ELSEVIER
Labour Economics 2 (1995) 275–297
LABOUR ECONOMICS

Formal and informal sector employment in urban areas of Bolivia

Menno Pradhan, Arthur van Soest*

Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

Received December 1992; final version received December 1993

Abstract

Earnings and labour market participation in urban areas of Bolivia are analyzed, using household level survey data from 1989. We distinguish between non-participation, formal sector work, and informal sector work, and estimate separate wage equations for informal and formal sector. Two types of models are analyzed: in the first, the informal sector is seen as a buffer zone between formal sector and non-participation, while in the second, there is no ordering among sectors. We find that accounting for selectivity substantially affects wage equation estimates. The direction of the selectivity effect is the same according to both models, but its magnitude varies, in particular for the informal sector. Other results are quite robust: wages are higher in larger local labour markets. In both sectors females of ethnic minorities are underpaid.

Keywords: Employment; Segmented labor markets

JEL classification: J23; J42

1. Introduction

This paper analyzes earnings and labour market participation in urban areas in Bolivia, using household survey data from 1989. Three labour market states

*Corresponding author. This paper benefitted from comments by Joop Hartog, Francois Laisney, Thierry Magnac, Bertrand Melenberg, Ruud Picavet, anonymous CentER and Labour Economics referees and seminar participants at Tilburg university and the 1992 E.S.E.M. conference in Brussels. Financial support has been provided by the Netherlands Organization for Scientific Research (NWO) for the first author and the Royal Netherlands Academy of Arts and Sciences (KNAW) for the second author.

SSDI 0927-5371(94)00007-T

are distinguished: not working, working in the informal sector, and working in the formal sector. We analyze the determinants of potential earnings in both sectors, considering, for example, returns to education and effects of local labour market conditions. Moreover, we analyze participation. What factors determine whether an individual works in the formal or informal sector? Are the two sectors competitive and is the difference between potential earnings sufficient to explain the sector someone is in? Or do other factors matter, such as (non-monetary) sector preferences or barriers to entry?

The distinction formal/informal has become the standard way to characterize urban labour markets in developing countries. The formal sector is subject to regulation, wages are paid on a regular basis, taxes are levied and explicit contracts between employers and workers exist. The informal sector is not subject to government regulation and is dominated by one person firms and small enterprises that employ few apprentices or hired labourers. The exact definition is arbitrary to some extent and depends on the specific research objectives (Hart, 1985). The ILO mission to Kenya in 1972 (Lubell, 1990) defined economic informality in a broad context using characteristics such as ease of entry, small scale of operation, family ownership, skills acquired outside the formal school system, and unregulated and more competitive markets.

Models analyzing labour market segmentation and returns to human capital in a two-sector framework, have been used extensively. See, e.g., Hartog and Oosterbeek (1993) for a developed country and Van der Gaag and Vijverberg (1988) for a developing country. For three labour market states, however, no established approach exists. Two approaches can be distinguished. In the first the informal sector is viewed as an intermediary sector between not working and the formal sector. This stems from the traditional staging hypothesis in, for example, the pioneering work of Fields (1975): formal sector employment is rationed and all informal sector workers would be better of in a formal sector job. Recent empirical evidence, however, shows that many informal sector workers favour their current status to formal sector employment (Thomas, 1992). In the second approach, competitive markets are assumed and the two sectors are treated symmetrically. Examples are Magnac (1991), who tests for labour market segmentation in Columbia, and Gindling (1991), analyzing urban labour markets in Costa Rica.

We consider both types of models and pay attention to the sensitivity of the results for the choice of model. In particular, estimated wage equations may differ, because the models lead to different corrections for selection bias (Heckman and Hotz, 1986). We choose between models on the basis of specification tests.

The organization of the paper is as follows. In Section 2 we provide some background information regarding the Bolivian economy and institutional setting, and describe the data. In Section 3, we introduce the two models.

Results are discussed and evaluated in Section 4. Conclusions are mentioned in Section 5.

