CSAE Working Paper WPS/1994-08

Family Background, Education and Employment in Urban Ethiopia

Pramila Krishnan

WPS/94.8

St Antony's College
Oxford, OX2 6JF

I would like to thank Tasfaye Selassie (ILCA) and the Ministry of Labour and Social Affairs, Addis Ababa for their help with the data used. I would also like to thank Paul Collier, Stefan Dercon and seminar participants at Oxford for their comments.

1. Introduction

A visitor to Addis Ababa is likely to remark not merely on the variety of urban employment but also on the fact that some employments such as shoe-shining seem so slight as to hardly be employment at all. This is a commonplace for any developing country and most economists have concluded that the labour market is segmented into formal and informal sectors; earnings in different sectors differ for workers of equal productivity and that entry into formal employment is rationed. Some have gone further and suggested that the informal sector is not homogenous and that it could be divided into an intermediate sector where entry is also restricted due to limited access to capital and skills and a residual free entry sector that includes shoe-shine boys[1]. Associated with such diversity, is the idea of differential access to jobs and that often access depends on kinship and other networks[2]. More casual observation by the same visitor might lead him to conclude that in a country with a very low enrolment in primary or secondary education and formal employment confined to the public sector, stringent selection must be at work restraining both education and employment. There is no excess supply of schooling; rationing is likely both through location of schools as well as through access by family background or through selective examinations. Similarly, access to employment could be limited by family connections or by selective recruitment on the part of firms. One possibility is that education is used as a screening device by firms particularly in the formal sector so that more highly educated workers are always preferred. This must result in a mis-match between occupations and skills and in the mis-allocation of labour within the formal sector. It also has an impact on the composition of the unemployed; if, as is likely, the educated are also from better-off families it will skew the distribution of the unemployed towards the better-off, educated who can afford to wait for formal sector employment[3].

---

[1]Gindling (1991) distinguishes between the public, private-formal and informal sectors and focuses on testing for segmentation. Both House (1992) and Fields (1990) argue that the diversity of outcomes must be accounted for in any research.

[2]Kannappan (1988) stresses the role such networks play, from providing information to even recruitment and employment. He argues that such networks are efficient if unequal in providing access.

[3]Collier and Lal (1986) argue that this view is too simple and is not sustained in Kenya. The model would suggest that those with the least education would not be recruited; however, this did not hold either in while-collar employment where tests supplemented education in selection- or in manual work where education did not matter. However, Alderman and Kozel (1989) find that the educated young (especially families receiving remittances from the Gulf

The stand taken in this paper is that segmentation is too slippery a concept to be useful or warrant a formal test but that the idea that labour cannot be aggregated across sectors is sensible in a developing country context[4]. The point of departure of the paper is the treatment of unemployment. I argue that it ought to be modelled together with the allocation of labour into the formal and informal sectors. The focus is on the process of labour allocation; to discover to what extent alternative earnings prospects as distinct from background or financial constraints influence the choice of occupation. Outcomes that result from such allocation may be unsatisfactory because access to the better jobs is unequal. Some kinds of households may thus be more likely to face poverty-level incomes than others. In particular, the importance of family background in determining both participation in employment and incomes is stressed. The impact of family background on earnings is usually seen as capturing the quality of the learning environment or proxying omitted ability. It has been suggested that it could also be interpreted as social stratification. For instance, other things being equal, more educated parents might be able to assure access to modern sector employment. Thus, in an imperfect labour market the direct effect of parents' education on the earnings of their children would not necessarily correspond to higher productivity of the children. The key issue therefore is whether it is possible to test the hypothesis that the direct effect of parents' education on childrens' education truly reflects an increase in productivity or not and the last part of this paper is devoted to this issue.

The paper considers the urban labour market in Ethiopia in 1990. This is a year of peculiar interest for it is the last year in power of the previous socialist government which had been in power for 18 years. However, it is also a typical year in that the supply of educational facilities in urban areas stayed the same over the previous decade[5]. In short, the educational structure did not change rapidly (or indeed at all) in this period and so warrants

---

states) prefer unemployment to informal sector employment.

[4]As demonstrated by Cain (1976) and Heckman and Hotz (1989) formal tests of segmentation are difficult to construct and are not informative. Rosenzweig (1988) in his discussion of urban labour markets emphasizes that such characterisation can stem from a variety of reasons and that testing for segmentation is not informative. Competition in the market does not rule out differences in wages across jobs or sectors. Specialization and selection by comparative advantage would mean that wages are specific to jobs or sectors - and hence that aggregation is not possible.

[5]The number of primary schools in 1983 was estimated as 7,215, about twice the number in 1974. However, little had changed since 1979/80. The total number of junior secondary and secondary schools is about 1000. The pupil-teacher ratios vary from 65 on average in primary schools to 41 on average in secondary schools. Participation rates vary from 40% at the primary schools to a low of 7-11% in senior secondary schools.

the cross-section basis of the analysis in this paper. Neither did labour market conditions change dramatically[6].

The following section presents a short-run microeconomic model of labour supply and earnings. The purpose is to build a description of the private sector labour market as well as the unemployed. The results of the estimation are used to build a complete picture of the role of family background in determining educational attainment and earnings in the public and private sectors of the Ethiopian labour market. In passing, the paper also examines the robustness of the usual methods employed in estimating models of polychotomous choice and selection.

2. The Model

The classic model of comparative advantage in the labour market by Roy (1951) assumes that differences in potential earnings determine the choice of occupation. These differences in earnings can be readily justified in a hedonic or linear characteristic approach to earnings as in Tinbergen (1951) and Rosen (1974). In their approach, different firms use attributes such as schooling or experience in production while workers with different latent attributes offer themselves in the market. These differences make workers most productive in particular jobs or sectors and offer possibilities for specialization. Such selection and specialization mean that the prices assigned to the latent attributes by the market are both non-linear and job or sector-specific as Rosen (1983) points out. This rules out the aggregation of labour across different occupations or sectors.

The model specified here is an extension of the Roy model to the case where workers choose occupations based on both the associated earnings as well as non-pecuniary aspects of the job (or sector) such as the status of the job and working conditions. Preferences for the non-pecuniary aspects of jobs vary across individuals and these preferences are correlated with their background characteristics such as their parents' social status, whether they are

[6]However, it is thought that since the takeover, increased pressure will be placed on the urban labour market. About a third of the demobilised soldiers returned to the rural areas, some were found employment but a substantial fraction have stayed in urban areas. This presence together with potential layoffs and an increase (whether one-off or a trend) in migrants from rural areas will change the character of the urban labour market.

married and so on. More important, however, these and other background characteristics might well play a role as constraints on entry into a particular sector. In particular, whether one is a migrant or not or whether one's father works for the government or in trade might well ease entry into particular jobs.

