.

This research has been funded by the Programme d'Analyses et de Recherches Economiques Appliquées au Développement International (PARADI) at the Université Laval (for the academic years 1993–94 for the first author, and 1993-94 and 1994-95 for the second author). The second author acknowledges also a TMR grant from the European Union which has allowed the completion of this project. We would like to thank participants in seminars at Journées PARADI and Université Laval in Canada; at CREST and DIAL in Paris; and at the Universities of Oxford, Keele and Dundee in the UK; and particularly I. Walker, D. Margolis, B. Fortin, T. Lemieux, F. Teal and S. Bond. We are grateful to DIAL–DSCN which made the data available to us.

Abstract: In developing countries, labour supply and activity choices are distorted by the existence of labour market imperfections restricting entry to the activity sectors or rationing the worked hours. The presence of decreasing returns to labour in the informal sector is another specific characteristic of labour market dualism in LDCs. Because of the existence of entry costs, hours rationing in the formal sector and the non-linear earnings function in the informal sector that is quasi-autarkic in labour, the usual separation theorems of the standard labour supply models do not apply. Using data from a sample of women in Yaoundé (Cameroon) in 1993, choosing their activity in different sectors, we show that formal and informal labour markets have features associated with dualism and market imperfections. By contrast with the simplified models in the literature, we estimate a non-linear non-separable integrated model of activity choice and labour supply of female workers, jointly with market imperfections. We identify the effects of explanatory variables of the preferences and the earnings functions in this simultaneous structural framework. The results show that the earnings function in the informal sector is strictly concave in hours, and positively related to the education and experience of workers and the capital of the family firm. The presence of young children has a negative impact on the latent labour supply whereas that of other female members makes easier the supply of labour which is consistent with the existence of substitutabilities in the domestic chores. Other children, age and marital status of the worker also influence the latent labour supply.

Résumé

Dans les pays en développement, l'offre de travail et les choix d'activité sont perturbés par la présence d'imperfections de marchés restreignant l'entrée dans les secteurs d'activité ou rationnant les heures travaillées. L'existence de rendements décroissants du travail dans le secteur informel est un autre trait spécifique du dualisme du marché du travail dans les PVD. En raison de l'existence de coûts d'entrée, de rationnement des heures dans le secteur formel et de fonctions de gain nonlinéaires dans le secteur informel quasi-autarcique en travail, les théorèmes usuels de séparation des modèles standards d'offre de travail ne s'appliquent pas.

A partir d'informations pour un échantillon de femmes à Yaoundé (Cameroun) en 1993, choisissant leur activité dans différents secteurs, nous montrons que les marchés du travail formels et informels ont des caractéristiques associées avec dualisme et imperfections de marché. A la différence des modèles simplifiés dans la littérature, nous estimons un modèle nonlinéaire nonséparable intégré de choix d'activité et d'offre de travail des travailleurs féminins, de façon jointe à des imperfections de marché. Nous identifions les effets des variables explicatives des préférences et des fonctions de gain dans ce cadre structurel simultané. Les résultats montrent que la fonction de gain dans le secteur informel est strictement concave par rapport aux heures, et positivement reliée à l'éducation et l'expérience du travailleur, et au capital de l'entreprise familiale. La présence de jeunes enfants a un impact négatif sur l'offre de travail latente, alors que celle d'autres membres féminins facilite l'offre de travail, ce qui est cohérent avec l'existence de substitutabilités dans les taches ménagères. Les autres enfants, l'age et le statut matrimonial influencent également l'offre de travail latente.

Contents

1. Introduction 1

2. The structure of the dualistic activity choice models 2
   2.1 The dualistic context 2

3. Data description 6
   3.1 Descriptive statistics 6
   3.2 Wage and non-labour income equations 9
   3.3 The participation model 11

4. The Model 12
   4.1 Definition of regimes 13
   4.2 The specification of the utility function 14
   4.3 Specification of the earnings function in the informal sector 15

5. Estimation of the Structural Model 16
   5.1 Some econometric difficulties 16
   5.2 The likelihood function and the exogenous variables 17
   5.3 Estimation results 18

6. Conclusion 22

Bibliography 23

Tables 25

Annexe 1 31

1. Introduction

Labour markets in developing countries are often characterised by dualism (Lewis (1954), Ranis and Fei (1961), Myint (1985)) rather than perfect competition. The existence of activities with diminishing return to labour in the traditional sector[1] and entry costs in the modern sector are a specific feature of this dualism, especially when agricultural production or the informal sector is involved (Ghatak and Ingersent (1984)). Often, the fundamental opposition is between a commercialised versus a non-commercialised context. The existence of a large gap between real wages in both sectors induces traditional workers to migrate into the modern sector. When modelling the labour supply or the activity of workers, the decreasing return to labour in the traditional sector naturally suggests agricultural household models (Sen (1966), Singh, Squire and Strauss (1986), Gronau (1986)), in which the rewards resulting from different technologies or markets can be compared. Several attempts at estimating such models have been carried out in the context of labour markets (Benjamin (1992), Jacoby (1993), Lemieux, Fortin and Fréchette (1994), Newman and Gertler (1994)).

The aim of this article is to extend the study of activity choice and labour supply models based on comparison of wage rates, to more general cases involving fixed costs, hours rationing and non-linear earning functions, for which the Roy model (Roy (1951)) and the separation of decisions are then no longer valid. Our general framework is particularly relevant for the study of dualistic labour markets in LDCs, and notably for the urban markets composed of formal and informal sectors. The emphasis of this paper is to propose a further step in the neoclassical modelling of non-separable labour behaviour under market imperfections.[2]

We present the structure of the dualistic activity choice and labour supply models in Section 2. We describe the sample of female workers in Yaoundé, capital of Cameroon, in 1993, and their activity sectors in Section 3, and we show the estimation results of participation choices. The labour market in Yaoundé displays dualistic features typical of developing countries, due to the importance of the informal sector. We describe the integrated non-separable model in Section 4. We derive in section 5 the likelihood of the structural model, discuss the explanatory variables and present the estimation results. We conclude in Section 6.

2. The structure of the dualistic activity choice models

2.1 The dualistic context

Labour markets in developing countries are generally characterised by dualism and market imperfections rather than perfect competition. On the one side, Ricardo (1815) first, then later Lewis (1954) and Ranis and Fei (1961) described the labour market in developing countries in a dualistic framework where a traditional sector with a low and variable return to labour and a low level of capital stock coexists with a modern sector with high wages and substantial accumulation

[1] Or even, as in the organisational dualism theory, with real wage related to traditional or subsistence consumption standard, rather than based on the marginal productivity calculus.

[2] However, we do not pretend to provide a realistic representation of the labour market in Cameroon. Neither shall we attempt to test the possible labour market segmentation, since the passage from one sector to another is considered as resulting in part from utility maximisation, accounting for various constraints which limit the mobility of workers, but do not make it impossible.

of capital. See Ranis (1988) for a survey of dualism in development theories. Saint-Paul (1996) extends this representation to industrial countries. On the other side, Rosenzweig (1988) notes that the main deficiency of the studies of dualism is the lack of interpretation of results in terms of behaviour. The same issues apply to the urban labour market of most developing countries, in which the informal sector plays the role of the traditional sector. This calls for the explicit modelling of sector choice and labour supply behaviour in a dualistic environment.

In this context, the urban workers choose their labour activity among varied types of work and earnings perspectives. The recent empirical literature distinguishes three different sectors: formal sector (or primary sector), informal sector (or secondary sector), and non-participation (see Dickens and Lang (1985, 1992); Hill (1988); Chiswick (1988); Magnac (1991); Lemieux, Fortin and Fréchette (1994); Tiefenthaler (1994); Hill (1994); Pradhan (1994)). By definition in the segmented market literature, the primary sector is characterised by high wages, high returns to education and to on-the-job training, and a very concentrated weekly hours schedule. Entry into the primary sector is rationed. Conversely, in the secondary sector, wages are generally low and earnings may be a non-linear function of the worked hours (originated from a technology with decreasing returns to labour). The rates of return to education or experience are low or non-existent in the secondary sector. The weekly hours are much more variable than in the primary sector. Even without specifying a model of market segmentation, this type of dichotomy between sectors seems to be empirically relevant, probably resulting from subjacent market imperfections. The main features of this dichotomy, such as the wage gap, are also consistent with the description of dualist labour markets (see Saint-Paul (1996)).

The representation of economic links between the different sectorial labour markets is far from obvious. Is it relevant to consider a single labour market as in the traditional neoclassical models (Killingsworth (1983)), or should specific market mechanisms be specified for each sector? In the latter case what are the characteristics of each market and of the earnings function associated with each type of job? The labour market segmentation theory (Reich, Gordon and Edwards (1973); Piore (1983); Mazumdar (1983)) rejects the hypothesis of perfect competition in labour markets in favour of a dualistic representation of labour markets. By contrast, we use the framework of neoclassical labour markets, although incorporating entry costs and hours rationing in the primary sector, and accounting for the quasi-autarky of labour in the informal sector. Thus, if segmentation or efficiency wages are present (see Saint-Paul (1996) for a survey of this literature), they are reduced here to the observation of wages in the formal sector that may be different from their marginal productivity, and to rationing and entry costs.

In our model, we describe the rationing of the job supply in the primary sector by entry costs. Another element of the restriction of choices in our model is the hours rationing in the primary sector where the set of wage-hours contracts offered is much more limited than in the secondary sector. To simplify the model we assume that there are entry costs only in the primary sector, which does not mean that every worker in the secondary sector is necessarily rationed since she may have chosen the secondary sector even without the entry costs because her productivity is higher in this sector. One could also imagine that there are different types of entry costs in the primary sector for workers coming from unemployment and from the secondary sector. We do not treat the rationing process in such detail and we consider in fact that the job rationing

corresponding to unemployed workers is exogenous so as to focus on the links between primary and secondary labour markets.[3]

Activity choice has been treated in the literature with models based on log-linear wages or log-linear shadow wages functions and where the endogeneity problems have been simplified by recursivity assumptions or by two-steps estimation methods as in Jacoby (1993). First of all, the Roy model (Roy (1951)) describes the sector choices as derived from the comparison of wage equations in each sector. Todaro (1971) extends this framework by incorporating the probability of obtaining a job in the formal sector, which is different to one due to entry restrictions. Chiswick (1988) believes that even when there exists an informal sector, the human capital earning function may still be a useful analytical tool. Finally, Magnac (1991) and Pradhan (1994) compare the different sector wage equations (or implicit wage equations) with a precise treatment of the selectivity problems.

Most applications measure the informal sector wage as the average hourly income. This is incorrect if the earnings function is a non-linear function of the worked hours, for example derived from a technology showing decreasing returns to labour, and if the informal sector is quasi-autarkic in labour, in the sense that the tasks for the worker's firm are carried out by the worker himself (sometimes by other members of the family). When the earnings function in the labour-autarkic informal sector is non-linear in hours, or when there are entry costs and hours rationing in the formal sector, wages are no longer the only information needed to explain sector choices. Particularly, in the informal sector, the marginal productivity of labour can be very different from the market wage rates of workers of the same qualification, since these workers are generally not accepted as perfect substitutes for family members, or for the worker under consideration, in the informal firm. This is due to the existence of moral hazard and adverse selection phenomena, which make the employment of unknown workers unattractive for the manager of the informal firm. Moreover, sector allocation and labour supply decisions are no longer separable. The correct specification is characterised by the comparison of indirect utility levels, specific to each sector. Given the characteristics in each sector, one simplifying strategy (different from separability assumptions) is to carry out the estimation in two steps to avoid a simultaneous non-linear estimation, even though the selectivity problem may not be perfectly controlled. Thus, Jacoby (1993) estimates a non-recursive agricultural household model where the time allocation of the household is decided in response to shadow wages, which are derived from a preliminary estimation of the agricultural technology. Finally, in a non-separable framework, Newman and Gertler (1994) estimate under limited information, by neglecting the budget constraint and specifying linear implicit log-returns. To our best knowledge, there has never been a full information estimation of a non-separable model with several sectors.

---

[3] In the literature, the entry costs in a specific sector are generally specified as a linear function of worker characteristics. For example, Magnac assumes that the costs of entry depend only on the variables appearing in the wage equation, arguing that these costs are determined by the demand side. In Dickens and Lang (1985), variable entry costs are also explained by the variables of the wage equations, among which is the ethnic group and a dummy variable for never-married people. Using data collected in the Philippines, Tiefenthaler (1994) finds that the entry costs are very different between the sectors, partly because of the substantial transport cost in the primary sector. Using data from Bolivia, Pradhan (1994) finds that education, ethnicity and location in an area with a high employment rate have a significant positive effect on the probability of rationing.