2. Economic background and data

To understand the role of the informal sector in Bolivia, some background information is useful. The Bolivian economy deteriorated rapidly from 1978 to 1985. In 1985, the annual inflation rates reached 12,000 percent (ILO, 1991) and GDP growth rates had been negative for five consecutive years. Open unemployment increased from 9.7 percent in 1981 to 18.0 percent in 1985 (ILO, 1991; persons aged 10 years and over). The government that took power in 1985 reacted to the situation with a 'New Economic Policy', to stabilize prices and reduce the role of the state. Policy measures included devaluation of the local currency, reduction of government expenditures, and a far reaching markets liberalization program. Prices and interest rates were freed and public sector wages were renegotiated. The policy of stabilizing prices was successful: from 1986 to 1989, consumer prices increased by only 16 percent per year. Economic growth however, failed to pick up. It reached 2.5 percent in the years 1986 to 1989, less than the growth in population. The unemployment rate hardly changed (19.0 percent in 1990).

During the economic slowdown between 1976 and 1987, the size of the informal sector in urban areas grew from 43 to 55 percent of the labour force (Velasco et al., 1989). There was a sharp rise just after the reforms: from 1985 to 1986, urban formal sector employment fell by 62,209 jobs, i.e. 14 percent. In the same year the informal sector grew with 116,704 people (UDAPE, 1991). This supports the view of the informal sector as a buffer sector. Income generated from the informal sector is needed in a period of economic recession.

Labour market regulation mainly affects the formal sector. In 1985, various labour regulations were abolished. Policy aimed at re-enforcing market powers in the labour market, establishing a closer link between work effort and pay. The policy change allowed for wage negotiations at firm level instead of sector level, eased hiring and firing of employees, and abolished most bonuses (such as subsidized food stores) for public sector workers. Minimum wage laws were maintained. In 1989, the minimum monthly wage was Bs 60. [1] It is mostly offered to unskilled public sector workers. Social security is compulsory for formal sector firms and workers. It covers health and disability insurance and old age pension.

---

[1] In Bolivianos; in 1989, 1 Boliviano was about 0.37 U.S. dollar.

2.1. Data

We use the second round of the Bolivian household survey (Encuesta Integrada de Hogares), drawn in 1989. This is a random sample of the urban population, administered yearly by the Bolivian National Bureau of Statistics (Instituto Nacional de Estadística), with assistance of the World Bank. The 1989 survey covers 7264 households in 8 urban centers. Household survey data are more appropriate than firm level data for measuring activity in the informal sector, since the latter often do not include information on non-listed firms or micro-enterprises, the bulk of the informal sector. We distinguish between formal and informal sector on the basis of worker status. See Appendix A for a discussion and for details on the construction of the data base, definitions of the variables, and sample statistics. Samples used for estimation consist of 6349 males and 7293 females from 19 to 65 years of age.

Hourly earnings are higher for males than for females. On average, informal earnings are higher than formal earnings. For females in particular, the variance of hourly earnings is much larger in the informal than in the formal sector. This corresponds to the notion that the informal sector is very heterogeneous (Fields, 1990). For both sexes, average hours worked in the informal sector exceed those in the formal sector.

Because of measurement problems, reported hourly earnings may underestimate the benefits of formal sector jobs compared to informal sector jobs. Fringe benefits are not included in reported formal sector wages. 57 percent of males and 72 percent of females working in the formal sector, reported to have received benefits other than the regular wage. The survey does not collect a monetary equivalent for these. 48 percent of males and 64 percent of females in the formal sector are enrolled in social security, the costs of which are partly born by the employer. Income taxes hardly play a role: in 1989, they represented 3.3 percent of total tax revenues (World Bank, 1989). The bulk of government revenues is collected trough value added taxes. Informal sector earnings may not always be measured net of costs. The questionnaire does not contain enough detail to correct for this. This type of problems may affect estimates of wage differentials between the two sectors, but will not affect the estimates of the (reduced form) sector allocation equations.

Another point of concern is the extent to which the data are representative. Obviously, rural areas are not represented, so that the data cannot be used to analyze migration from rural to urban areas and the choice between employment in the agricultural or the urban sectors, the issues in the early theoretical models [Fields (1975), for example]. Because rural areas are not represented, the survey data deviate substantially from official nationwide statistics in some respects. For example, activity rates (the fraction of people working or looking for work), amount to 84 percent for males and 51 percent of females in the survey. Official nationwide estimates are 93.4 and 26.5 in 1990.