Let Y_{ij} be the potential earnings of person i if sector j is chosen, X_i a vector of human capital variables and ε_i and w_i represent the unobserved talents and tastes of person i. Let V_{ij} denote the value of choosing sector j for person i and Z_i consist of the vector of exogenous variables, notably background and personal characteristics related to the non-pecuniary aspects of jobs.

$$\begin{aligned} Y_{ij} &= y_j(X_i,\varepsilon_i) \quad i=1,2,....n \\ U_{ij} &= u(y_j,Z_i,\omega_i) \quad j=1,2..k \end{aligned} \tag{1}$$

In order to introduce the entry constraints, I use a linearized index function;

$$\begin{aligned} V_{ij} &= U_{ij} + \lambda_j' B_i \\ \text{where} \quad U_{ij} &= \alpha_j' Y_{ij} + \beta_j' Z_i + \omega_i \end{aligned} \tag{2}$$

The value function, V_{ij}, consists of a linearized function of potential earnings and some background variables affecting preferences, while B consists of those background variables that might affect entry. It is the maximum value function obtained in a static optimization of the utility function subject to the entry constraints while the lagrange multiplier denotes the marginal effect of relaxing the constraints posed by background. A person is found in a particular sector if the value function attains its maximum in that sector. The earnings functions within each sector, y_j, are functions of human capital and unmeasured ability; human capital variables influence productivity at home as well. They are specified as semi-log linear functions of human capital variables.

It might well be argued that the effects of constraints and preferences are not readily separable; this cannot be distinguished in such a model. The error terms are assumed to reflect differences in the personal valuation of their home productivity or differences in endowments of specific skills in each sector. The amount of structural information to be gleaned from the actual jobs is limited by the unobservables — ie talent and financial

constraints. In general, it would not be possible to tell if a person was to be found in a particular sector because he was particularly able at that job or because it was easy for him to obtain it through money or family connections. What can be known, by estimating selectivity-corrected earnings is how the net effect of the unobservables operate to find him in that job and with additional information on the workings of the labour market we could produce a better notion of how the jobs and skills are matched. Moreover, as will be seen in the following sections, a careful examination of selectivity-adjusted and unadjusted earnings equations might allow these effects to be distinguished. The worker chooses the classification that maximises the value and is observed in only one sector. The emphasis of the model is on the idea that if observed assignments in sectors are non-random samples of the population, observed earnings will also be non-random samples of potential earnings and the returns to schooling observed in such earnings functions will be biased estimates. The purpose is to obtain selectivity-corrected earnings functions for each sector that will give accurate estimates of the return to schooling.

Finally, it is assumed that the joint distribution of the unobservables is known. In the case of more than two sectors there are two possible distributional assumptions that are tractable. The first and most popular is the multivariate logistic or Gumbel that yields the familiar multinomial logit. The second is the multivariate normal which yields the multivariate probit model which is intractable if there are more than four choices (Maddala, (1983)). Both models can be derived from a model of utility maximization. The multivariate logit is usually employed for it is easy to estimate. Even so, its tractability comes at considerable cost: the substantive assumption that the relative probability of any pair of choices is completely independent of amy other alternatives in the choice set, also known as the assumption of independence of irrelevant alternatives. In other words, there is no room to accommodate closeness or similarity of some alternatives.

Another issue, quite apart from the assumptions of the model, crops up in the estimation of the selectivity terms in the earnings functions. The literature (Maddala, 1983) suggests two methods. A recent paper by Schmertmann (1994), finds that both suffer from serious deficiencies and can be mis-leading. The one method (Lee, 1978) places unacceptable restrictions on the sign of the selectivity coefficient; the other, (Dubin and McFadden, 1984;

Hay, 1980), is likely to have high variances in the face of multicollinearity between regressors, which is usually present[7].

In this paper, the multinomial logit is used as a starting point for the estimates. At the next stage, the assumption of the irrelevance of independent alternatives is checked and the model is re-specified by aggregating two of the choices. The model is then re-estimated using a multivariate probit with the appropriate selectivity corrections, avoiding the problems in estimation in the logit model. Normality is a substantive assumption; but it does offer the advantage of allowing correlation between the unobserved attributes of choices[8].

3. Data, Estimation and Results

The data are drawn from the 'Survey of Adolescent Fertility, Reproductive Behaviour and Employment Status of the Youth Population in Urban Ethiopia', conducted in June 1990. The survey covered 4216 persons between 15 and 30 years old, in 55 urban centres in Ethiopia. It contains information on employment, occupation, education, demographic characteristics and some household characteristics. There is also information on background; on parents' education and occupation.

Table 1 provides some summary data on the entire sample. Of the sample 62% are women: explained mainly by the higher rate of female migration from rural areas, the recruitment of young men by the army and finally, by the possibility that young men are less likely to register with the local *kebele*[9] and hence not be covered by the survey. The table

---

[7]Gindling (1991) tests for segmentation of the urban labour market in Costa Rica. His test turns on the significance of the selectivity-corrections obtained using Lee's method which places unacceptable restrictions on the covariance between errors or the selectivity coefficient.

[8]This distributional assumption has its own burden and has come under increased scrutiny. Heckman (1990), Manski (1988, 1990) both discuss semi- and non-parametric approaches to correcting selectivity biases. Newey *et al.* (1990) present evidence on semi-parametric approaches using kernel methods and conclude that the results obtained are similar to parametric methods but with weaker selectivity effects. The estimates produced in this paper are parametric estimates obtained by using starting values from a maximum score estimation. The values do not differ significantly from a (fully) parametric estimation.

[9]The *kebele* is the urban dwellers' association which undertook most administrative functions and registered residents. Ration cards for use in controlled-price shops were issued by the *kebeles* and to register at the employment exchange required letters from the local kebele. In addition, in May 1983, the National Military Service Proclamation was issued which provided for conscription of young men between 18 and 30. At 16, a young man was required to register with his kebele. A report by Africa Watch (De Waal, 1991), describes the various methods used for conscription in urban

Table 1:
The Distribution of the Sample across Activities by Education and Gender

Activity (%)	Education					Parents' Education			
	male	female	gradel	grade2	grade3	mother		father	
						medn0	medn1	fedn0	fedn1
Student	43.6	36.6	5	23	64	23	15	33	35
Housewife	-	17.6	15	15	21	8	4	18	17
Self-employed	6.8	2.3	12	17	34	24	15	27	20
Private sector	9.0	1.7	12	16	42	36	25	24	27
Domestic service	-	5.5	18	11	7	7	8	15	8
Public sector	15.7	9.0	3	14	74	20	12	35	27
Unemployed	27.0	25.0	7	12	60	18	12	30	31
Total	37.7 (1588)	62.3 (2628)							

medn0, fedn0 = literate or completed primary/junior school
medn1, fedn1 = completed secondary school or higher
Grade1 = 1 if literate/completed primary school
Grade2 = 1 if completed junior school
Grade3 = 1 if completed secondary school or higher

provides the distribution of the sample across different activities and sectors. About half the sample are outside the labour force as either students or housewives. About 15% are in the public sector, 26% are unemployed and the remainder are distributed evenly between private work, self-employment and domestic work. The students are clearly from well-educated families; on average, 35% of the fathers in the sample of unemployed are likely to have secondary or post-secondary education. The unemployed and the public sector employees are similar in background (as captured by the parents' level of education) and the distribution of education is similar to that of the families of the students. Housewives and domestics (almost all of whom are female), however, come from families with far less education: only 12-15% of the mothers in the sample are likely to have had any education. The similarities between

areas, including quotas required to be filled by kebeles. The report describes regular rounds of conscription between 1984 and 1988. In particular, following the capture of Tigray by the Tigrayan Peoples Liberation Front in February 1989, the conscription campaign was stepped up.