In the informal sector, a labour earnings function which is concave in the worked hours is a natural assumption. Lemieux, Fortin and Fréchette (1994) use this specification for the estimation of a labour supply model in Canada, where activities in both the formal sector and the informal sector are simultaneously possible. Even if the importance of accounting for the technology of the firm in the informal sector is acknowledged, Rosenzweig (1988) argues that the broad diversity of products in the informal sector makes the estimation of this technology difficult. Nonetheless, he insists on the usefulness of estimating the family firm technology to allow for all inputs, and not to estimate only an earnings function. The heterogeneity of technologies in the urban informal sector necessitates very rich information to carry out the estimation. This explains why an earnings function linear in hours is generally estimated without explicitly considering the family production function (for example Hill (1994)). Newman and Gertler (1994) find that the marginal return to farm work and the marginal rate of substitution depend on family interactions. However, in order to limit the size of our problem we do not study these interactions. We estimate an earnings function that is concave in hours for the secondary sector, simultaneously with the labour decisions and activity choice of the female worker. This yields more efficient estimates and a more accurate treatment of the sample selectivity than with the two steps method using the shadow prices.

Another difficult question is the definition of sectors. ILO (1972) discusses the definition of formal and informal sector in Kenya, which raises many difficulties. Rosenzweig (1988) notes that every earnings functions comparison is very sensitive to the number, the definition and the separation level of the sectors. Using data from the Philippines, Tiefenthaler distinguishes four sectors (formal, informal, piece-work, non-participation) and so does Magnac (formal, informal, non-participation, unemployment) with data from Columbia. Most other authors limit themselves to three sectors (primary, secondary, non-participation) or unemployment. Besides, Taubman and Wachter (1986) consider that distinctions across sectors are determined as much by data availability as by the theory. Finally, Mead and Morrisson (1996) show that the term 'informal sector' has very different meanings following the analysts. We consider first a nomenclature of five sectors for descriptive dichotomic participation models, then we use only the three usual sectors for the structural model. Even if we distinguish unemployment and non-participation in the probit models, we do not model endogenously the unemployment situation in the structural model, because of the complexity of its likelihood. We account for unemployed workers in this model provided that this rationing is not related with the determinants of activity choices.[4]

As it is recognised that the female labour supply is more flexible than the male labour supply (see Killingsworth (1983)), particularly concerning the choice of non-participation, we focus on the choices of female active members in the household.[5] This higher variability will make easier the identification of the likelihood of the model. Of course, a similar study for men would be possible.

---

[4] A comparison of public and private modern sectors would also have been interesting to assess the influence of connections rather than competences in obtaining public jobs, since the difference in human capital in these sectors is smaller than between the formal and the informal sector. However, the sub-sample of private sector workers is too small (6.5 per cent of the whole sample) to be treated separately here.

[5] Female members working simultaneously in two different jobs are rare in our sample (8 per cent of workers) and we neglect this element of the activity choice.

From a more general point of view, the question under study is how to account for the heterogeneity of jobs in the female labour supply models. Killingsworth and Heckman (1986) underline that despite the interest of this question, very few studies try to introduce job heterogeneity in female labour decisions. When the heterogeneity of jobs is taken into account, other characteristics of jobs than wage, such as fringe benefits, working conditions and other non-pecuniary characteristics, are also chosen.[6] Hours rationing and entry costs are potentially two major non-pecuniary characteristics of jobs and are concentrated in the primary sector in our structural model. Unfortunately, the observation of labour supply provides only partial information about structural parameters of the labour supply schedule. However, the observation of hours supplied to discrete sectors enables us to partially account for differences in characteristics of jobs.[7]

---

[6] Dickens and Lang (1985) judge that these characteristics are generally inferior in the secondary sector. However, Pradhan (1994) finds that women are more ready to participate in the informal sector than in the formal one for the same wage. In any case the difference in the wage rate between the sectors may reflect the compensation for the difference of labour disutility.

[7] Several questions of economic policy can be treated in the framework of dualistic labour markets in LDCs. The type of model that we propose allows the investigation of these questions, while accounting for the whole set of interactions between decisions and sectors, and for the labour market imperfections. Since the aim of this paper is not to provide a realistic representation of the Cameroon labour market, nor policy recommendations, we do not simulate or study any specific policy. We mention, however, several possible policy investigations, as a general motivation for the study of this type of model. The assessment of the effects of the fixation of public wages is a natural usage of such model.

Firstly, the fixation at a high level of wage rates in the public sector is a potential source of imperfection of the labour market. Indeed, for obvious political reasons, the wage rate in the public sector rarely corresponds to the marginal productivity of labour. Besides, by definition, the public activities are not directly related to the system of private markets and therefore their allocations and prices do not correspond to the general equilibrium allocation of the economy.

Secondly, many African governments have lowered the wage rates of public servants in response to budgetary crises and structural adjustment plans. Manpower regulation through dismissals, fixation of worked hours and modification of entry barriers in the public sector has also been implemented. These policies are likely to entail migrations of workers from the public sector to the informal sector, unemployment or non-participation categories.

Finally, the LDCs' governments and the development agencies are increasingly concerned by the development of the informal sector both as an alternative source of the economic growth and as an employment sector preferable to unemployment for workers rejected from the primary sector. Consequently, the development projects aiming at improving the productivity and the labour absorption capacity of the informal sector are in expansion.

In practice, the size and the direction of the migrations of manpower implied by all these policies, as well as the importance of misallocations of labour, are almost unknown and the use of models, not accounting for the specific characteristics and imperfections of these markets, may provide misleading conclusions.

3. Data description

3.1 Descriptive statistics

Cameroon in 1993 is a lower-middle-income economy (770 US$ of GNP per capita) at the intersection of Western and Central Africa, with a large access to the gulf of Guinea. With 12.5 million inhabitants and 475 thousand square kilometres, it is a medium-sized African country. At the time of the survey, the economy is in recession with an average annual 7.3 per cent decrease of per capita GNP from 1985, and every production sector is affected (agriculture, wood, oil and industry). Oil production is in decline and the 1992–93 coffee and cacao crops have fallen because of public funding problems met by peasants (-29 per cent for robusta coffee, -32 per cent for arabica coffee, -12 per cent for cacao). This bad conjuncture entails a permanent pressure on the employment in rural and urban areas. The government is not able to comply with its commitments, particularly with the IMF, which suspends its funding and asks for dismissals in the public sector. Moreover, the World Bank strongly supports the privatisation and restructuration of public firms, which is also likely to entail manpower migrations out of the formal sector in Yaoundé, mostly composed of public administrations or firms.

The data used in this study are taken from a survey aimed at studying the employment conditions of the population of Yaoundé, capital of Cameroon (DIAL-DSCN (1993)). The survey has been realised in collaboration by DIAL (Développement des Investigations sur Ajustement à Long Terme, research centre in Paris) and the DSCN at the Ministry of the Republic of Cameroon. 11172 people in 1961 households have been surveyed from a sampling plan based on spatial teledetection. The collection took place in January–February 1993. The devaluation of the Franc CFA (FCFA) in January 1994 does not affect the information collected that corresponds to preceding periods. The survey provides information concerning the labour market status and the activity of every individual aged over 10. Among 7 865 individuals of more than 9 years old, 2 947 are occupied active members and 832 unemployed. We consider the population of married women or single female heads of household aged between 18 and 55, leaving us with 1225 individuals.

Individuals in the primary sector (27 per cent of the population) are wage earners, employed in the public sector, either in an administration (64 per cent) or in a state firm (12 per cent), or employed in a private firm (24 per cent). The secondary sector workers (18 per cent of the population) are those who claim to be working either as a small-firm employer (less than 10 per cent of the secondary sector workers) or as an artisan.

DIAL-DSCN (1993b) shows that the activity conditions in the informal sector are very precarious, with less than 20 per cent of the informal production unit with premises specifically devoted to the activity. More than three third of these production units are limited to one worker. Only 4 per cent employ more than three workers. Only 10 per cent of workers in the informal sector are wage earners, and among them only one-third receive a fixed wage, the other ones often being paid with respect to the task carried out. Moreover, most of salaried workers in the informal sector have parental links with the owner of the firm, while the remaining wage earners have generally as well strong social and ethnic links with this owner. Finally, more than two-thirds of these jobs have been obtained from connections. All these elements strongly suggest that the labour market for wage earners in the informal sector is extremely reduced and characterised by overwhelming imperfections. An acceptable approximation seems to consider that in most cases,

the informal activities are labour-autarkic activities for the informal workers. In the rare cases when another worker, who is a wage earner, is present in the informal unit, his/her labour is likely to be imperfectly substitutable with the labour of the considered female worker. Indeed, moral hazard and adverse selection phenomena are pervasive in this context, where the owner of the firm or the person directly responsible for the activity (here mostly the considered female worker) has a much more intense interest in the productivity of the informal firm, than a supplementary wage earner, difficult to supervise with accuracy.

Unemployed workers (41 per cent of the population) are individuals reporting to have been searching for a job at least once during the four weeks before the interview, individuals claiming that no jobs are available as a reason for not searching and individuals reporting to suffer from under-employment because of bad economic conditions. 14 per cent of female workers decide not to participate in the labour market.[8]

We summarise this situation by assuming in the model that salaried workers belong to the primary sector, whereas the secondary sector is composed of self-employed workers. Because of the importance of public administration in Yaoundé, most (76 per cent) of salaried workers are civil servants. The majority of these individuals work around 40 hours a week. The wage rates are determined by a global pay scale which is negotiated with the government. Primary sector jobs in the private sector often share similar characteristics. In the secondary sector, the worker determines freely her labour supply and the technology of the individual firm may exhibit diminishing marginal return to labour.

Table 1 gives details of the socio-economic characteristics of the sample by sector of activity. The values obtained for these statistics compare well with what is known of the labour market in Yaoundé (see DIAL-DSCN (1993a, 1993b)) and guaranties an excellent quality of the data. Column 'secondary sector (a)' refers to observations for which the precise income is known. Column 'secondary sector (b)' refers to the observations in the secondary sector for which only the range of income is given. The two sets of observations may correspond to very different jobs, secondary sector (b) being associated with the less formal jobs where the earnings are more volatile, and secondary sector (a) with more stable earnings, but in the informal sector. This distinction is also interesting because workers in the secondary sector (a) may have earnings related to a rigid rate (possibly intermediary between their marginal and their average productivity because of the variability of this productivity with time), while workers in the secondary sector (b) may have earnings following a more flexible schedule that may be closest to their actual marginal productivity, and even may better reflect a non-linear (in hours) earning function originated from diminishing returns to labour.[9] The figure provided for the income per month corresponds to the mean between the bounds of the range chosen.

---

[8] Note that the characteristics of non-participant and unemployed workers are very similar. The unavoidable and common inaccuracy in the definition of these classes of workers may lead to confusion between unemployed workers who have a strong need of work, and workers who would be ready to work only a few hours or in very favourable conditions, perhaps unrealistic in the Cameroon context. Several tries with different definitions of the categories of unemployed and non-participant do not seem to change substantially the results, perhaps because of the homogeneity of these two categories.

[9] Another possibility would be to consider different industrial sectors, but such details are outside the scope of this paper.

Unemployed and non-participants are on average younger than employed women. The average education level in the primary sector (11.9 years) is almost twice the one in the secondary sector (6.48 years) and is higher than in the non-participant and the unemployed sector (8.54 years). 63 per cent of the population have attended at least a secondary education. The husband's age is much higher for non-working women (33 years), while the husband's education level is lower for women in the secondary sector (5.85 years). The worker's experience is higher in the primary sector than in the secondary one, and much higher than for non-participant and unemployed people, which suggests some permanency of these latter statuses for many female workers. The number of female members in the household from 13 to 18 years old is lower for non-working women. This suggests the possibility of a substitutability in the domestic work of these members and that of female worker. The number of children under four years old is higher among the non-working women, which is consistent with the fact that women often refuse to work in order to devote more time to child care. The number of dependants[10] can be quite large – the maximum is 24 – but generally one or two dependants are present in the households. The number of dependants is higher when the female worker is in the primary sector, and lower when she does not work. This could be the consequence of the general social level of the household, since traditionally the rich households have to take in charge some of their family members initially extra to the household.[11] The married women are less active which is consistent with the fact that they can rely on their husband's income.