3. Models

The main objective of the paper is to analyze factors driving participation, sector choice and earnings in the formal and informal sector. The models consist of two wage equations, one for each sector, and two reduced form equations explaining the selection mechanism. They differ with respect to the latter. The number of hours worked is not considered. Wages are hourly wage rates, obtained by dividing total earnings by number of hours worked.

In the first model, the choice between working in the formal sector, working in the informal sector, and non-participation, is modelled using ordered probit. In the second model, multinomial logit is used. The selection equations are in reduced form, in the sense that the wage rate is not included as an explanatory variable. Wage effects are thus indirectly reflected through, for example, age and education effects. The model is also reduced form in the sense that we consider someone's actual state only. Information on preferred labour market state or job search is not taken into account. We do not disentangle effects through preferences from those through rationing, costs of search, etc.

3.1. Ordered probit selection model

In the first model, the three labour market states are ordered: participation in the formal sector – participation in the informal sector – non-participation. An underlying latent variable can be interpreted as an indicator of formality. Non-participation includes being engaged in household production which is associated with the lowest level of formality.

We do not explicitly specify an underlying structural economic model leading to this ordering. The economic interpretation is the staging hypothesis of the Fields (1975) model: informal sector employment is inferior to formal sector employment; informal sector earnings exceed unemployment income. In equilibrium, a move from unemployment to the informal sector involves an increase in earnings at the cost of a reduced efficiency of search. In this view, economic formality is associated with increasing economic activity and a decreasing search effort [see also Todaro (1989, p. 268)]. In a dynamic setting, new labour market participants or immigrants from rural areas would first accept a job in the informal sector, and simultaneously look for formal sector work.

The formal representation of the model is:

$$Y = Z\delta_1 + \varepsilon_3 \qquad \begin{array}{lll} Y < \alpha_1 & \text{if} & \text{working in formal sector,} \\ \alpha_1 < Y < \alpha_2 & \text{if} & \text{working in informal sector,} \\ \alpha_2 < Y & \text{if} & \text{non-participant.} \end{array} \tag{1}$$

The subscript indicating the individual is suppressed. Y is the latent variable, the inverse of the 'degree of formality'. $\varepsilon_3 \sim N(0,1)$, independent of Z. Z is a vector of individual, family, and regional characteristics. Since the system is reduced form (wage rates are eliminated), Z contains all variables in the wage equations. In addition, Z contains taste shifters that do not result from potential earnings differences. By means of normalization, Z contains no constant term.

In the standard version of the model stated above, α_1 and α_2 are constant across the sample. Identification requires only one of them to be constant, however. Keeping the other α constant limits model flexibility: choice probabilities depend on Z through the one index $Z\delta_1$ only; the probability of informal sector employment would depend on Z only through the non-linearity of the distribution function of ε_3. This is a general drawback of the standard ordered probit model. To allow for more flexibility, we parametrize α_2:

$$\alpha_2(Z) = \alpha_1 + \exp(Z\delta_2). \tag{2}$$

According to (1) and (2), the probability of formal sector employment is determined by $Z\delta_1$ only, but the choice between informal employment and non-participation depends on both $Z\delta_1$ and $Z\delta_2$. Incorporating the exponential term in $\alpha_2(Z)$ guarantees $\alpha_2 > \alpha_1$.

3.2. Multinomial logit selection model

In the second model, no a priori ordering among the three states is assumed. The model can be interpreted in terms of utility maximization. Let Y_i be the indirect utility associated with participation in sector i. We assume:

$$Y_i = Z\delta_i + \eta_i \qquad \begin{array}{lll} \max\{Y_1, Y_2, Y_3\} = Y_1 & \text{if} & \text{working in formal sector,} \\ \max\{Y_1, Y_2, Y_3\} = Y_2 & \text{if} & \text{working in informal sector,} \\ \max\{Y_1, Y_2, Y_3\} = Y_3 & \text{if} & \text{non-participant.} \end{array} \tag{3}$$

Here $\eta_i \sim EV(I)$ (an extreme value type I distribution), and η_1, η_2, η_3 independent. Z is the same as in the ordered probit model. Alternative i is chosen if its utility exceeds that of all other alternatives. Of course, since actual and preferred state do not necessarily coincide, the interpretation of utility maximization should not be taken literally. If someone prefers but cannot find a formal sector job, Y_1 will be small. Thus Y_i reflects both rationing and preferences.