Table 2:
Descriptive Statistics

Variable	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.	Mean	Std.Dev.
	labour force (2014)		unemployed (1105)		public sector (492)		private sector (110)		self-employed (171)		domestics (146)	
Human capital variables												
AGE	20.61	3.89	19.33	3.33	23.56	3.35	21.68	3.87	21.63	4.16	19.15	3.32
GRADE1	0.07	0.25	0.07	0.26	0.03	0.34	0.12	0.32	0.12	0.32	0.18	0.38
GRADE2	0.13	0.34	0.12	0.33	0.14	0.34	0.16	0.37	0.17	0.38	0.10	0.31
GRADE3	0.53	0.49	0.54	0.49	0.74	0.39	0.42	0.49	0.34	0.47	0.07	0.26
Background												
MEDN0	0.17	0.37	0.17	0.37	0.19	0.39	0.15	0.36	0.14	0.35	0.08	0.27
MEDN1	0.11	0.31	0.11	0.32	0.11	0.32	0.06	0.24	0.09	0.29	0.07	0.26
FEDN0	0.29	0.45	0.27	0.44	0.34	0.47	0.23	0.42	0.26	0.44	0.15	0.35
FEDN1	0.27	0.44	0.28	0.44	0.26	0.44	0.27	0.44	0.19	0.40	0.08	0.27
FOCC	0.15	0.35	0.14	0.35	0.18	0.38	0.15	0.36	0.07	0.28	0.04	0.21
LIVE1	0.84	0.35	0.87	0.33	0.82	0.38	0.65	0.47	0.84	0.36	0.24	0.43
MIGRANT			0.08	0.28	0.09	0.28	0.23	0.42	0.14	0.35	0.41	0.49
YRES	5.95	4.98	5.45	4.77	6.59	5.11	5.94	4.97	6.88	5.53	4.08	3.47
ELECTRIC	0.88	0.31	0.86	0.34	0.94	0.23	0.94	0.22	0.83	0.37	0.89	0.30
RADIO	0.73	0.44	0.74	0.43	0.74	0.43	0.77	0.42	0.59	0.49	0.82	0.38
CLOCK	0.08	0.28	0.08	0.28	0.10	0.30	0.10	0.30	0.04	0.19	0.16	0.37
COOKER	0.19	0.39	0.21	0.41	0.19	0.39	0.09	0.28	0.09	0.29	0.30	0.46
PHONE	0.14	0.35	0.16	0.37	0.10	0.31	0.18	0.38	0.10	0.30	0.28	0.45
FRIDGE	0.10	0.30	0.11	0.31	0.07	0.27	0.11	0.32	0.09	0.30	0.23	0.42
TV	0.13	0.33	0.15	0.35	0.10	0.30	0.09	0.28	0.10	0.30	0.28	0.45
Wages/hour, hours/day, days/week, income/year												
HRS	8.50	2.5	-	-	7.9	2.49	8.83	2.47	8.80	2.58	9.80	2.41
DAYS	6.22	1.0	-	-	6.1	0.88	6.38	0.89	5.82	1.59	6.88	0.49
WAGE1	1.0	2.17	-	-	1.34	2.70	0.47	0.37	1.07	1.73	0.17	0.38
ANNINC	1958	3066	-	-	2472	3434	1168	817	2268	3576	459	564

Grade1=1 if literate/primary school
Grade2=1 if completed junior school
Grade3=1 if completed secondary school or higher

Medn0, Fedn0=1 if mother/father literate or completed primary/junior school
Medn1, Fedn1=1 if mother/father completed secondary school or higher
Focc=1 if father is in white-collar job

Leve1=1 if living at home with parents
Other variables are dummies taking the value 1 if the household possesses the goods

those unemployed and those in public sector employment show up more clearly in Table 2. About half of the unemployed have completed secondary school or higher, as have 62% of those in the public sector compared with an average of 35-42% for the private sector and the self-employed.

The public sector workers are slightly older, more likely to be married and more likely to be living with their parents. Seventy-five percent had completed junior school or less as compared with about half for the private sector or the self-employed. In profile, they are very similar to the unemployed. The father is more likely to be literate or possess some schooling.

The unemployed are very similar to the public sector workers in background. More than half have been unemployed for over three years (and over 80% of them have completed junior or secondary school). Most preferred non-manual or clerical employment[10]. Data on the duration of unemployment was available for about half the sample of unemployed; unsurprisingly, migrants and those living with their parents are likely to be unemployed longer.

Of the self-employed, 65% are sales-workers (retailers or street-vendors), about 12% are tailors or food and beverage sellers and the rest work as mechanics, shoemakers, carpenters, masons and in services. About half relied on their own savings to start the business, while 20% used an inheritance and the rest borrowed from friends and family. About three-quarters had invested less than Birr 500 in the business.

Women in domestic service (for almost all are women), are the most likely to have migrated in search of a job and earn far less (a third of the private sector's earnings on average) than any of the other groups. They are far less likely to have completed any level of schooling, and their parents are more likely to be illiterate. Oddly enough, they are more likely to own a cooker or a radio or even a fridge than any of the other groups[11].

The workers in the private sector have, on average, a lower level of education, earn less than those in the public sector or those self-employed and are more likely to have

[10]The effect of consumer goods shortages on labour supply has been investigated for rural labour markets but rarely in urban labour markets. A study using data on Soviet workers (Mokhtari and Gregory, 1993) suggests that the presence of quantity constraints causes workers to reduce their hours of work. It is conceivable that a similar rationing of consumer goods affects partcipation by educated unemployed in Ethiopia.

[11]There is no distinction made in the survey between live-in domestics and others. It is possible that such assets are available only to live-in domestics.

migrated in search of work than either of these groups. Like the domestic workers they are less likely to live with their parents.

4. Estimation

The first stage is the estimation of a model of entry into the different types of work for pay. There are four sectors of interest: the public, the private, self-employment and unemployment. Domestic service has been excluded since it appears to be a very particular choice and quite independent of allocation in the rest of the labour market. Another option would be to model the choices of men and women separately but a cursory attempt at that appeared not to yield results that were particularly different[12]. At this stage, since I have four possible choices, I use a multinomial logit model. These estimates are presented in Table 3. Unemployment is included as an option and the estimates are based on contrasts between work for pay and being unemployed[13].