Several elements militate in favour of the existence of entry costs or market imperfections, when investigating the simple descriptive statistics for different sectors. Having been born in a city other than Yaoundé seems to be a disadvantage for working in the secondary sector. A preceding job in an administration appears to favour employment in the primary sector. The sector of the father, or of the husband's father, may influence the present sector of the female worker. Women coming from a peasant family are more likely to be non-working or to be employed in the secondary sector rather than in the primary sector. On the contrary, when her father had worked in the services, the female worker has more chances to be employed in the primary sector. Moreover, the ethnic group of the woman, and even more the ethnic group of her husband, play a role in the determination of the activity sector. Bassa or Bakoko people are more likely to work firstly in the primary sector, secondly in the secondary one. By contrast, Bamiléké people are more concentrated in the secondary sector.

The level of household assets is higher in the first type of secondary sector (1188 FCFA) than in the second type (796 FCFA), which is associated with less productive activities. Moreover, this assets level is closer to the one observed in the primary sector (879 FCFA). Non-working women correspond to a much lower level of assets (204 FCFA).

---

[10] Dependants are defined as individuals surveyed living with the household who are neither the head of household nor his/her spouse nor the children of the head of household.

[11] From a more general point of view, the link household composition with activity sector and labour supply may be resulting in part from the endogeneity of household composition. In particular, members of the family that originate from outside the household domestic unit may be supported and incorporated in the richer households. Unfortunately, our data do not enable us to correct for these parasite phenomena, and we neglect them in our comments.

The number of worked weeks per year is higher in the primary sector (45.23 weeks) than in the secondary sector (41.37 weeks). On the other hand, the number of worked weekly hours is higher in the secondary sector (42.25 hours) than in the primary one (38.94 hours). The variance of the worked hours in the two sectors corresponds to what is recognised in the literature as the relatively stronger rigidity of hours in the primary sector. In the primary sector, 81 per cent of the population works from 30 to 48 hours a week, with a strong mode at about 40 hours where we observe 65 per cent of the observations. This justifies considering the hours worked per week as rationed (generally at about 40) in the primary sector. In the secondary sector, we find 38 per cent of the population working from 30 to 48 hours a week and the mode is at 60 hours where we see 10 per cent of the observations.[12]

DIAL-DSCN (1993a, 1993b) provides empirical elements leading us to consider that despite the market imperfections, the workers have considerable autonomy in their choice of activity sector. Indeed, 25 per cent of unemployed people who lost their job chose to quit voluntarily. Less than 20 per cent of informal independent workers chose the informal sector because they had not found a salaried job in the modern sector. This suggests that entry and exit into and out of the activity sector is often voluntary, even if there exist barriers to entry. Note also that two-thirds of the informal workers have obtained their job through family or friends, and only one-quarter by directly applying for a job.

3.2 Wage and non-labour income equations

As mentioned earlier, labour income can be reported in two different ways in the questionnaire: either the individual reports the exact amount earned per month (in 1000 Francs CFA) or indicates among several predetermined ranges, where her income is situated. We estimate now wage equations for the two sectors, with the purpose of progressing in the preliminary statistical analysis of the data, and not to conduct a structural analysis of wage determination. This justifies that we use simpler specifications or methods, than what could be possible. In the case of the primary sector, we estimate a log-wage equation on the continuous observations using an Instrumental Variables method to account for the possible endogeneity of experience. Mills ratios, included to investigate the possible selectivity bias for participation in the considered sector, are not significant and have been excluded.[13] Statistics of the predictions from the 2SLS estimates of this wage equation are shown in Table 1. The same type of log-wage equation has been estimated for the secondary sector with the average hourly income instead of the wage. Henderson (1983), find that the wage equations are not significantly different for wage and self-employed workers in Iran. Other authors such as Chiswick found different results for the two sectors. In fact, there is a trade-off between the need to compare similar activities and the risk of introducing selection bias by restricting the set of occupations. Our estimates contribute to this debate and confirm the structural difference between rewards in the formal and the informal sector. Table 2 shows the

---

[12] This picture is not altered substantially if we account for the variation in terms of weeks per year since 86 per cent of the workers in the primary sector and 70 per cent of the workers in the secondary sector work no less than 40 weeks per year (with a mode at 48 weeks, with 42 per cent of the observations in the primary sector and a mode at 52 weeks in the secondary sector with the same proportion of the observations).

[13] Moreover, in the public sector, since the log-wage equation attempts to recover the pay scale of civil servants that depends mainly upon qualification, age and experience, it should not be very sensitive to variations of the sample of workers if the estimation is accurate enough. In that case a regression in levels describing the salary scale would have been possible, but less direct to compare with the log-wage equation in the informal sector.

estimation results of the log-wage equations.[14] The results of the Hausman tests show that experience and education can be considered as exogenous in every log-wage equations, which suggests favouring OLS estimates.[15]

Ignoring the selectivity problem, the coefficient of the variable 'years of education' in the log-wage equation can be interpreted as the rate of return to investment in schooling (see Willis (1986)). The OLS estimates show that it is higher in the primary sector (0.071) than in the secondary sector (0.048) where it is not significantly different from zero.[16] The coefficient of the dummy for primary education is never significant in any sector while this education level could be believed to be well adapted to the standard types of task that are carried out in the informal sector. The coefficient of the dummy for higher education is significatively positive for the two sectors and higher in the secondary sector and permits controlling for very qualified jobs concerning only a few workers.[17] Also, as suggested sometimes in the segmented labour market theory, the positive effect of the working experience is only positive in the primary sector.

Once experience is included, the age plays a notable positive role only in the secondary sector. Note also that the wage is much better explained in the primary sector than in the secondary one. The average predicted unearned income and the average predicted weekly wage rate of the female worker are higher in the primary sector than in the secondary one. It is interesting to see that whatever the activity sector in which the female is working, on average her predicted primary sector wage rate is always higher than her predicted secondary sector wage rate which is consistent with dualism. However, primary workers gain more (300 Franc CFA more on average) by being in this sector than do the other categories (110 to 150 FCFA on average). This is consistent with the theory of comparative advantages and suggests that a model for which workers balance their costs and advantages in different sectors, associated with a selection by firms of workers according to their productivity, may fit the data well.

Observed unearned income is the sum of the woman's unearned income and the total income of the husband and of the dependants (if any are present).[18] The result of the estimation of the

---

[14] The number of observations differs here a lot from that in the descriptive statistics because we used only the most precise indicators for wages and unearned incomes in the estimates, systematically eliminating the 'zero' observations for incomes and imprecise data for wages. Indeed, we believe that for this type of data measurement error is a major problem and if not treated would lead to meaningless estimates. That is another reason for not treating explicitly the endogenous selectivity of the sample in these equations since some attrition bias results already from the calculus of indicators.

[15] Furthermore, the limited accuracy of 2SLS estimates indicates clearly that our instrument set is insufficient for a satisfactory prediction of experience and education.

[16] The return of education in the informal sector appears to be much too high in the 2SLS estimates (27 per cent). This is in fact an illusion due to the loss of accuracy caused by the instrumentation. Indeed, the coefficients of education are not significant in the 2SLS estimations.

[17] The no-education and the secondary education categories have been excluded in the equation.

[18] Some women are still observed with no unearned income. We assume that this is due to the non-observation rather than to the non-existence of unearned income. Since the existence of a significant selection effect using a conventional Heckman 2-step method is rejected, all the predictions are obtained using the ordinary least squares estimates on complete observations only.

unearned income equation is presented in Table 3. This is an auxiliary equation used to generate the variable 'predicted unearned income'. The number of children, their age, the age and the education level of the husband, and the size of the husband's father's firm are shown to influence the level of unearned income.

3.3 The participation model

In a first step towards the estimation of a full model of activity choices, we describe the participation process, for each type of activity considered. This preliminary examination of the data based on probit models enables us to include a large set of explanatory variables. The population under study is composed of married women who are distributed in five sub-populations: 1) formal sector; 2) informal sector with precise income (type a); 3) informal sector with inaccurate income (type b); 4) non-participating women; 5) unemployed women. The participation choices in 1, 2, 3 have been studied separately with respect to alternatives 4 and 5. In this analysis, we ignore the selectivity problem involved in the simultaneous consideration of all alternatives. That will be treated in the next section by the specification of a structural model.[19]

We present in Table 4 the results of probit estimations, in which the list of exogenous variables includes the explanatory variables of preferences, fixed costs and wage rate. The heterogeneity of the behaviour of the five sub-populations under study appears in these results.[20]

Everything else being constant, married women participate less in the different working activities (against non-participation). However, when the age and education of the husband are included, the coefficient of the dummy for married women becomes most of the time non-significant and is only positively significant for the choice of working in the formal sector against non-participation. Workers whose husbands are old have a lower probability of participating in the formal sector against non-participation, which may be related to a traditional mentality of the head according to which women should stay at home and not join the labour market. The husband's education is not significant.[21] The presence of children lowers labour participation in the formal sector against unemployment, especially for children between 4 and 10 years old. This may be related to the caring for children at home by mothers. The 'inverse U curve' describing the effects of age is almost always significant, which corresponds to a usual feature of life cycle models.

The presence of substitutabilities in domestic work, between women and girls or young women inside the household, is suggested by the positive effect of variables 'number of female members under 13 years old' and 'number of female members from 13 through 18 years old', on participation of women in the primary sector against non-participation and unemployment.

---

[19] The same instruments as for the unearned income equation have been used to derive predicted values for the dummy variables specific to the preceding job of the female worker (among these variables is the dummy for administration). Indeed, these variables may be related to the same type of choice as those under study. We use the predicted variables in the estimation.

[20] Note that this was not always obvious in Table 1, especially when the two types of secondary sector, or the two types of non-working status, are compared.

[21] The workers whose husbands are well educated participate less in the second type of informal sector, when predicted wages are added to the equation.

The number of education years has positive effects on participation in the formal sector against non-participation or unemployment, although it is associated with a lower probability of participating in the second type of informal sector against unemployment. Dummies for primary or higher education are less influential. Experience is positively associated with participation in the formal or informal sectors against unemployment, and with participation in the informal sector against non-participation. The effects of education and experience may be a mixture of their influences on preferences and on wage rates or labour productivity, and are therefore difficult to interpret. The predicted unearned income of the female worker is related to a lower probability of participating in the second type of informal sector against non-participation, which may be caused by an income effect on labour supply.

The effects of the other variables suggest that it would be possible to explain the fixed cost by socio-economic characteristics related to the social connections of workers.[22] The probability of entry into the primary sector against non-participation is higher (conditionally to other independent variables) for women originated from another city than Yaoundé, and lower against unemployment for women whose husband's ethnic group is Ewondo or whose father worked in services. The probability of entry into the secondary sector against unemployment is lower for women whose husband's ethnic group is Bassa, Bakoko or Bamiléké. The probability of participation in the second type of informal sector against non-participation or unemployment is lower for women originated from cities other than Yaoundé, and whose husbands' ethnic group is Bamiléké, although it is higher for women of the Bamiléké ethnic group.

4. The Model

We now present a model where we explicitly treat the selectivity caused by the choice between the primary sector, the secondary sector and the non-working status. We simultaneously model the activity choice, the entry cost into the primary sector and the labour supply, accounting for hours rationing in the primary sector. The earnings function in the secondary sector is concave with respect to working hours. The error terms in preferences, fixed costs and earnings functions of the informal sector can be correlated because of unobserved heterogeneity and measurement errors jointly influencing these three fundamental elements of the model. However, these improvements of the modelling necessitate a drastic diminution of the number of exogenous variables, and the number of sectors considered. We also consider that the wage rate in the primary sector is exogenous because of institutional constraints and to simplify the estimation.[23] In spite of all these simplifications our approach corrects for deficiencies of elementary models that do not fully consider the fundamental problems of non-linearity, endogeneity and selectivity involved in the estimation of a structural model.

[22] Reasonable interpretation of the effects of the socio-economic variables in terms of entry costs has been confirmed by D. Njinkeu, an economist from the University of Yaoundé.

[23] We assume that the decisions are mainly taken by the female worker alone. A collective decision framework for the household would have been more general, as in Chiappori (1988), but it would have entailed identification difficulties, making impossible the estimation of the model.