Normalization requires one δ_i to be constant (we choose $\delta_3 = 0$). Define

$$\eta_i^* = \max(Y_j) - \eta_i \quad (j = 1, 2, 3, j \neq i). \tag{4}$$

Domencich and McFadden (1975) show that the probability of being in state i equals

$$P_i = \text{prob}(\eta_i^* < Z\delta_i) = F_i(Z\delta_i) = \frac{\exp(Z\delta_i)}{\sum_{j=1,2,3}\exp(Z\delta_j)}. \tag{5}$$

Here F_i is the distribution function of η_i*, which depends on $Z\delta_j$, $j \neq i$.

3.3. *Wage equations*

The natural logarithm of the potential hourly wage rate w_i in sector i is modeled as

$$\ln(w_i) = X_i\beta_i + \varepsilon_i, \quad i = 1(\text{formal}), 2(\text{informal}). \tag{6}$$

X contains explanatory variables: personal characteristics (human capital variables), and variables describing the condition of the labour market by urban area. ε_i is a normally distributed error term.

3.4. *Error structure*

In the ordered probit case, the three error terms ε_1, ε_2 and ε_3 are assumed to be jointly normally distributed with zero mean and a full covariance matrix. The covariance of ε_1 and ε_2 is not identified, since we observe one wage at most.

In the multinomial logit case, we follow Lee (1982) [see also Maddala (1983, p. 273)]. Let

$$\varepsilon_{3i}^* = \Phi^{-1}F_i(\eta_i^*) = J_i(\eta_i^*). \tag{7}$$

Here Φ^{-1} is the inverse of the standard normal distribution function. (7) implies that $\varepsilon_{3i}^* \sim N(0,1)$. Alternative i is chosen if $\varepsilon_{3i}^* < J_i(Z\delta_i)$. We assume that, for $i = 1, 2$, $(\varepsilon_i, \varepsilon_{3i}^*)$, is bivariate normal with mean $(0,0)$ and covariance matrix Σ_i, with $\Sigma_i(2,2) = 1$. These assumptions make it convenient to estimate the multinomial logit model and wage equations with a two-step method or maximum likelihood.

3.5. *Identification*

Both models are generalizations of two-sector self-selection models, of which the model of Roy (1951) is the seminal example. Identification of this type of models is discussed in Heckman and Honoré (1990) and Heckman (1990). Results in the latter can straightforwardly be extended to our models:

First, the complete model is identified under the specific distributional assumptions on the errors, even without exclusion restrictions on the regressors in one or more of the equations. This changes if distributional assumptions are relaxed and replaced by, for instance, the semiparametric assumptions of independence of errors from regressors and zero error means or medians. For this case, Heckman (1990, p. 314) proves identification under the assumption that at least one of the regressors in the selection part is excluded from the wage equations. This condition is satisfied in our case: Z contains variables referring to family composition and other family income, not included in X. In practice, whether results depend on distributional assumptions, may thus depend on the explanatory power of the additional regressors in the selection equations.

Moreover, semiparametric estimation of the constant terms in the wage equations is harder than estimating slope coefficients. Semiparametric identification of the constant term in, say, the wage equation of sector 1, relies on the assumption that a nonzero population fraction has a probability close to one of being selected into sector 1. As a consequence, semiparametric estimates of the constant term would be inaccurate if few observations have selection probability close to one. In other words, parametric estimates of constant terms may be particularly sensitive for the choice of distributional assumptions.

3.6. Net dissavings

To allow for an income effect on labour supply, some measure for full income or income excluding earnings is included in Z. For consistency with a life cycle framework (Blundell and Walker, 1986), the measure should be corrected for savings: in a two-stage budgeting framework, the within-period allocation of leisure and consumption is conditional on full expenditures in the same period, or, equivalently, on net dissavings, defined as consumption expenditures minus earnings. We use net dissavings per capita, assuming that the household's net dissavings are shared equally among family members. In this theoretical framework, net dissavings may be treated as an exogenous variable. However, if unobservable factors affect net savings and labour supply, treating net savings as exogenous leads to biased results. We therefore have instrumented for net savings. Non-labour income is the main instrument. Details are described in Appendix B.