The estimated model includes human capital variables such as age and education, personal and demographic characteristics such as sex, whether the person is married (MARR), or is a migrant who came in search of work (MIGG1). It also includes the number of years resident in the town (YRES which is equal to age if the person is a lifetime resident) and whether the person lives at home with the parents (LIVE1). The other variables are background characteristics including the mother's education (MEDNL), the father's occupation (FOCC1 which takes the value 1 if the father is in white collar work and has secondary school education) and the father's education (FEDN0 which takes the value 1 if

[12]The results presented are conditional upon participation. Over half the sample are non-participants, divided between housewives and students. A probit of the choice of participation showed an insignificant selectivity effect. Assuming that partcipation and the occupational choice are independent (admittedly heroic but if not opening up evil estimation problems) means that the estimation of the earnings function remains unbiased. A sufficient condition for this to hold is that the unobservables in the participation decision and occupational choice are independent.

[13]Heckman and Sedlacek (1990) find that allowing for a non-market sector is very important in fitting wage distributions. It might be argued that unemployment is a slippery concept in a developing country and the number of non-participants, including students should be used instead. I have restricted the unemployed to those who either said they were actively looking for work or said they did not have work. Both housewives and students have other support networks which the unemployed may not possess.

the father has completed primary school). The variable, REGION, captures the effect of living in the capital, Addis Ababa, as opposed to other urban areas[14].

Table 3:
Multinomial Logit Model with Four Outcomes

Log-Likelihood.....		-1560.547	Restricted (Slopes=0) Log-L =		-1961.644	
Chi-Squared (39).....		802.1937	Significance Level.....		0.00000	
Variable	Coefficient	Std. Error	t-ratio	Prob/t/≥x	Mean of X	Std.Dev. of x
Estimates for the public sector versus unemployment						
Constant	-5.9455	0.5231	-11.365	0.00000		
Human capital variables						
AGE1	2.6929	0.1962	13.727	0.00000	2.0815	0.39132
GRADE1	0.49467	0.3733	1.325	0.18510	0.61563E-01	0.24042
GRADE2	1.2312	0.2679	4.595	0.00000	0.13865	0.34567
GRADE3	1.1931	0.2207	5.406	0.00000	0.60814	0.48830
Personal/Demographic characteristics						
SEX	-0.43798	0.1322	-3.312	0.00093	0.55353	0.49726
MARR	1.4059	0.2086	6.740	0.00000	0.13383	0.34056
LIVE1	-1.5269	0.2265	-6.740	0.00000	0.89454	0.30723
MIGG1	-0.38916	0.2721	-1.430	0.15267	0.85653E-01	0.27993
YRES1	0.36018E-03	0.1484E-03	2.426	0.01525	-603.54	488.93
REGION	-0.13432	0.1410	-0.953	0.34077	0.34208	0.47453
Family Background variables						
MEDNL	0.63580E-01	0.1489	0.427	0.66948	0.29818	0.45758
FEDN0	0.15195	0.1450	1.048	0.29470	0.30300	0.45968
FOCC1	0.29977	0.2210	1.356	0.17496	0.10707	0.30928
Private sector employment relative to unemployment						
Constant	-2.7353	0.7239	-3.779	0.00016		
Human capital variables						
AGE1	1.3055	0.3021	4.321	0.00002	2.0815	0.39132
GRADE1	0.43770	0.4000	1.094	0.27390	0.61563E-01	0.24042
GRADE2	0.30579	0.3474	0.880	0.37867	0.13865	0.34567
GRADE3	-0.42937	0.2891	-1.485	0.13756	0.60814	0.48830

cont ...

[14]Fifty-five towns were surveyed but they differ markedly in size and importance and an attempt to aggregate them into groups produced little in the way of contrasts. The use of other regional dummies made no difference and since Addis versus other towns presents the sharpest contrast it was decided to retain this variable.

Table 3 cont ...

Variable	Coefficient	Std. Error	t-ratio	Prob/t/≥x	Mean of X	Std.Dev.of x
Personal/Demographic characteristics						
SEX	-0.88335	0.2183	-4.046	0.00005	0.55353	0.49726
MARR	1.2690	0.3237	3.920	0.00009	0.13383	0.34056
LIVE1	-2.2568	0.2811	-8.028	0.00000	0.89454	0.30723
MIGG1	0.68785	0.3323	2.070	0.03846	0.85653E-01	0.27993
YRES1	-0.64995E-04	0.2641E-03	-0.246	0.80563	-603.54	488.93
REGION	0.17393	0.2283	0.762	0.44606	0.34208	0.47453
Family Background variables						
MEDNL	-0.13210	0.2754	-0.480	0.63147	0.29818	0.45758
FEDNO	-0.13203	0.2552	-0.517	0.60486	0.30300	0.45968
FOCC1	0.30751	0.3960	0.776	0.43746	0.10707	0.30928
Self-employment relative to unemployment						
Constant	-3.0991	0.6110	-5.072	0.00000		
Human capital variables						
AGE1	1.6147	0.2435	6.632	0.00000	2.0815	0.39132
GRADE1	0.11635	0.3181	0.366	0.71455	0.61563E-01	0.24042
GRADE2	-0.20028	0.2697	-0.742	0.45779	0.13865	0.34567
GRADE3	-1.2731	0.2300	-5.536	0.00000	0.60814	0.48830
Personal/Demographic characteristics						
SEX	-1.2036	0.1811	-6.647	0.00000	0.55353	0.49726
MARR	0.76158	0.2896	2.630	0.00854	0.13383	0.34056
LIVE1	-1.1241	0.2879	-3.905	0.00009	0.89454	0.30723
MIGG1	0.18511	0.3090	0.599	0.54917	0.85653E-01	0.27993
YRES1	-0.17082E-03	0.2110E-03	-0.810	0.41814	-603.54	488.93
REGION	-0.62113E-01	0.1886	-0.329	0.74192	0.34208	0.47453
Family Background variables						
MEDNL	0.21659	0.2171	0.998	0.31842	0.29818	0.45758
FEDN0	-0.26339E-01	0.2005	-0.131	0.89551	0.30300	0.45968
FOCC1	-0.26764	0.3884	-0.689	0.49083	0.10707	0.30928

YRES1 = YRES/10
AGE1 = AGE/10

The point of estimating a multinomial logit model of choice is to examine whether the sectors are particularly different. It is often suggested that self-employment is a entirely different choice and it is usually treated separately. An obvious question, upon inspecting the estimates is whether any of the alternatives are similar or even observationally identical. Using the test

developed by Cramer and Ridder (1991), I checked whether private sector employment and self-employment could be pooled[15]. The results are in Table 4 below.