4.1 Definition of regimes

The preferences of the individual over consumption C and worked hours (h) are represented either with a direct utility function (U(c, h)), or with the corresponding indirect utility function (Ψ(w, y)) where w is the wage rate and y is the appropriate measure of exogenous income. We assume that the indirect utility function is strictly monotonically increasing in y and strictly monotonically increasing and strictly quasi-convex in w. We denote h(w, y) for the Marshallian labour supply function obtained from Ψ(w, y) through the Roy identity. The labour earnings function in the secondary sector, R(h) is such that $\frac{dR(h)}{dh} > 0$ and $\frac{d^2 R(h)}{dh^2} \leq 0$, allowing for decreasing return to labour. Only the labour of the individual enters in the subjacent production function, which is consistent with the very imperfect labour market associated with informal activities. We consider that the fixed cost (fc) of entry into the primary sector is given, while we disregard the fixed cost of entry in the secondary sector. Any job in the primary sector is supposed to be described by a fixed number of hours $\bar{h}$ accounting for the hours rationing in this sector, and a given hourly wage rate w. As said above, simultaneous activities are excluded. If the labour earnings function R(h) were linear in h, if there were no rationing in the primary sector, and if there were no fixed costs, the comparisons in utility would be reduced to a comparison in wages. In that case, the Roy model as described by Magnac (1991) would be the appropriate way to model the sector choices. In our case, because of the existence of fixed costs, of hours rationing and of the non-linearity of R(h), the sector choice is derived from the comparison of the sector-dependent indirect utilities (as in Hausman (1985)). Thus, the activity and labour decisions of workers have a similar structure to those of non-linear discrete continuous choice models (Hanemann (1984)).

The observed position is summarised in (s^{obs}, h^{obs}, R^{obs}), s^{obs} being the observed sector of activity (1, 2 or 3), h^{obs} being the observed labour supply, and R^{obs} being the observed non-wage activity income. (s^{obs}, h^{obs}, R^{obs}) is related to the latent model in the following way where the unobserved variables are replaced by a zero.

$$
(1) \qquad (s^{obs}, h^{obs}, R^{obs}) = \begin{cases} (1, \bar{h}, 0) & \text{if (primary sector):} \\ & \text{a) } U(\bar{c}, \bar{h}) \geq \Psi(w^v(h^*), \mu^v(h^*) + y_0) \\ & \text{b) } U(\bar{c}, \bar{h}) \geq U(y_0, 0), \\ (2, h^*, R^*) & \text{if (secondary sector):} \\ & \text{a) } \Psi(w^v(h^*), \mu^v(h^*) + y_0) \geq U(\bar{c}, \bar{h}), \\ & \text{b) } \Psi(w^v(h^*), \mu^v(h^*) + y_0) \geq U(y_0, 0), \\ (3, 0, 0) & \text{if (non-participant):} \\ & \text{a) } U(y_0, 0) > \Psi(w^v(h^*), \mu^v(h^*) + y_0), \end{cases}
$$

b) $U(y_0, 0) > U(\bar{c}, \bar{h})$,

where $w^v(h) = \frac{dR(h)}{dh}$ and $\mu^v(h) = R(h) - h\, w^v(h)$ are respectively the virtual wage and the net virtual activity income associated with the implicit linear budget constraint in the secondary sector. $\bar{c} = w\,\bar{h} + y_0 - fc$ is the consumption expenditure in the primary sector.

$h^* = h(w^v(h^*), \mu^v(h^*) + y_0)$ is the optimal labour supply in the secondary sector obtained from the Roy identity, $R^* = R(h^*)$ is the non-wage activity income in the secondary sector.

We consider that the case of unemployed workers is exterior to this structural model. Indeed, the comparison of their indirect utilities implies that, without job rationing, they would be included among workers in the primary or secondary sectors. Then, the corresponding observations should satisfy:

$$U(\bar{c}, \bar{h}) \geq U(y_0, 0) \text{ or } \Psi(w^v(h^*), \mu^v(h^*) + y_0) \geq U(y_0, 0).$$

However, to explain their unemployment status we would have to include in the model the entry costs into the primary and secondary sectors, specific to these workers. Indeed, one may expect that it is generally more difficult to find a job when coming from unemployment status, which may be a signal of low productivity. Even when it is not the case, including this category of workers in the model would complicate it too much, hampering an already complex estimation process. We choose rather to consider the job-rationing process leading to unemployment as exogenous, which is equivalent to excluding unemployed workers from the sample used in the estimation.

4.2 The specification of the utility function

Several specifications of the utility function have been used in the labour supply literature. The most common is certainly the linear form in income and wage (see Hausman (1981) and Stern (1986)). Yet this form presents some deficiencies which makes it difficult to use in a structural model: low-income individuals may exhibit a negative marginal rate of substitution between consumption and hours of work, and for a sufficiently large wage rate, consumption becomes an inferior good. As a simple alternative to Hausman's form, we use of the following class of utility functions that has recently been suggested by Blundell (1990).

$$(2) \qquad \Psi(w, y_0) = \left[y_0 + \frac{\alpha}{\gamma+1} w + \frac{\beta}{\gamma+1} w \cdot \left(\ln(w) - \frac{1}{\gamma+1} \right) \right] \cdot w^{\gamma}$$

where imposing $\beta > 0$, $0 > \gamma > -1$ is sufficient to ensure the convexity of preferences, a strictly positive marginal rate of substitution and the normality of consumption good. Using Roy identity we obtain the following labour supply function:

$$(3) \qquad h = \alpha + \beta \ln(w) + \gamma \frac{y_0}{w}.$$

Heterogeneity of preferences can be introduced through the intercept term α. When the labour supply is strictly positive, the marginal rate of substitution $\omega(c, h)$ and the virtual income $\mu(c, h)$,

associated with the separating hyperplane of the optimisation programme, can be defined as the roots of the system of equations:

$$(4) \qquad h = \alpha + \beta \ln(\omega(c, h)) + \gamma \frac{\mu(c, h)}{\omega(c, h)},$$

$$(5) \qquad c = \omega(c, h)\, h + \mu(c, h).$$

The direct utility function can then be obtained by

$$(6) \qquad U(c, h) = \Psi(\omega(c, h), \mu(c, h)),$$

which yields the following marginal utilities:

$$(7) \qquad \frac{\partial U(c, h)}{\partial c} = \omega(c, h)^{\gamma}, \quad \text{and} \quad \frac{\partial U(c, h)}{\partial h} = -\omega(c, h)^{\gamma+1}$$

which have the appropriate signs if and only if $\omega(c, h)$ is strictly positive.

4.3 Specification of the earnings function in the informal sector

The choice of a functional form is driven by crucial arguments. Firstly, we must be able to obtain decreasing returns to labour over a sufficiently wide range of hours in order to obtain a unique equilibrium in the secondary sector. Secondly, to allow for non-participation, the marginal return to labour at zero hours must be bounded.[24] The specification of the earnings function that we use here has been introduced by Rosen (1976). It is such that:

$$(8) \qquad R(h) = h \exp(\xi + \delta h),$$

where ξ and δ are parameters. Constant returns are obtained when $\delta = 0$, and provided we are willing to bound the time endowment of any individual by T, positive marginal returns to labour and decreasing returns to labour are obtained if $-1/T < \delta < 0$. The logarithm of the average income relates to the log-average-wage equation for which we have:

$$(9) \qquad \ln\left(\frac{R(h)}{h}\right) = \xi + \delta h.$$

The implicit wage rate in the secondary sector is:

$$(10) \quad R'(h) = (1 + \delta h).R(h)/h .$$

[24] For instance while the Cobb-Douglas specification meets the first requirement easily, it never meets the second one since marginal return to labour tends to infinity when hours tend to zero.

Adding exogenous explanatory variables and an error term to ξ is a simple way to introduce deterministic and stochastic variation into this equation.[25] Since the marginal return to labour at zero hours is bounded, this specification allows for non-participants. Finally the virtual net activity income is obtained by subtracting the implicit value of labour input.

$$(11) \qquad \mu^{v}(h) = R(h) - h\,R'(h) = -\delta\, h\, R(h)$$

5. Estimation of the Structural Model

5.1 Some econometric difficulties

Several econometric difficulties have to be coped with. They guide the functional and stochastic specification of the model:

- the selectivity bias has been well documented in the literature for the standard labour supply model (see Heckman and MaCurdy (1986)). Chiswick (1988) proposes to incorporate in an analysis of earnings those labour force participants who are self-employed, and thus avoids the bias in the earnings functions coefficients that arises when the sample is restricted to wage earners. We go a step further by explicitly including in the global model the endogenous selectivity rules and a non-linear earnings function for the secondary sector. We treat the selectivity bias by estimating an integrated model, in which all sector choices are simultaneously estimated;
- identification conditions have to be imposed in switching response models. For instance, in the case of probit model, the coefficients of explanatory variables are only identified up to a multiplicative constant. Such identification problems are much less likely to occur when a non-linear earnings function is used, since some collinearity problems disappear. However, because of the highly non-linear character of our model, it is not possible to globally identify a large set of parameters with accuracy. This explains why the number of exogenous variables is quite limited in this model. Moreover, the correct derivation of the likelihood from the error terms distribution implies to impose the logical coherency property to the model (see Gouriéroux, Laffont, Monfort (1980)). This is satisfied by insisting that parameter values must lead to a unique optimum for each value of the error terms and vice versa. This condition is obtained firstly by assuming strict concavity of the utility function and the earnings function, and by the fact that the event defined by the equality of utility levels in two different sectors has a null probability.

5.2 The likelihood function and the exogenous variables

The full model is given by system (1) with the functional form (2) for the utility function and the equation (8) as the earnings function in the secondary sector. The maximum likelihood method used in the estimations is fully discussed in Appendix 1. The stochastic specification involves three terms related to the following structural parameters of the model: the heterogeneity of preferences, α, the scale of the revenue function, ξ, and the fixed cost, fc. α, ξ and fc are supposed to be linear in the error terms and some characteristics of the individual, and are assumed to be distributed according to a trivariate normal distribution.

[25] The heterogeneity could also have been introduced through the parameter δ but this would complicate an already difficult estimation.

Because of the high non-linearity of the model, it is not possible to identify with accuracy all the parameters related to the explanatory variables when they are included at the same time in preferences, technology and entry cost. Therefore, we decided to introduce each explanatory variable in only one of these structures (preferences, technology, entry cost) and where it seems to be the more relevant.

Earnings (or wage rate) are explained by the usual variables coming from human capital theory (age and age squared, education level, experience), but also, in the case of the informal sector, by the characteristics of the individual firm (here the value of the capital stock of the firm, approximated by the assets of the household). We have special interests in the role of children that provide incentives to the mother to take care of them, and on the other hand in the role of other female members who could substitute themselves to the female workers in the domestic activities. The complex sociological phenomena occurring in the households for these considerations lead us to assume that the preferences of the female worker depend on the number of children and the number of female members, both categories split up into two age classes. The preferences depend also on the age of the woman and a dummy variable for married women, since age is a major determinant of preferences and the presence of a husband is expected to have a strong influence on female worker decisions.

Because of identification difficulties, it has not been possible to include the explanatory variables of entry costs in the estimation of the complete model, as they have been included in the dichotomic participation models in section 3.3.[26]

5.3 Estimation results

Table 5 provides one typical estimation of the model.[27] The parameters β and $(1/\gamma)$, which are related to the economic variables, i.e. predicted wage and predicted unearned income, are well determined.

---

[26] The measurement unit of the fixed cost is Franc CFA. Its variance may also be considered as a measure of heterogeneity of the budget constraint specific to the primary sector. Finally, a dual interpretation of the fixed cost suggests that it partially accounts for the unobserved heterogeneity in the rationing of hours (around 40 per week) in the primary sector.

Initially, the entry costs were explained by the ethnic group of the female worker or of her husband. Indeed, the social obstacles or privileges involved in entry into activity sectors are likely to depend on the characteristics of the parents or of the family, as much as on the personal characteristics of the female worker. The other explanatory variables of the entry costs are the geographical origin (Yaoundé, other cities than Yaoundé, countryside), the social origin (activity sector of the father or of the husband's father), and the preceding firm (administration, large private firm, small private firm...). However, since the choice of the preceding firm is likely to be endogenous, we have instrumented these variables in the estimation. Unfortunately, if the model with explanatory variables of fixed costs converges, the variance–covariance matrix of its parameters is not estimable because of poor identification.