Table 4:
Test of Pooling

The null hypothesis is that the two states (private sector and self-employment) have the same regressor coefficients

H0: $b_{priv} = b_{self}$

Log-likelihood for pooled model =	-1386.04	
Log-likelihood for original model =	-1560.547	
Restricted log-likelihood (under H0) =	-1386.04-188.100 =	-1574.14
Likelihood ratio value =	27.1	
Test statistic with 24 df =	32.67 (5%)	

Conclusion: fail to reject H0

The conclusion, that the two states can be pooled suggests two interpretations. Most current writing on employment in developing countries suggests that the informal sector is heterogenous in both opportunities and earnings and that an important research issue is to conceptualise this heterogeneity[16]. The statistical finding that quite different groups can be pooled on the basis of education, demographic characteristics and background is that standard measures are unlikely to reveal such heterogeneity as does exist. The other interpretation is that if states can thus be pooled might indicate that the opportunities are part of a continuum in the private sector and that self-employment in such an economy is part of the whole rather than a segment apart. Given that the data do not contain any information on household structure, income or assets it is not possible to draw any firm conclusions about this result. It does, however, allow the comparison between sectors to be reduced to a three-way comparison between unemployment, public employment or private employment.

[15]Testing for the assumption of independence of alternatives in this model proved infeasible. Testing whether alternatives are identical is a much stronger test; and comparisons of the other groups suggested they could not be pooled. This does not rule out similarity in unobserved attributes of alternatives.

[16]House (1992) emphasises the need for capturing the diversity in occupations. "The diversity of activities in the informal sector must also be accounted for. Evidently, the more sub-sectors incorporated into the model, the greater becomes the complexity of conceptualizing behavioural patterns." He concludes that such modelling requires greater in-depth information and data from well-designed household surveys.

The second step was to test for assumption of the independence of irrelevant alternatives, using the Hausman test. A standard problem with this test is that it is often difficult to obtain the inverse of the variance-covariance matrix; a problem that cropped up in this case as well. However, the fact that the comparison is now reduced to a three-way one makes it feasible to move away from the logistic distribution with its strong assumptions about the independence of alternatives to the multivariate normal which is more satisfactory for it allows alternatives to be correlated (and therefore for the assumption to be tested more readily). The other advantage of moving away from the logistic distribution is that the selectivity-corrections in the earnings functions can be estimated without the problems of restrictions on their signs or inefficiency in estimates.

Consider the following model of sectoral choice with three sectors: public, private and the non-market (unemployed/non-participants) sector. As before, V_{ij}, represents the linearized index of value of person i in sector j. Let V_{pub} denote the value of being in the public sector, V_{pri}, the value in the private sector and V_{unem}, that while unemployed. A person is found in the public sector if V_{public} is larger than $V_{private}$ or V_{unem}. Since only parameters of contrasts of the value are identified, the value of the non-market sector is normalized to zero. The error terms then follow a bivariate normal distribution and the structure of the wage function is as specified earlier.

5. Results

An examination of Table 3, with the estimates of the multinomial model, reveals that background as captured by the parent's education and occupation does not affect entry at all. Education, however, is very important in determining entry, especially into the public sector. In order to check whether educational outcomes are influenced by background, I estimated a simple probit model of education using the parents' education and occupation as explanatory variables and controlling for age, sex and region. The results, in Table 5, show that background is very important in explaining education.

the education variables in order to see the direct effect of background on entry into sectors (relative to being unemployed).

Table 6:
Estimates of the Bivariate Probit Model

Log-Likelihood..... -1449.219
N[0,1] used for significance levels.

Variable	Coefficients	Std. Error	t-ratio	Prob/t/≥x	Mean of X	Std.Dev of X
The public sector relative to unemployment						
Constant	-3.3385	0.2.593	-12.873	0.00000		
AGE1	1.4312	0.1068	13.403	0.00000	2.0815	0.39132
SEX	-0.93006E-01	0.6902E-01	-1.347	0.17783	0.55353	0.49726
MARR	0.52319	0.1011	5.175	0.00000	0.13383	0.34056
MIGG1	-0.52356	0.1303	-4.018	0.00006	0.85653E-01	0.27993
YRES1	-0.11710E-01	0.4689E-02	-2.497	0.01252	14.675	8.0667
REGION	-0.60893E-01	0.7285E-01	-0.836	0.40326	0.34208	0.47453
LIVE1	-0.34624	0.1130	-3.065	0.00218	0.89454	0.30723
MEDNL	0.15118	0.7975E-01	1.896	0.05802	0.29818	0.45758
FEDN0	0.18935	0.7570E-01	2.501	0.01237	0.30300	0.45968
FOCC1	0.26272	0.1335	1.968	0.04904	0.10707	0.30928
The private sector relative to unemployment						
Constant	-0.65520	0.259	-2.523	0.01163		
AGE1	0.36233E-01	0.1144	0.317	0.75153	2.0815	0.39132
SEX	-0.38205	0.7573E-01	-5.045	0.00000	0.55353	0.49726
MARR	0.69847E-01	0.1143	0.611	0.54125	0.13383	0.34056
MIGG1	0.63264	0.1336	4.632	0.00000	0.85653E-01	0.27993
YRES1	0.13249E-01	0.5371E-02	2.467	0.01364	14.675	8.0667
REGION	0.30126E-02	0.7877E-01	0.038	0.96949	0.34208	0.47453
LIVE1	-0.56362	0.1163	-4.845	0.00000	0.89454	0.30723
MEDNL	-0.54792E-01	0.8935E-01	-0.613	0.53973	0.29818	0.45758
FEDN0	-0.11411	0.8224E-01	-1.387	0.16530	0.30300	0.45968
FOCC1	-0.17687	0.1480	-1.195	0.23201	0.10707	0.30928
RHO (1,2)	-0.98572	0.2653E-01	-37.148	0.00000		

The results of the estimation of the bivariate probit model of entry are in Table 6. The estimates show that background is very important in determining entry into the public sector. Educated parents and a father who works in a white collar job increase the chance of being

Table 5:
Ordered Probit Model of Educational Outcomes

(Dependent variables; =1 if completed grade1, =2 if completed junior school, =3 if completed secondary school or higher)

Log-likelihood.....	-1829.456
Restricted (Slopes=0) Log-L.....	-1988.759
Chi-Squared (7).....	318.6075
Significance Level.....	0.0000000

Cell Frequencies for Outcomes

Cell	Count	Rel.Freq.
0	358.	0.19165
1	115.	0.06156
2	259.	0.13865
3	1136.	0.60814

Variable	Coefficients	Std. Error	t-ratio	Prob/$t \geq x$	Mean of X	Std.Dev of X
Constant	0.49258	0.1585	-3.108	0.00188		
MEDN0	0.32084	0.8591E-01	3.735	0.00019	0.18148	0.38552
MEDN1	0.57097	0.1223	4.669	0.00000	0.11670	0.32115
FEDN0	0.57674	0.6974E-01	8.270	0.00000	0.30300	0.45968
FEDN1	0.77708	0.8803E-01	8.828	0.00000	0.28480	0.45144
FOCC	0.33489	0.9523E-01	3.517	0.00044	0.16006	0.36676
REGION	0.12257	0.6232E-01	1.967	0.04921	0.34208	0.47453
AGE1	0.42069	0.7101E-01	5.924	0.00000	2.0815	0.39132
MU (1)	0.23468	0.2091E-01	11.222	0.00000		
MU (2)	0.68058	0.3133E-01	21.772	0.00000		

The results are not unexpected. Parental education might be acting as proxy for the effects of other variables such as income and wealth. It might also be strongly correlated with unmeasured ability or capture the omitted effect of the quality of the learning environment at home[17]. The implication for the model of entry into a sector is that the level of education chosen is endogenous to the choice of job or sector in the labour market[18].