[27] Because of all the non-linearities involved in the model, the computation is time-consuming. The precision required for the integration and for the root finding makes it variable. One week of calculation time with GAUSS on a SUN-SPARC station is a typical duration for one estimation. The starting values are obtained after the estimation of the model on a 30 per cent random sample of the final dataset. Several 30 per cent sub-samples have been tried and the values obtained were similar to one another. The region of convergence had also to be reached by grid method and likelihood concentration methods so as to compensate the inaccuracy of the gradient method near the optimum. Note also that the algorithm for a similar model with hours freely chosen in the primary sector does not converge. See the appendix 1 for more details about the optimisation of the likelihood.

The number of young children under 4 years of age has a negative effect on the propensity to work, which may reflect their need for caring from the female worker. By contrast, the number of children between 4 and 10 of age, the number of female members and the fact that the individual is married,[28] provide incentives to work. The presence of other female members may ease the substitutability of the female worker's labour in domestic activities. Age has no significant effect on work decisions. This may be due to compensation effects in the life cycle. Including both age and age squared has not been possible since these variables are too much correlated to be accurately identified in this model.

Fixed costs are identified by the structure that we impose on preference and technology. Because of the highly non-linear character of the model, we have not been able to obtain estimates for the variance–covariance matrix of the parameters when explanatory variables have been included in the fixed costs, even when the convergence is reached.[29] The constant fixed-cost coefficient is not significant. However, its estimate is 4.4 per cent of the non-labour exogenous income for primary sector workers, although respectively 11.2 per cent and 10.9 per cent for workers in secondary sectors of types a and b, and respectively 6.6 per cent and 8.3 per cent for non-participants and unemployed workers. This is consistent with the fact that the entry cost contributes to the selection of workers since it corresponds to a relatively lower financial burden for actually observed primary workers. It is remarkable that workers for whom the fixed cost is a relatively larger financial constraint, compared to their financial capacity without wage income, and that may not access to the primary sector, seem rather to be rejected in the secondary sector than in inactivity. The variance σ_{fc} is large and indicates that the amount of fixed costs any individual face can be very variable.

The estimates of coefficients of the earnings function are the most interesting. The household assets, which are a proxy for the capital stock of the firm, are statistically significant, which reveals the inadequacy of the standard wage rate approach in this context. The experience and the education of the female worker positively influence the productivity, which was not the case in the wage rate equation of the primary sector.[30] The parameter δ is significantly negative, which implies that the earnings function is strictly concave and confirms the relevance of our simultaneous non-linear modelling approach. Then, the estimation of the structural model, consistent with the theoretical framework, provides evidence for the non-linearity of the technology of the informal sector. However, the estimated curvature of the earnings function is very moderate. This may be caused by two features of the studied context. First, as our

[28] It has been suggested that the latter effect may be related to the bad economic conjuncture, since married women may be losing advantages linked to the husband's job, while non-married workers are not subject to this shock.

[29] Note that to the best of our knowledge, nobody has yet succeeded in estimating fixed-cost functions in such non-linear models. This may explain why Newman and Gertler (1994) choose to eliminate the budget constraint in their model and rely on comparisons of implicit log-wages based on linear specifications. One reason of this difficulty is that only very rich datasets can identify with accuracy the parameters both in the linear budget constraint and in a non-linear production function included in this budget constraint. Indeed, the corresponding Fisher information matrix shows columns that are likely to be very correlated, leading to its quasi-singularity and the impossibility of estimating the variance–covariance matrix of parameters.

[30] However, as we said above, this may also come from the impossibility of identifying the effects of these variables in preferences.

preliminary statistical analysis shows, the informal sector is an aggregate of heterogeneous activities, with some workers receiving very variable remunerations and others paid at almost a fixed wage rate. The contribution of these last workers in the likelihood pushes the estimates towards a null curvature. Second, as in organisational dualism theories, the earnings of the female worker may be shared with other members of the family, and average implicit wage or a fraction of the implicit wage may sometimes be closer to the characteristics of the decision process, than the actual implicit wage. Nonetheless, the linear implicit log-wage equations, which are used in the literature and in section 3, are likely to be a source of misspecification.[31]

Finally, the correlation coefficients between preferences, fixed cost and technology are all significant, which supports the joint modelling of these structures as a response to the existence of unobserved heterogeneity influencing at the same time preferences, technology and entry costs.[32]

Table 6 provides several useful statistics based on the separation structure of the optimisation programme, for the mean households, in seven sub-populations of interest. They are estimates of elasticities of labour supply and demand; elasticities of earnings; wages and incomes. They characterise potential labour responses, incomes and wages, for mean worker in each sector, under the assumption that there is no change of regime, and in reference to a simplified set of constraints (the implicit separating hyperplane associated with the optimisation programme). It must be remembered that those statistics are related to latent variables, and do not generally correspond to observed elasticities.[33] They provide, however, useful analytical information about the properties of the estimated system. We distinguish the statistics associated with latent consumer and latent producer programs. We calculate these statistics, using as the value of h, the average observed labour supply and the mean characteristics in every sub-population. For the non-participants and the unemployed, we account for the fact that their reservation rate should be higher than the implicit proposed wage rates (in absence of fixed costs and hours rationing), by using the maximum of the predicted wage rates of the primary and the secondary sectors wage rates (in practice the primary sector wage rate).

a. Latent consumer program:

Definitions

$e_1 = \beta/h - (\gamma\, y_0)/(w\, h)$ *is the elasticity of the latent labour supply with respect to the observed wage rate.*

$e_{2,i} = X_i b_i / h$ *is the elasticity of the latent labour supply with respect to the characteristic* X_i *of the preferences.*

---

[31] Linear reduced-form models are still valid competitors of the structural approach, since they permit the introduction of a large number of explanatory variables even if the interpretation of their results is difficult. The impossibility of including all explanatory variables in our framework is the reason why we do not attempt to compare our results with a linear reduced approach. We prefer to focus on the structural properties of our model rather than on the robustness of the effects.

[32] However, it must be recalled that this approach, because of identification difficulties, may hide effects of some explanatory variables of the entry cost.

[33] Observed elasticity should account for the non-separability of the system and changes of regime.

$e_2 = \alpha / h$ denotes the sum of these elasticities for all characteristics of the preferences (including the constant term).

$e_3 = (\gamma\ y_0\ )/(w\ h)$ is the elasticity of the latent labour supply with respect to the exogenous income.

b. Latent producer program:

The earnings function in the formal sector is bilinear in actual wage and hours, and does not present any difficulty. We consider the earnings function in the informal sector as a latent earnings function, under the assumption that there is no change of regime.

Definitions

$e_4 = 1+\delta h$ is the latent earnings (from the informal production) elasticity with respect to the worked hours.

$e_{5,j} = Z_j c_j$ is the latent earnings (from the informal production) elasticity with respect to the characteristic Z_j of the technology.

$e_5 = \xi$ denotes the sum of these latent earnings elasticities, for all the characteristics of the technology (including the constant term).

Other useful statistics can be derived from the profit maximisation programme of the latent producer, where the restricted profit function is $\Pi(h) = R(h) - w\ h$. Since the earnings function is strictly concave, the optimal labour demand[34] is given by the first-order condition:

$$R'(h) = w, \text{ i.e. } (1 + \delta\ h) \exp(\xi + \delta\ h) = w.$$

The differentiation of this equation yields the following elasticities of the labour demand for the latent producer sub-program.

Definitions

$e_{6a} = w / [\ h\ \delta\ (2 + \delta\ h) \exp(\xi + \delta\ h)\ ]$ is the elasticity of labour demand with respect to the observed wage rate w.

$e_{6b} = (\ 1 + \delta\ h) / [\ h\ \delta\ (2 + \delta\ h)]$ is the elasticity of labour demand with respect to the shadow wage rate in the informal sector.

$e_{7,k} = - Z_k c_k\ (1 + \delta\ h) / [\ h\ \delta\ (2 + \delta\ h)\ ]$ is the elasticity of labour demand with respect to the characteristic Z_k of the technology.

$e7 = -(1 + \delta\ h) / [\ h\ \delta\ (2 + \delta\ h)\ ]$ is the sum of these elasticities of labour demand, for all the characteristics of the technology.

c. Implicit wages and incomes

Definitions

w_1 is the predicted wage rate in the formal sector. w_2 is the predicted wage rate in the informal sector. w_3 is value of the marginal productivity in the informal sector, which can be considered as the shadow wage rate in the labour-autarkic informal sector.

[34] Equal to the labour supply in the global model at the equilibrium.

R_0 is the exogenous income. R_1 is the exogenous income minus the entry cost in the primary sector. R_3 is the earnings level given by the earnings function in the informal sector, for the observed worked hours.

d. Results:[35]

The wage elasticities (e_1) of the labour supply in the primary sector (0.371), or in the secondary sector (0.493 and 0.485) compare well with traditional figures (Killingsworth (1983), Blundell (1992)). They correspond to virtual situations where only the implicit budget constraint (therefore without fixed costs or hours rationing) is considered. Since $\gamma < 0$, e_1 is higher for high exogenous incomes and lower for high labour supply levels, other variables being maintained constant. e_1 is higher in the secondary sector because the wage rate is much lower, which is far from fully compensated by a lower exogenous income and larger worked hours.

The elasticity of the labour supply with respect to the exogenous income, (e_3), is lower in absolute value in the primary sector (-0.186) than in the secondary sector (-0.308, -0.299). This is due, as before, to a higher wage rate in the primary sector. Naturally, for the non-participants and the unemployed, the wage elasticity tends to plus infinity and the exogenous income elasticity tends to minus infinity, when the worked hours tends to zero.

The elasticity of earnings with respect to worked hours, e_4, is almost similar for the mean household in any sector, and close to 1. This result comes from the very moderate concavity of the earnings function. The elasticity of the latent labour demand, with respect to the implicit wage rate, e_{6b} (calculated from the estimated marginal productivity at mean hours), is always strongly negative, and close for the two activity sectors, since the mean levels of worked hours are similar in both sectors. The elasticity of the latent labour demand with respect to the observed wage, is still stronger in absolute value, and lower in absolute value for the workers observed in the informal sector (-9.7 and -10.7), than for the workers observed in the formal sector (-19.9). This suggests, that the adjustment of labour allocation in the informal sector may be substantially more flexible with the introduction of perfect markets in each sector, assuming that these markets clear at the respective observed mean values of wages.

The implicit marginal wage rate in the informal sector, w_3, is always lower than the mean predicted wage rates in both sectors. It is much higher (more than twice) for workers observed in the formal sector, than for workers observed in the informal sector, due to the higher education, experience and assets of the former.

Finally, the potential activity income coming from the informal activities of the female workers, (R_2), is on average relatively small compared to the exogenous income (R_0), even when subtracting the entry cost into the formal sector (fc). Indeed, beyond the non-activity incomes of the households, the labour incomes of the other members of the family, and particularly of the husband, constitute most of the family income. This suggests that a further analysis of female activity decisions could benefit from the consideration of activity decisions of other members.

---

[35] To shorten the exposition, we show in table 6 only aggregate elasticities e_2, e_5 and e_7, which we do not comment, instead of elasticities with respect to elementary characteristics of preferences or technology, as would be possible. The latter can be easily calculated from the provided descriptive statistics.

6. Conclusion

In the dualist context of developing countries, labour supply and activity choices are likely to be distorted by the existence of labour market imperfections. Modelling and testing these phenomena raises complex econometric issues, specially when one allows for the presence of a non-linear earnings function in the informal sector, allied to quasi-autarky of labour in this sector, and entry costs and hours rationing in the formal sector. This explains why no models dealing jointly with these characteristics have been estimated until now. This paper is a first step toward such a task.

We consider a cross-section sample of female workers from Yaoundé (Cameroon) allocated to three different sectors. The first sector is that of modern salaried activities where the wages are high, but for which fixed costs and labour time rationing exist. The second sector is that of informal self-employed activities, for which there exist diminishing returns to labour. The third sector is non-participation. The analysis of the data shows features specific to dualistic labour markets.

We estimate a joint model of the activity choice and the labour supply, where the error terms in the preferences, the fixed cost and the return to labour in informal technology may be correlated. In spite of technical difficulties, this approach is proven tractable and our results provide evidence for the concavity of the earnings function in the informal sector with respect to the hours. We identify the explanatory variables of the preferences and the earnings function in this simultaneous structural estimation, and derive elasticities for different sectors.