The bivariate probit model of sectoral choice was estimated in the light of the finding that background determines educational outcomes. The sectoral choice function now omits

[17] However, it must be mentioned that entry into higher education is rationed. The percentage of secondary school students who are admitted to colleges or universities has declined, from 16% in 1981 to about 10% in 1990.

[18]For instance, Kenny *et al.* (1979) deal with this by modelling years of college education as endogenous to the choices in the labour market.

in the public sector. Background in the form of the parents' education or occupation appears to have no impact in the private sector at all. There is a slight negative effect to having a father in a white-collarjob and it is emphasised that the estimates reveal marginal effects relative to being unemployed. There is a strong effect from living with one's parents. It is negative in both cases indicating that the unemployed tend to live with their parents and so have a safety net[19]. Migrants are more likely to enter the private sector - but the variable measuring years in residence suggests that people who have been resident longer are more likely to find jobs in the private sector. This is exactly the opposite in the public sector where migrants are less likely to find work - but the probability of entry falls with length of residence (or search). The variable REGION, which captures the effects of living in the capital, Addis Ababa is significant for private sector employment. Clearly, the capital does possess the most opportunities for small enterprise.

Gender does not matter in the public sector; but women are much less likely to enter the private sector. This finding, together with the observation that domestic service workers are almost all women suggests that there is occupational segregation. Marriage and corresponding need to set up an independent household differentiates the unemployed from the public sector workers.

The correlation coefficient is negative indicating that the unobserved attributes of jobs or skills which propel a worker into public sector work also raise the likelihood of being unemployed relative to being in the private sector.

Table 7 provides estimates of the selectivity-corrected earnings functions. These include the (log of) age (AGEL), gender (SEX), and education (GRADE, GRADE2, GRADE3). Age and gender do not affect earnings in the public sector. There are returns to experience in the private sector but women earn far less. The estimated returns to post-secondary education are extremely high especially in the public sector and are twice that in the private sector. The private sector demonstrates high returns for junior school education as opposed to the public sector where the returns are negligible. Age and gender are insignificant in the public sector but very significant in the private sector. The selectivity

[19] Clearly, this variable is likely to be endogenous. This was tested using (an extension of) the Smith and Blundell (1986) test of (weak) exogeneity. However, the results were inconclusive because the variance-covariance matrix of the errors proved to be singular. The variable is used to proxy the existence of family networks or family support in the absence of information on assets or income.

coefficients are very significant indicating that omitted ability or quality of education matters in determining returns[20].

Table 7:
Wage Functions Corrected for Selectivity Bias

Wages in the public sector

Selected sample contains 492 observations.

Adjusted R-squared = 0.3029957E+00
Estimated disturbance standard deviation = 1.08670
Estimated correlation with selection equation A = 0.546224
Estimated correlation with selection equation B = 0.572855

Variable	Coefficients	Std. Error	t-ratio	Prob/t/≥x
Constant	-0.96807	3.504	-0.276	0.78236
AGE1	0.32161	1.007	0.319	0.74954
RGION	-0.97887E-02	0.1434	-0.068	0.94558
SEX	-0.12492	0.1397	-0.894	0.37130
MARR	-0.11192	0.1658	-0.675	0.49965
GRADE1	0.10269E-01	0.2566	0.040	0.96808
GRADE2	0.79318E-01	0.1311	0.605	0.54520
GRADE3	0.79200	0.1217	6.507	0.00000
Lambda-A	-0.70684	0.3893	-1.815	0.06945
Lambda-B	1.3193	0.5995	2.201	0.02776

6. The Impact of Background on Earnings

The casual observation of the visitor that intergenerational mobility is low in Ethiopia is the hypothesis of interest in this section. The results obtained in the selectivity-corrected equations display high returns to education in both sectors. The selectivity correction incorporates the non-linear effect of background and demographic variables upon earnings. Clearly, family background is important in both educational attainment and earnings and this section presents evidence of the direct effect of background on earnings.

[20]The coefficients cannot readily be used to conclude the presence of positive comparative advantage in this bivariate probit model because the correlation coefficients is high and significant (Fishe et al, 1981). It must be emphasised that this is quite unlike the possible interpretation in the usual probit model.

The estimates in Table 8 include parents' education in the selectivity-adjusted equations. There is no (further) direct effect on earnings indicating that any externality due to the parents' education disappears once selection is accounted for.

Table 7 cont ...

Wages in the private sector
Adjusted R-squared = 0.1911036E+00
Selected sample contains 281 observations.
Estimated disturbance standard deviation = 1.24053
Estimated correlation with selection equation A = 0.268423
Estimated correlation with selection equation B = 0.281255

Variable	Coefficients	Std. Error	t-ratio	Prob/t/≥x
Constant	-6.4208	2.473	-2.596	0.00942
AGEL	1.5228	0.7577	2.010	0.04446
REGION	0.20192	0.2535	0.797	0.42568
SEX	-0.76629	0.2731	-2.805	0.00503
MARR	0.22192E-01	0.3539	0.063	0.95001
GRADE1	-0.68744E-01	0.1834	-0.375	0.70782
GRADE2	0.40715	0.1724	2.362	0.01819
GRADE3	0.30882	0.1554	1.988	0.04686
Lambda-A	-0.38557	0.6774	-0.569	0.56924
Lambda-B	0.72896	0.4192	1.739	0.08201

Table 8:
Effect of Background on Wages after Correcting for Selectivity

Wages in the public sector Adjusted R-squared = 0.3026430E+00

Variable	Coefficients	Std. Error	t-ratio	Prob/t/≥x
Constant	-2.4011	3.398	-0.707	0.47982
AGEL	0.71366	0.9678	0.737	0.46086
SEX	-0.13624	0.1193	-1.142	0.25342
MARR	-0.58543E-01	0.1542	-0.380	0.70413
GRADE1	-0.59535E-02	0.2438	0.024	0.98052
GRADE2	0.55244E-01	0.1347	0.410	0.68181
GRADE3	0.75705	0.1222	6.194	0.00000
REGION	-0.23832E-01	0.1215	-0.196	0.84448
MEDNL	0.60716E-01	0.1282	0.474	0.63569
FEDNL	0.77835E-01	0.1058	0.736	0.46186
Lambda-A	-0.54968	0.3685	-1.492	0.13574
Lambda-B	1.1835	0.5347	2.213	0.02688

Table 8 cont ...