Since there is no perfectly substitutable worker in the labour market for the female worker in the informal sector (notably because of supervision costs), the concavity of the earnings function implies that the consumption–leisure, the choice of activity sector and the production–labour decisions are not separable. This element, joined to hours rationing and entry costs in the formal sector, means that the second welfare theorem does not readily apply and it is not possible easily to decentralise a welfare objective through the use of a general wage policy only. Policies specific to each sector based on their observable characteristics and not only on prices or wages may then be an alternative to consider.

Bibliography

BENJAMIN, D. (1992), 'Household Composition, Labour Markets, and Labour Demand: Testing for Separation in Agricultural Households Model', *Econometrica*, 60, 287–322.

BLUNDELL, R. (1990), 'Evaluating Structural Microeconometric Models of Labour Supply', Working Paper, University College London.

BLUNDELL, R. (1992), 'Labour Supply and Taxation: a Survey', *Fiscal Studies*, 13, 15–40.

CHIAPPORI, P.A. (1988), 'Rational Household Labour Supply', *Econometrica*, 56: 63–89.

CHISWICK, C.U. (1988), 'Analysis of Earnings from Household Enterprises: Methodology and Application to Thailand', *Review of Economics and Statistics*, February.

DIAL–DSCN (1993a), 'Conditions d'Activité de la Population de Yaoundé: Premiers Résultats', *Technical Report*.

DIAL–DSCN (1993b), 'Le Secteur Informel à Yaoundé, Enquête 1-2-3. Phase 2', *Technical Report*, November.

DICKENS, W.T. and K. LANG (1985), 'A Test of Dual Labour Market Theory', *American Economic Review*, 75, 792–805.

DICKENS, W.T. and K. LANG (1988), 'The Reemergence of Segmented Labour Market Theory', *AEA Papers and Proceedings*, May.

DICKENS, W.T. and K. LANG (1992), 'Labour Market Segmentation Theory: reconsidering the Evidence', *National Bureau of Economic Research Working Paper*, 4087.

GHATAK, S. and K. INGERSENT (1984), *Agriculture and Economic Development*, Johns Hopkins University Press.

GOURIEROUX, C., J.-J. LAFFONT and A. MONFORT (1980), 'Coherency Conditions in Simultaneous Linear Equation Models with Endogenous Switching Regimes', *Econometrica*, 48, 675–695.

GRONAU, R. (1986), 'Home Production – A Survey', chapter 5 in O. Ashenfelter and R. Layard, *Handbook of Labour Economics*, Elsevier Science Publishers.

HANEMANN, W.M. (1984), 'Discrete/Continuous Models of Consumer Demand', *Econometrica*, 52, 541–61.

HAUSMAN, J.A. (1980), 'The Effect of Wages, Taxes and Fixed Costs on Women's Labour Force Participation', *Journal of Public Economics*, 14, 161–194.

HAUSMAN, J.A. (1985), 'The Econometrics of Nonlinear Budget Sets', *Econometrica*, 53, 1255–1882.

HECKMAN, J. and B. HONORÉ (1990), 'The Empirical Content of the Roy Model', *Econometrica*, 58, 1121–1149.

HECKMAN, J. and T.E. MACURDY (1986), 'Labour Econometrics', chapter 32 in Z. Griliches and M.D. Intriligator, *Handbook of Econometrics*, Elsevier Science Publishers.

HECKMAN, J. and G. SEDLACEK (1985), 'Heterogeneity, Aggregation and Market Wage Functions: An Empirical Model of Self-Selection in the Labour Market', *Journal of Political Economy*, 93, 1077–1125.

HENDERSON, J.W., (1983), 'Earnings Functions for the Self-Employed', *Journal of Development Economics*, 13, 97–102.

HILL, M.A. (1988), 'Female Labour Force Participation in Developing and Developed Countries – Consideration of the Informal Sector', *Review of Economics and Statistics*, February.

HILL, M.A. (1994), 'Female Labour Supply in Japan. Implications of the Informal Sector for Labour Force Participation and Hours of Work', *Journal of Human Resources*, XXIV, 1.

I.L.O. (1972), 'Employment Incomes and Equality: A Strategy for Increasing Productive Employment in Kenya', *International Labour Office*, Geneva.

JACOBY, H.J. (1993), 'Shadow Wages and Peasant Family Labour Supply: An Econometric Application to the Peruvian Sierra', *The Review of Economics Studies*, 60, 903–921.

KILLINGSWORTH, M.R. (1983), *Labour Supply*, Cambridge University Press.

KILLINGSWORTH, M.R. and J. HECKMAN, (1986), 'Female Labour Supply: A Survey', chapter 2 in O. Ashenfelter and R. Layard, *Handbook of Labour Economics*, Elsevier.

LEMIEUX, T., B. FORTIN and P. FRÉCHETTE (1994), 'The Effect of Taxes on Labour Supply in the Underground Economy', *American Economic Review*, 84, 231–254.
LEWIS, W.A. (1954), 'Economic Development with Unlimited Supplies of Labour', *Manchester School of Economics and Social Studies*, 131–191.
MAGNAC, T. (1991), 'Segmented or Competitive Labour Markets', *Econometrica*, 59, 165–187.
MAZUMDAR, D. (1983), 'Segmented Labour Markets in LDCs', *American Economic Review*, AEA Papers and Proceedings, May, vol. 73, no. 2.
MEAD, D.C. and C. MORRISSON (1996), 'The Informal Sector Elephant', *World Development*, vol. 24, no. 10.
MONFORT, A. (1980), *Cours de Probabilités*, Economica, Paris.
MYINT, H., (1985), 'Organizational Dualism and Economic Development', *Asian Development Review*, vol. 3, no. 1, pp. 25–42.
NEWMAN, J.L. and P.J. GERTLER (1994), 'Family Productivity, Labor Supply, and Welfare in a Low Income Country', *Journal of Human Resources*, XXIX, no. 4.
PIORE, M.J. (1983), 'Labour Market Segmentation: To What Paradigm Does it Belong?', *American Economic Review*, AEA Papers and Proceedings, May, vol. 73, no. 2.
PRADHAN, M. (1994), 'Labour supply in urban areas of Bolivia and the role of the informal sector', Ph.D. thesis.
RANIS, G. and J. FEI (1961), 'A Theory of Economic Development', *American Economic Review*, 56, 533–558.
RANIS, G. (1988), 'Analytics of Development: Dualism', chapter 4 in H. Chenery and T.N. Srinivasan, *Handbook of Development Economics. Volume I*, Elsevier.
REICH, M., D.M. GORDON and R.C. EDWARDS (1973), 'A Theory of Labour Market Segmentation', *American Economic Review*, May, vol. 63, no. 2.
RICARDO, D. (1815), *Works and Correspondence of David Ricardo*, vol. 1–4, Cambridge University Press.
ROSEN, H.S. (1976), 'Taxes in a Labour Supply Model with Joint Wages–Hours Determination', *Econometrica*, 44, 485–507.
ROSENZWEIG, M.R. (1988), 'Labour Markets in Low-Income Countries', chapter 15 in H. Chenery and T.N. Srinivasan, *Handbook of Development Economics*, Elsevier Science Publishers.
ROY, A. (1951), 'Some Thoughts on the Distribution of Earnings', *Oxford Economic Papers*, 3, 135–146.
SAINT-PAUL, G. (1996), *Dual Labour Markets*, MIT Press.
SEN, A.K. (1966), 'Peasants and Dualism with or without Surplus Labour', *Journal of Political Economy*, October.
SINGH, I., L. SQUIRE and J. STRAUSS (1986), *Agricultural Household Models*, The World Bank.
STERN, N. (1986), 'On the Specification of Labour Supply Functions', in R. Blundell and I. Walker, eds, *Unemployment, Search and Labour Supply*, Cambridge University Press, Cambridge.
TAUBMAN, P. and M.L. WAECHTER (1986), 'Segmented Labour Markets', chapter 21 in O. Ashenfelter and R. Layard, *Handbook of Labour Economics*, Elsevier Science Publishers.
THOMAS, J.J. (1992), *Informal Economic Activity*, LSE Handbooks in Economics, Harvester Wheatsheaf.
TIEFENTHALER, J. (1994), 'A Multisector Model of Female Labour Free Participation: Empirical Evidence from Cebu Island, Philippines', *Economic Development and Cultural Change*, July.
TODARO, M.P. (1971), 'Income Expectations, Rural–Urban Migrations and Employment in Africa', *International Labour Review*, vol. 104, no. 5, November.
WILLIS, R. (1986), 'Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions', chapter 10 in O. Ashenfelter and R. Layard, *Handbook of Labour Economics*, Elsevier Science Publishers.

Table 1: Descriptive statistics by sector of activity, for female workers

	Primary sector	Secondary sector (a)	Secondary sector (b)	Secondary sector (complete)	Non-participant	Unemployed	Inactivity sector (complete)	Whole population
Age	36.37 (6.8)	35.57 (8.3)	34.75 (7.9)	35.16 (8.07)	32.01 (9.2)	28.77 (6.7)	29.58 (7.54)	32.43 (8.1)
Years of experience	11.45 (6.65)	8.50 (8.11)	8.32 (7.89)	8.40 (7.98)	1.73 (4.56)	1.33 (4.01)	1.43 (4.16)	5.41 (7.29)
Years of education	11.92 (4.0)	6.61 (3.6)	6.36 (4.0)	6.48 (3.78)	7.83 (4.6)	8.78 (3.3)	8.54 (3.70)	9.07 (4.2)
Husband's age	27.7 (20.4)	28.03 (20.6)	28.68 (21.6)	28.36 (21.04)	37.02 (13.4)	31.64 (14.4)	32.99 (14.38)	30.71 (17.66)
Husband's years of education	9.60 (7.9)	5.79 (5.6)	5.90 (5.58)	5.85 (5.59)	9.52 (5.70)	9.12 (5.4)	9.22 (5.45)	8.70 (6.4)
Primary education	0.11 (0.32)	0.59 (0.49)	0.65 (0.48)	0.62 (0.49)	0.46 (0.50)	0.34 (0.48)	0.37 (0.48)	0.35 (0.48)
Higher education	0.28 (0.45)	0.009 (0.09)	0.044 (0.21)	0.027 (0.16)	0.10 (0.30)	0.072 (0.26)	0.079 (0.27)	0.12 (0.33)
Number of female members between 13 and 18 years	0.67 (0.86)	0.45 (0.72)	0.54 (0.71)	0.50 (0.73)	0.35 (0.66)	0.37 (0.70)	0.37 (0.69)	0.47 (0.75)
Number of female members between 0 and 12 years	1.15 (1.28)	1.27 (1.22)	1.14 (1.28)	1.20 (1.25)	1.03 (1.24)	1.18 (1.23)	1.14 (1.24)	1.15 (1.25)
Number of children under 4 years	0.47 (0.73)	0.51 (0.71)	0.54 (0.71)	0.53 (0.71)	0.74 (0.84)	0.71 (0.75)	0.72 (0.77)	0.62 (0.75)
Number of children between 4 and 10 years	1.16 (1.19)	1.26 (1.31)	1.23 (1.15)	1.24 (1.23)	1.02 (1.18)	1.24 (1.32)	1.18 (1.29)	1.19 (1.25)
Number of dependent members	1.85 (2.37)	1.41 (2.93)	1.54 (2.04)	2.76 (2.16)	1.04 (1.68)	1.36 (1.80)	1.28 (1.77)	1.47 (2.11)
Dummy for married woman	0.67 (0.47)	0.68 (0.47)	0.68 (0.47)	0.68 (0.47)	0.93 (0.25)	0.86 (0.34)	0.88 (0.32)	0.79 (0.41)
Origin from other cities than Yaoundé	0.62 (0.49)	0.56 (0.50)	0.44 (0.50)	0.50 (0.50)	0.61 (0.49)	0.63 (0.48)	0.62 (0.49)	0.60 (0.49)
Preceding job in administration	0.25 (0.43)	0.10 (0.30)	0.11 (0.32)	0.11 (0.31)	0.15 (0.36)	0.17 (0.37)	0.16 (0.37)	0.18 (0.38)
Sector of husband's father = agriculture	0.20 (0.40)	0.28 (0.45)	0.29 (0.46)	0.28 (0.45)	0.30 (0.46)	0.35 (0.48)	0.33 (0.47)	0.29 (0.45)
Sector of father = agriculture	0.24 (0.43)	0.39 (0.49)	0.39 (0.49)	0.39 (0.49)	0.32 (0.47)	0.32 (0.47)	0.32 (0.47)	0.31 (0.46)
Sector of father = services	0.36 (0.48)	0.22 (0.41)	0.18 (0.38)	0.20 (0.40)	0.25 (0.43)	0.29 (0.45)	0.28 (0.45)	0.29 (0.45)
Husband's ethny = Bassa and Bakoko	0.08 (0.27)	0.018 (0.13)	0.044 (0.21)	0.031 (0.17)	0.060 (0.24)	0.086 (0.28)	0.079 (0.27)	0.071 (0.26)
Husband's ethny = Bamiléké	0.15 (0.36)	0.20 (0.40)	0.21 (0.41)	0.20 (0.40)	0.30 (0.46)	0.33 (0.47)	0.32 (0.47)	0.25 (0.44)
Husband's ethny = Ewondo	0.10 (0.30)	0.18 (0.39)	0.19 (0.40)	0.19 (0.39)	0.14 (0.34)	0.18 (0.38)	0.17 (0.37)	0.15 (0.36)
Ethny = Bamiléké	0.20 (0.40)	0.28 (0.45)	0.34 (0.48)	0.31 (0.46)	0.30 (0.46)	0.36 (0.48)	0.34 (0.48)	0.30 (0.46)
Weeks per year	45.23 (8.30)	41.03 (13.63)	41.71 (14.46)	41.37 (14.02)	4.57 (13.96)	2.71 (9.69)	3.17 (10.94)	21.52 (22.98)
Hours per week	38.94 (8.47)	42.64 (23.21)	41.88 (20.39)	42.25 (21.78)	0 (0)	9.73 (10.48)	9.73 (10.48)	39.69 (15.88)
Assets (*)	878.50 (2720)	1188.76 (5209)	795.64 (2080)	989.58 (3943)	301.40 (2079)	171.96 (1006)	204.42 (1356.39)	530.22 (2443)
Unearned income (predicted) (*)	61.02 (64.9)	24.07 (24.6)	24.60 (29.2)	24.34 (27.01)	40.74 (43.37)	32.49 (26.21)	34.56 (31.57)	39.81 (44.64)
Secondary sector wage (predicted) (*)	0.49 (0.47)	0.19 (0.13)	0.20 (0.19)	0.19 (0.16)	0.23 (0.23)	0.19 (0.17)	0.20 (0.19)	0.28 (0.31)
Primary sector wage (predicted) (*)	0.79 (0.54)	0.34 (0.23)	0.35 (0.26)	0.35 (0.25)	0.34 (0.26)	0.34 (0.20)	0.34 (0.22)	0.46 (0.40)
Observed wage (*)	67.02 (112.07)	33.40 (68.44)	–	–	–	–	–	45.77 (94.72)
Number of observations	330	111	114	225	168	502	670	1225

* : in thousands of FCFA.

Table 2: Log-wage equations

	Primary sector		Secondary sector	
	OLS	2SLS	OLS	2SLS
Constant	-1.914 (5.3)	-1.468 (0.05)	-3.355 (5.9)	-5.731 (2.3)
Age	-0.0020 (0.3)	-0.057 (1.15)	0.0228 (1.7)	0.0229 (0.8)
Years of education	0.0707 (3.60)	0.140 (0.6)	0.0483 (1.3)	0.268 (1.1)
Experience	0.0398 (3.78)	0.113 (1.7)	0.0119 (1.0)	0.049 (0.7)
Primary education	-0.206 (1.2)	-0.295 (0.2)	0.0123 (0.05)	1.127 (0.8)
Higher education	0.653 (4.4)	0.101 (0.06)	1.362 (2.6)	-0.161 (0.08)
R^2	0.536	–	0.205	–
Number of observations	188	188	108	108

Wages are measured in thousands of FCFA per week. Hausman Test statistic for the endogeneity of experience and education: 6.791 for primary sector, 1.088 for secondary sector, 15.378 for the whole set of participants. The instruments are the number of children, the number of children aged from 0 to 3, from 4 to 10, from 11 to 15, the education level of the parents, the occupation of the father, the age and education of the husband.

Table 3: Non-labour income equations

	Heckman two-steps estimates		Ordinary Least Squares	
Variables	**Estimate**	**(t-value)**	**Estimate**	**(t-value)**
Constant	2.582	(2.0)	1.067	(3.3)
Number of children	0.071	(1.6)	0.099	(3.2)
Number of children from 0 to 3 years old	-0.194	(2.5)	-0.146	(2.7)
Number of children from 4 to 10 years old	-0.0928	(1.7)	-0.086	(1.9)
Number of dependants	0.008	(0.4)	0.027	(1.7)
Husband's age	0.0256	(2.3)	0.016	(2.4)
Husband's years of education	0.0857	(3.7)	0.096	(5.6)
Husband's primary education	-0.115	(0.7)	-0.145	(1.1)
Husband's higher education	0.146	(0.77)	0.233	(1.7)
Husband's father employed	-0.406	(1.6)	-0.259	(1.4)
Husband's father self-employed	0.128	(1.2)	0.157	(1.9)
Husband's father employed in public sector	0.252	(1.1)	0.052	(0.4)
Husband's father employed in small private firm	-0.731	(2.0)	-0.673	(2.4)
Husband's father employed in large private firm	0.347	(2.0)	0.322	(2.3)
Age	-0.0007	(0.08)	0.003	(0.5)
Years of education	0.035	(1.5)	0.021	(1.2)
Primary education	0.032	(0.2)	-0.046	(-0.4)
Higher education	0.208	(1.1)	0.227	(1.5)
Father employed	0.0209	(0.11)	0.132	(1.0)
Father self-employed	-0.0424	(0.4)	-0.047	(0.6)
Father employed in public sector	-0.137	(0.9)	-0.165	(1.3)
Father employed in small private firm	-0.181	(0.5)	0.035	(0.1)
Father employed in large private firm	0.0470	(0.3)	0.0041	(0.03)
Dummy for married woman	-1.534	(1.4)	-0.281	(0.8)
Mill's ratio	-1.492	(1.2)	–	–
R^2	0.421		0.420	
Number of observations	956		956	

Table 4: **Estimates of dichotomic choices (Probit Maximum Likelihood)**

Choices / Explanatory variables	1/4	1/5	2/4	2/5	3/4	3/5	Whole sample
Constant	-8.3167 (-4.40)	-8.5467 (-6.40)	-4.0872 (-2.38)	-3.7725 (-2.89)	-4.2819 (-2.42)	-4.5316 (-3.45)	-6.1652 (-7.17)
Married woman	1.8990 (2.13)	0.4296 (0.77)	-0.5870 (-0.68)	-0.0558 (-0.09)	-1.3175 (-1.42)	-0.8350 (-1.30)	0.1599 (0.40)
Age	0.4561 (4.40)	0.3867 (5.09)	0.2756 (2.86)	0.1773 (2.38)	0.3141 (3.19)	0.2508 (3.29)	0.3181 (6.47)
Age squared	-0.0068 (-4.62)	-0.0053 (-4.86)	-0.0041 (-3.11)	-0.0023 (-2.22)	-0.0047 (-3.49)	-0.0035 (-3.24)	-0.0044 (-6.32)
Number of children between 0 and 3 years	-0.1033 (-0.77)	-0.0444 (-0.46)	-0.0930 (-0.69)	0.0045 (0.04)	-0.0620 (-0.45)	0.0396 (0.37)	-0.0397 (-0.59)
Number of children between 4 and 10 years	-0.0802 (-0.81)	-0.1479 (-2.32)	0.0946 (0.96)	-0.0532 (-0.74)	0.0387 (0.37)	-0.1108 (-1.46)	-0.0667 (1.43)
Number of female members under 13 years	0.0277 (0.31)	0.1202 (1.85)	-0.0459 (-0.45)	0.0417 (0.55)	-0.1144 (-1.14)	-0.0305 (-0.40)	0.0410 (0.87)
Number of female members between 13 and 18 years	0.2583 (1.81)	0.1633 (1.88)	0.0594 (0.40)	-0.0172 (-0.17)	0.0424 (0.27)	0.0324 (0.32)	0.0617 (0.96)
Husband's age	-0.0675 (-3.67)	-0.0144 (-1.21)	-0.0089 (-0.53)	0.0054 (0.44)	0.0057 (0.34)	0.0156 (1.24)	-0.0082 (-1.02)
Husband's education	-0.0063 (-0.24)	0.0068 (0.34)	0.0039 (0.14)	-0.0293 (-1.26)	0.0386 (1.22)	0.0266 (1.05)	0.0005 (0.04)
Education	0.0976 (2.46)	0.1035 (3.39)	-0.0094 (-0.21)	-0.0473 (-1.30)	-0.0372 (-0.78)	-0.0802 (-2.21)	0.0156 (0.73)
Dummy for primary education	-0.4671 (-1.56)	0.0807 (0.38)	0.0863 (0.30)	0.1504 (0.71)	0.1495 (0.51)	0.2519 (1.18)	0.1524 (1.06)
Dummy for higher education	0.0749 (0.23)	0.0432 (0.18)	-1.1542 (-1.57)	-0.4335 (-0.96)	0.0732 (0.15)	0.3944 (1.15)	0.2573 (1.38)
Unearned income	-0.0001 (-0.14)	0.0016 (0.75)	-0.0045 (-1.50)	0.0035 (1.19)	-0.0103 (-2.98)	-0.0006 (-0.20)	-0.0002 (-0.33)
Origin from another city than Yaoundé	0.3238 (1.69)	-0.0159 (-0.12)	0.0597 (0.30)	0.0064 (0.05)	-0.6142 (-2.97)	-0.3688 (-2.46)	-0.0720 (-0.79)
Bamiléké woman	0.3783 (0.63)	-0.0804 (-0.27)	0.9815 (1.07)	0.3588 (1.10)	1.4366 (2.13)	0.9686 (3.20)	0.3600 (1.62)
Husband's ethny = Bassa or Bakoko	0.1799 (0.46)	-0.2676 (-1.12)	-0.1850 (-0.32)	-0.9036 (-2.14)	0.3673 (0.83)	-0.3195 (-1.04)	-0.2482 (-1.32)
Husband's ethny = Bamiléké	-0.7652 (-1.23)	-0.2665 (-0.81)	-1.0466 (-1.09)	-0.6005 (-1.65)	-1.1511 (-1.62)	-0.9925 (-2.88)	-0.5144 (-2.10)
Husband's ethny = Ewondo	-0.2204 (-0.77)	-0.3850 (-1.90)	0.0935 (0.33)	-0.2239 (-1.07)	-0.2755 (-0.85)	-0.3664 (-1.61)	-0.2928 (-2.08)
Sector of father = agriculture	-0.0315 (-0.13)	-0.0231 (-0.14)	-0.1828 (-0.74)	-0.0638 (-0.36)	-0.0295 (-0.12)	0.0249 (0.14)	0.0242 (0.21)
Sector of father = services	-0.0217 (-0.09)	-0.3022 (-1.90)	0.3837 (1.48)	0.1193 (0.65)	0.0836 (0.30)	-0.2853 (-1.49)	-0.0650 (-0.56)
Sector of husband's father = agriculture	0.2710 (1.25)	-0.1310 (-0.88)	0.2651 (1.27)	-0.0239 (-0.15)	0.2788 (1.29)	-0.0625 (-0.38)	0.0071 (0.07)
Preceding job in administration	-0.4960 (-0.97)	0.5190 (1.43)	-1.0721 (-1.42)	-0.0861 (-0.17)	-0.4356 (-0.62)	0.6671 (1.35)	0.2214 (0.78)
Experience	0.2162 (0.02)	0.1479 (12.4)	0.1154 (6.10)	0.0887 (7.22)	0.1176 (6.08)	0.093 (7.72)	0.1365 (15.6)

t-values are given in parenthesis

Table 5: Maximum likelihood estimates

Mean log-likelihood = -3.64. Number of observations = 1206. t-values in parenthesis.