Wages in the private sector				Adjusted R-squared = 0.1897258E+00
Variable	Coefficients	Std. Error	t-ratio	Prob/t/≥x
Constant	-6.4509	2.369	-2.692	0.00709
AGEL	1.5428	0.7343	2.101	0.03564
SEX	-0.77821	0.2654	-2.932	0.00336
MARR	0.19136E-01	0.3446	0.056	0.95572
GRADE1	-0.58459E-01	0.1822	-0.321	0.74838
GRADE2	0.39519	0.1698	2.328	0.01991
GRADE3	0.29569	0.1482	1.995	0.04608
REGION	0.19311	0.2463	0.784	0.43300
MEDNL	0.18401	0.2990	0.615	0.53833
FEDNL	-0.40123E-01	0.2036	-0.197	0.84380
Lambda-A	-0.36749	0.6516	0.564	0.57277
Lambda-B	0.70103	0.4143	1.692	0.09065

Table 9:
Wage Functions Unadjusted for Selectivity (but including family background)

Wages in the public sector						Adjusted R-squared = 0.2922587E+00
Variable	Coefficients	Std. Error	t-ratio	Prob/t/≥x	Mean of X	Std.Dev of X
Constant	-6.6936	0.7693	-8.701	0.00000		
AGEL	1.8709	0.2461	7.603	0.00000	3.1493	0.14838
GRADE1	-0.11395	0.2178	-0.523	0.60111	0.32520E-01	0.17756
GRADE2	0.35152E-01	0.1473	0.239	0.81144	0.14024	0.34759
GRADE3	0.76384	0.1267	6.029	0.00000	0.74390	0.43692
SEX	-0.13739	0.6681E-01	-2.056	0.04028	0.50407	0.50049
MARR	0.86420E-01	0.8016E-01	1.078	0.28156	0.30691	0.46168
MEDNL	0.11471	0.7614E-01	1.507	0.13258	0.31504	0.46500
FEDNL	0.14485	0.7509E-01	1.929	0.05433	0.61382	0.48737
REGION	-0.68928E-01	0.7232E-01	-0.953	0.34100	0.32114	0.46739

cont ...

Table 9 cont ...

Wages in the private sector Adjusted R-squared = 0.1629783E+00

Variable	Coefficients	Std. Error	t-ratio	Prob/t/≥x	Mean of X	Std.Dev of X
Constant	-4.2638	1.022	-4.174	0.00004		
AGEL	1.0931	0.3374	3.240	0.00134	3.0575	0.18814
GRADE1	-0.62310E-01	0.1961	-0.318	0.75095	0.11744	0.32252
GRADE2	0.47216	0.1732	2.726	0.00684	0.17082	0.37702
GRADE3	0.40095	0.1482	2.705	0.00726	0.37367	0.48464
SEX	-0.58430	0.1237	-4.723	0.00000	0.38434	0.48731
MARR	0.76263E-01	0.1660	0.459	0.64638	0.17794	0.38314
MEDNL	0.22980	0.1503	1.529	0.12739	0.23132	0.42243
FEDNL	0.21524E-01	0.1297	0.166	0.86836	0.48399	0.50064
REGION	0.22892	0.1261	1.822	0.06958	0.33808	0.47390

Table 10:
Wages Functions Unadjusted for Selectivity (excluding family background)

Wages in the public sector Adjusted R-squared = 0.282555E+00

Variable	Coefficients	Std. Error	t-ratio	Prob/t/≥x	Mean of X	Std.Dev of X
Constant	-6.4409	0.7684	-8.382	0.00000		
AGEL	1.8081	0.2463	7.340	0.00000	3.1493	0.14838
GRADE1	-0.11366	0.212	-0.519	0.60431	0.32520E-01	0.17756
GRADE2	0.86854E-01	0.1469	0.591	0.55462	0.14024	0.34759
GRADE3	0.85236	0.1234	6.905	0.00000	0.74390	0.43692
SEX	-0.13119	0.6717E-01	-1.953	0.05139	0.50407	0.50049
MARR	0.64708E-01	0.8013E-01	0.808	0.41974	0.30691	0.46168
REGION	-0.67623E-01	0.7271E-01	-0.930	0.35280	0.32114	0.46739

Wages in the private sector Adjusted R-squared = 0.1604422E+00

Constant	-4.0853	1.017	-4.015	0.00008		
AGEL	1.0416	0.3365	3.096	0.00217	3.0575	0.18814
GRADE1	-0.68562E-01	0.1959	-0.350	0.72660	0.11744	0.32252
GRADE2	0.50721	0.1711	2.964	0.00330	0.17082	0.37702
GRADE3	0.45187	0.1401	3.226	0.00141	0.37367	0.48464
SEX	-0.55172	0.1222	-4.513	0.00001	0.38434	0.48731
MARR	0.79909E-01	0.1661	0.481	0.63088	0.17794	0.38314
REGION	0.24505	0.1260	1.945	0.05284	0.33808	0.47390

In Tables 9 and 10, I present the simple (unadjusted) OLS estimates of returns to education; 0 Table 9 includes parents' education while Table 10 does not. The comparison is illuminating. The estimates in Table 9 show that the father's education has a strong positive effect on earnings, while the mother's education is (weakly) significant in both sectors. As pointed out by Heckman and Hotz (1989)[21], the meaning of such a finding is ambiguous. It could mean that background proxies the quality of education or it might indicate that society is stratified and the labour market is nepotistic and only the better-off have access to education. The Heckman and Hotz remark was made in reference to their finding of high (direct) returns to background; however, these were (uniquely) unadjusted for a Heckman selectivity correction and were the results of an OLS estimation. Since no returns to background can be observed after correcting for selection, an unambiguous conclusion can be drawn that background effects are returns to family connections rather than proxies for a better learning environment at home.

The comparison of the constant terms in Tables 7 and 10 (or Tables 8 and 9) also bolsters this conclusion. The constant terms in such a regression can be interpreted as base wages before allowing for return to human capital. Correction for selectivity drives these base wages (or sector-specific) rents further apart, indicating that the rent to being in the public sector increases after accounting for selection, supporting the idea that there are returns to family connections.

The returns to (own) education in either sector fall slightly with the addition of the direct effect of background, but the difference is insignificant. It is clear that the selectivity effects, if seen primarily as omitted variable effects are dominated by the parents' background. The fact that returns to education fall but negligibly suggests that returns to background are, in the main, returns to family connections.

Another possibility is that the impact of background is primarily an econometric artifact. It can be shown that if schooling is measured with error, standard estimates of returns are biased downwards. Furthermore, if omitted variables like ability are correlated with the measurement errors and are actually included in the regression, the bias downward

[21]Heckman and Hotz examined the effect of background on earnings for a sample of men in Panama and obtained similar results. Lam and Schoeni (1993) ask whether family background is responsible for the high returns to education in Brazil and find that returns to schooling fall by a third. They suggest that assortive mating is a more plausible explanation than nepotism in the labour market.

actually increases (Griliches, 1977; Lam and Schoeni, 1993)[22]. The evidence here is that in both sectors, the estimated returns hardly fall which is not consistent with the presence of measurement error in schooling.