α	**Constant**	52.71
		(6.5)
	Number of children between 0 and 3 years	-5.52
		(2.9)
	Number of children between 4 and 10 years	4.25
		(3.1)
	Married woman	7.71
		(2.2)
	Number of female members over 13 years	4.65
		(2.4)
	Number of female members between 0 and 13 years	4.15
		(2.9)
	Age	-1.00
		(0.05)
	β	6.25
		(5.4)
	$1/\gamma$	-12.24
		(11.8)
ξ	**Constant**	-3.05
		(17.6)
	Assets	0.030
		(2.1)
	Education	1.10
		(11.4)
	Experience	0.044
		(5.9)
	δ	-0.0022
		(2.9)
	Cf	2.69
		(0.4)
Variance–Covariance Matrix	σ_{α}	32.8
		(17.2)
	σ_{ξ}	0.89
		(10.5)
	σ_{fc}	109.9
		(12.4)
	$_{\alpha\xi}$	-0.44
		(5.4)
	$_{\alpha,fc}$	0.23
		(2.9)
	$_{\xi,fc}$	-0.66
		(2.6)

A few aberrant observations (19) have been discarded from the sample so as to reach the convergence point. To ease the optimisation algorithm behaviour, some variables have been rescaled: hours are per week, wages in thousands of FCFA, assets in millions of FCFA, education and age in dizains of years. This must be taken into account when interpreting the above parameters.

Table 6: Elasticities, wages and incomes for different sectors

Statistics	formal	informal of type a	informal of type b	informal (complete)	non-participants	unemployed	inactivity (complete)	whole sample
e_1 labour supply w.r.t. wage rate	0.37	0.493	0.485	0.497	+∞	+∞	+∞	0.46
e_2 labour supply w.r.t. char. preferences	1.9	1.911	1.904	1.907	+∞	+∞	+∞	1.98
e_3 labour supply w.r.t. non-labour income	-0.2	-0.31	-0.3	-0.31	-∞	-∞	-∞	-0.27
e_4 earnings w.r.t. worked hours	0.93	0.926	0.926	0.926	1	1	0.999	0.93
e_5 earnings w.r.t. char. technology	-1.2	-1.91	-1.96	-1.94	-2.153	-2.021	-2.04	-1.79
e_{6a} labour demand w.r.t. observed wage rate	-20	-9.72	-10.7	-9.98	-∞	-∞	-∞	-17.4
e_{6b} labour demand w.r.t. implicit wage rate	-6.5	-6.5	-6.51	-6.5	-∞	-∞	-∞	-6.75
e_7 labour demand w.r.t. char technology	-7.8	-12.4	-12.8	-12.6	-∞	-∞	-∞	-12.1
w_1 wage rate in primary sector	0.79	0.34	0.35	0.35	0.34	0.34	0.34	0.46
w_2 wage rate in secondary sector	0.49	0.19	0.2	0.19	0.23	0.19	0.2	0.28
w_3 implicit wage rate in secondary sector	0.26	0.127	0.121	0.124	0.116	0.133	0.13	0.143
R_0 non-labour income inactive	61	24.07	24.6	24.34	40.74	32.49	34.56	39.81
R_1 non-labour income minus entry cost	58.3	21.38	21.91	21.65	38.05	29.8	31.87	37.12
R_2 potential activity income in secondary sector	9.39	4.61	4.39	4.49	0	0	0	5

Annexe 1

The next sections describe the contributions to the likelihood of each regime.

$$(12)\quad \begin{pmatrix} \alpha \\ \xi \\ fc \end{pmatrix} \sim N\left(\begin{pmatrix} \bar{\alpha} \\ \bar{\xi} \\ \bar{fc} \end{pmatrix}, \sum \equiv \begin{pmatrix} \sigma_\alpha^2 & \rho_{\alpha\xi}\,\sigma_\alpha\,\sigma_\xi & \rho_{\alpha fc}\,\sigma_\alpha\,\sigma_{fc} \\ \rho_{\alpha\xi}\,\sigma_\alpha\,\sigma_\xi & \sigma_\xi^2 & \rho_{\xi fc}\,\sigma_\xi\,\sigma_{fc} \\ \rho_{\alpha fc}\,\sigma_\alpha\,\sigma_{fc} & \rho_{\xi fc}\,\sigma_\xi\,\sigma_{fc} & \sigma_{fc}^2 \end{pmatrix} \right)$$

Primary sector workers

This regime is such that the following inequalities are fulfilled:

$$(13)\quad U(\bar{c}\,(fc), \bar{h};\, \alpha) \geq \Psi(w^v\,(h^*\,(\alpha, \xi);\, \xi),\, \mu^v(h^*\,(\alpha, \xi);\, \xi) + y_0;\, \alpha),$$

with $h^*\,(\alpha, \xi)$ such that:

$$h^*\,(\alpha, \xi) = h(w^v(h^*(\alpha, \xi);\, \xi),\, \mu^v\,(h^*(\alpha, \xi);\, \xi) + y_0\,;\, \alpha),$$

$$U(\bar{c}\,(fc), \bar{h};\, \alpha) \geq U(y_0,\, 0;\, \alpha),$$

where we have extended our notation to take into account the dependence on the error terms in α, ξ and fc and where w^v and μ^v are short notations for $\omega(.,.)$ and $\mu(.,.)$ defined in (4) and (5).[36] The probability to observe a primary sector worker is:

$$(14)\quad L_1 = \int_{-\infty}^{+\infty} \int_{-\infty}^{fc_1^M(\alpha)} {}_2\left[\frac{e_\alpha}{\sigma_\alpha}, \frac{e_{fc}}{\sigma_{fc}}, \rho_{\alpha,\,fc}\right] {}_2\left[\frac{e_{\xi 1}^M - [e_\alpha,\, e_{fc}]\,\Sigma_{11}^{-1}\,\Sigma_{12}}{\sqrt{\sigma_\xi^2 - \Sigma_{21}\,\Sigma_{11}^{-1}\,\Sigma_{12}}}\right] \frac{1}{\sigma_\alpha\,\sigma_{fc}}\, dfc\, d\alpha$$

where

$$e_\alpha = \alpha - \bar{\alpha},\quad e_{fc} = fc - \bar{fc},\quad e_{\xi 1}^M = \xi_1^M\,(\alpha,\, fc) - \bar{\xi}.$$

Φ is the standard normal cumulative distribution function and ${}_2[.,.,\rho]$ is the bivariate standard normal probability density function. The matrices Σ_{11}, Σ_{12}, Σ_{21} are relevant sub-matrices of Σ following results on the conditional distribution of multivariate normal random variables (see Monfort (1983)).

Self-employed workers

If for the secondary sector workers we assume that the earnings function depends non-linearly on the labour supplied, which is endogenous to our model, the selectivity issue cannot be ignored. The description of the probability for this regime follows a very similar route. The following equations characterise a self-employed worker:

---

[36] Given a value of α, the second inequality gives us an upper bound for the fixed cost, i.e. $fc_1^M(\alpha)$. This quantity does not depend on the scale of production. Provided the Slutsky condition is satisfied this bound is unique. Therefore, for any value of the fixed cost less than the upper bound $fc_1^M(\alpha)$ and any value of α, the upper bound for the production scale can be defined, i.e. $\xi_1^M(\alpha, fc)$. Note that in the optimisation programme, the upper and lower bounds in the integrals (in this regime and the other ones) are numerically computed using a combination of bisection and Newton-Raphson method.

(15) $\Psi(w^{v}(h^{*}(\alpha, \xi); \xi), \mu^{v}(h^{*}(\alpha, \xi); \xi + y_0; \alpha) \geq U(\bar{c}(fc), \bar{h}; \alpha)$,
$h^{*}(\alpha, \xi) = h(w^{v}(h^{*}(\alpha,\xi); \xi), \mu^{v}(h^{*}(\alpha, \xi); \xi); \alpha)$,
$R^{*}(\alpha, \xi) = R(h^{*}(\alpha, \xi); \xi)$.

Given the values of α and ξ which solve the last two equations, the inequality enables us to define a unique lower bound for the value of the fixed cost, i.e. $fc_2^{m}(\alpha, \xi)$. The contribution to the likelihood is:

(16) $$L_2 = \left[-\frac{e_{fc.2}^{m} - [e_\alpha, e_\xi] \sum_{11}^{-1} \sum_{12}}{\sqrt{\sigma_{fc}^2 - \sum_{21} \sum_{11}^{-1} \sum_{12}}} \right] \; {}_2\left[\frac{e_\alpha}{\sigma_\alpha}, \frac{e_\xi}{\sigma_\xi}, \; {}_{\alpha}, \xi \right] J^{-1}$$

where $e_\alpha = \alpha - \bar{\alpha}$, $e_\xi = \xi - \bar{\xi}$, $e_{fc.2}^{m} = fc_2^{m}(\alpha, \xi) - \bar{fc}$, and where J is the determinant of the Jacobian matrix of the transformation of the error terms α and ξ into the observed values h^{*} and R^{*}. While the sub-matrices Σ_{ij} (i = 1; j = 1, 2) are different from their counterparts of the previous section they are defined with the same convention where i correspond to the lines number of the block and j to its column number.

(17) $$J = \frac{{}^{2} h(1 + \delta h)^2}{(1+\delta h)\,((1+\delta h)(1+\gamma) - \beta\delta(2+\delta h)) - \gamma + \delta\gamma\; y_0(2+\delta h)},$$

with $= e^{\xi + \delta h}$.

The likelihood of the self-employed workers can be changed slightly to account for the fact that the revenue lies inside a given range. Since the scale of the revenue function lies in a range $[\xi^{m}, \xi^{M}]$, the likelihood L_{2b} is such that:

(18) $$L_{2b} = \int_{\xi^m}^{\xi^M} \left[-\frac{e_{fc.2}^{m} - [e_\alpha, e_\xi] \sum_{11}^{-1} \sum_{12}}{\sqrt{\sigma_{fc}^2 - \sum_{21} \sum_{11}^{-1} \sum_{12}}} \right] \; {}_2\left[\frac{e_\alpha}{\sigma_\alpha}, \frac{e_\xi}{\sigma_\xi}, \; {}_{\alpha\xi} \right] J_{2b}^{-1} \, d\xi,$$

where J_{2b} results from the changing in variables from α to h^{*}.

Non-participants
The regime is described by the following inequalities:

(19) $U(y_0, 0; \alpha) > U(\bar{c}(fc), \bar{h}; \alpha)$
$\omega(y_0, 0; \alpha) > w^{v}(0; \xi)$.

Given α, it is possible to define a lower bound for the fixed costs, $fc_3^{m}(\alpha)$, and an upper bound for the scale of the revenue function, $\xi_3^{M}(\alpha)$. Therefore, the likelihood becomes:

$$(20)\quad L_3 = \int_{-\infty}^{+\infty} 2\left[\frac{e^{M}_{\xi,3} / \sigma_\xi - \rho_{\alpha\xi}\, e_\alpha / \sigma_\alpha}{\sqrt{1-\rho^2_{\alpha,\xi}}}, -\frac{e^{m}_{fc,3} / \sigma_{fc} - \rho_{\alpha,fc}\, e_\alpha / \sigma_\alpha}{\sqrt{1-\rho^2_{\alpha,fc}}}, -\frac{\rho_{fc,\xi} - \rho_{\alpha,fc}\,\rho_{\alpha,\xi}}{\sqrt{1-\rho^2_{\alpha,fc}}\sqrt{1-\rho^2_{\alpha,\xi}}}\right] \left[\frac{e_\alpha}{\sigma_\alpha}\right] \frac{1}{\sigma_\alpha}\, d\alpha,$$

where $e^{m}_{fc,3} = fc^{m}_{3}(\alpha, \xi) - \bar{fc}$ and with obvious notations.

Our model can account for the unemployed workers, that is, individuals who are willing to work but are rationed at zero hours, provided that the rationing mechanism does not depend on the determinant of the choice. In that case, the probability to be observed as a participant in the primary sector is the product of the probability of being unconstrained at zero hours and L_1. The probability to be observed as a participant in the secondary sector is the product of the probability of being unconstrained at zero hours and L_2. The probability to be unemployed is the product of the probability of being constrained at zero hours and $(1 - L_3)$. Finally, the resulting likelihood contribution must be divided by the probability of being unconstrained so as to account for the truncature of the sample caused by the elimination of unemployed workers. Under the assumption of exogeneity of the rationing process, maximum likelihood estimation of the parameters of the model outlined above, and thus, based on the marginal likelihood only, yields consistent and efficient estimates. Different probabilities of being rationed can also be adopted for unemployed workers in formal and informal sectors, provided that these probabilities are exogeneous with respect to our structural model.