7. Conclusions

Family background plays an important part in determining occupation in urban Ethiopia. Returns to education are very high even after controlling for omitted effects through the use of family background. Parental education has no direct effect on earnings nor does it affect returns to own education significantly. If it did capture the effect of unobserved characteristics that increase labour productivity or the inherited part of ability one would expect returns to own education to fall as well. The fact that after adjusting for selection there are no direct effects at all strongly supports the fact that access to formal sector employment is restricted by background and family connections.

The relationship between unemployment and employment, especially formal employment suggests that unemployment is not, by itself, a clear indicator of poverty. Employment in the private sector is chosen by those with lower levels of education than in either unemployment or the public sector. Clearly, both public sector employees and the unemployed come from a similar background and are more likely to have access to support networks.

This relationship is likely to have changed since the takeover in 1991. Domestic workers who have very low incomes are primarily single women who have migrated in search of jobs. It is possible that more men will now migrate to urban centres for they were more likely to have been deterred from doing so in the 1980s. Both demobilised soldiers (whose income support from foreign donors has ended) and those without family support in urban

[22]In the equation, A represents ability which is unmeasured (or indeed any omitted variable) while B represents family background. β_s is the true coefficient on schooling while β_s^* is the estimated returns to schooling. If background and ability are highly correlated, the returns estimated move closer to the true effect.

$$y_{ij} = \beta_0 + \beta_s S + (\beta_a A_{ij} + \varepsilon_{ij})$$
$$plim\ \beta_{s.b}^* = \beta_s + \beta_a(1 - \rho_{ab.s}^2)$$

areas are likely to seek informal or private sector employment. In short, the composition of the unemployed and urban poor will change dramatically from the educated, better-off unemployed to a much poorer, less-skilled group for it is not clear that opportunities in the private sector will be available unless the private sector is encouraged. It is clear that outcomes in the private sector are relatively homogenous (except for domestic workers who are unmarried women who earn far lower wages and work longer hours than any other group). There is no case for suggesting that self-employment is a particularly different outcome from wage employment or indeed that the self-employed possess (measured) characteristics that differ profoundly from the rest in private work. If there is heterogeneity in the informal sector it ought to be captured using more sophistication than is usual in a labour market survey. Surveys would do better to obtain information on the structure of households in urban areas, mobility, family assets, movements in and out of work and the importance of informal networks. Education and demographic characteristics alone, cannot pick up the alleged heterogeneity of activities that both casual observers and the experts comment on.

Finally, a brief remark about the methods of estimation. Polychotomous choice models are notoriously difficult to estimate and the corrections for selectivity usually employed are not reliable. Furthermore, given the strong restrictions on testing for similarity of sectors it is difficult to see that the standard specification of sectors and choice can reveal heterogeneity in the private informal sector.

References

Akerlof, G. and J. Yellen (1986) 'Efficiency Wage Models of the Labour Market', London: Cambridge University Press.

Alderman, H. and V. Kozel (1989) 'Formal and Informal Sector Wage Determination in Urban Pakistan', Living Standards Measurement Study Working Paper, No. 65, Washington DC: World Bank.

Collier, P. and D. Lal (1986) *Labour and Poverty in Kenya 1900-1980*, Oxford: Oxford University Press.

Fishe, R.P.H., R.P. Trost, and P.M. Lurie (1981) 'Labor Force Earnings and College Choice of Young Women: an Examination of Selectivity Bias and Comparitive Advantage', *Economics of Education Review*, Vol. 1, No. 2, pp. 169-91.

Gindling, T.H. (1991) 'Labour Market Segmentation and the Determination of Wages in the Public, Private-Formal and Informal Sectors in San Jose, Costa Rica', *Economic Development and Cultural Change.*

Griliches, Z. (1977) 'Estimating the Returns to Schooling: Some Econometric Problems', *Econometrica*, Vol. 45, pp. 1-22.

Heckman, J.J. (1990) 'Varieties of Selectivity Bias', *American Economic Review*, Papers and Proceedings, Vol. 80, No. 2, pp. 313-19.

Heckman, J.J. and V.J. Hotz (1989) 'An Investigation of the Labour Market Earnings of Panamanian Males: Evaluating the Sources of Inequality', *The Journal of Human Resources*, Vol. 21, No. 4, pp. 507-63.

Heckman, J.J., and G.L. Sedlacek (1990) 'Self-Selection and the Distribution of Hourly Wages', *Journal of Labour Economics*, Vol. 8, No. 1, pp. 329-63.

House, W.J. (1993) 'Priorities for Urban Labour Market Research in Anglophone Africa', *Journal of Developing Areas*, Vol. 27, pp. 49-68.

Kannappan, S. (1989) 'Urban Labour Markets and Development', *World Bank Research Observer.*

Lam, D. and R.F. Schoeni (1993) 'Effects of Family Background on Earnings and Returns to Schooling: Evidence from Brazil', *Journal of Political Economy*, Vol. 101, No. 4, pp. 711-37.

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

Mokhtari, M. and P.R. Gregory (1993) 'Backward Bends, Quantity Constraints and Soviet Labour Supply: Evidence from the Soviet Interview Project', *International Economic Review*, Vol. 34, No. 1, pp. 221-43.

Newey, W.K., J.L. Powell, and J.R. Walker (1990) 'Semiparametric Estimation of Selection Models: Some Empirical Results, *American Economic Review*, Papers and Proceedings, Vol. 80, No. 2, pp. 324-29.

Roy, A. (1951) 'Some Thoughts on the Distribution of Earnings', *Oxford Economic Papers*, Vol. 3, pp. 135-46.

Rosen, S. (1974) 'Hedonic Prices and Implicit Markets: Product Differentiation in Pure Consumption', *Journal of Political Economy*, Vol. 82, pp. 34-55.

— (1978) 'Substitution and Division of Labour', *Economica*, Vol. 45, pp. 235-50.

— (1983) 'A Note on the Aggregation of Skills and Labour Quality', *Journal of Human Resources*, Vol. 28, No. 3, pp. 425-31.

Rosenzweig, M. (1988) 'Labour Markets in Low-Income Countries', in H. Chenery and T.N. Srinivasan (eds), *Handbook of Development Economics*, Holland: Elsevier Science Publishers.

Schmertmann, C.P. (1994) 'Selectivity Bias Correction Methods in Polychotomous Sample Selection Models', *Journal of Econometrics* (forthcoming).

Smith, R.J., and R.W. Blundell (1986) 'An Exogeneity Test for a Simultaneous Equation Tobit with an Application to Labour Supply', *Econometrica*, Vol. 54, No. 3, pp. 679-85.

Tinbergen, J. (1951) 'Some Remarks on the Distribution of Labour Incomes', *International Economic Papers*, Vol. 1, pp. 195-207.