3.7. Estimation

If Z, including net dissavings, is exogenous, the models can be estimated by full information maximum likelihood, or with a consistent two-step estimator. The latter can be used to obtain starting values for maximizing the likelihood.

Differences between the two sets of estimates can be used to carry out a Hausman test for model misspecification. Details on both estimators are mentioned in a technical appendix, available upon request from the authors.

Allowing for endogeneity of net dissavings as described in Appendix B, does not substantially complicate the estimation. In an auxiliary first OLS step, an equation explaining net dissavings is estimated. OLS residuals are then added to the systematic parts of the selection equations. Non-zero coefficients of these residuals indicate endogeneity. See Appendix B. Standard errors of ML have been corrected for the uncertainty in the first step estimates using Newey (1984).

4. Estimation results

Tables with detailed estimation results of two-step and ML estimations, are in the technical appendix, available upon request. The following discussion is based on the, more efficient, ML results.[2]

4.1. Wage equations

In Table 1 we present estimates for the wage equations. The estimates of the age pattern and the constant term in the informal sector are sensitive to the choice of selection model. For most other explanatory variables the estimated coefficients in both models are comparable in sign and significance.

The coefficients on the age variables (*age*) are significant (at 5 percent level) only for females in the formal sector. A maximum is reached at about 46 years of age. Females in ethnic minorities are significantly underpaid in both sectors according to both models. For males, the signs are the same, but significance levels are lower.

The variables *econ act* (economic activity) and *unemployment* describe local labour markets conditions. The effect of *econ act*, which is used as a proxy for the size of the local labour market, is positive and significant in all cases. The local unemployment rate has a negative effect on earnings, and the effect is largest in the informal sector. This suggests that the informal sector is more competitive, and that less favourable labour market conditions have a greater negative impact on earnings in the informal than in the formal sector.

Table 1 also contains estimates of the error variances of the wage equations and of error covariances between wage and selection equations. The informal

[2] The terminology in the sequel ignores the preliminary step instrumenting for per capita net dissavings. The two step estimator thus is actually a three step estimator. Similarly, what we call maximum likelihood is not ML of the full model, but the two step estimator with OLS in the first step, and quasi generalized ML on the rest of the model in the second step.

Table 1

ML estimates wage equations[a]

	Ordered probit				Multinomial logit			
	Males		Females		Males		Females	
	Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal
cnst	0.042	0.221	− 1.777*	0.273	− 0.058	− 1.961*	− 1.887*	1.874*
	(0.174)	(0.278)	(0.288)	(0.331)	(0.168)	(0.302)	(0.278)	(0.387)
age	0.026*	− 0.005	0.081*	0.009	0.031*	0.095*	0.083*	− 0.024
	(0.008)	(0.012)	(0.012)	(0.014)	(0.008)	(0.013)	(0.012)	(0.015)
age square/100	− 0.007	0.030*	− 0.088*	0.000	− 0.013	− 0.094*	− 0.092*	0.034
	(0.010)	(0.015)	(0.017)	(0.017)	(0.010)	(0.015)	(0.016)	(0.018)
inter	0.082	0.052	0.220*	0.256*	0.082	0.077	0.227*	0.275*
	(0.049)	(0.061)	(0.088)	(0.071)	(0.048)	(0.065)	(0.088)	(0.076)
medio	0.245*	0.284*	0.523*	0.360*	0.248*	0.204*	0.540*	0.529*
	(0.041)	(0.054)	(0.075)	(0.064)	(0.040)	(0.058)	(0.073)	(0.069)
midtech	0.440*	0.377*	0.813*	0.566*	0.441*	0.215	0.853*	1.001*
	(0.078)	(0.136)	(0.104)	(0.121)	(0.076)	(0.137)	(0.100)	(0.131)
hightech	0.582*	0.371*	1.028*	0.813*	0.582*	0.143	1.079*	1.375*
	(0.100)	(0.169)	(0.141)	(0.245)	(0.099)	(0.167)	(0.137)	(0.251)
normal	0.078	− 0.419*	0.972*	0.125	0.106	− 0.610*	1.044*	1.208*
	(0.083)	(0.166)	(0.121)	(0.175)	(0.080)	(0.166)	(0.112)	(0.161)
university	0.827*	0.600*	1.311*	0.718*	0.854*	0.227*	1.371*	1.407*
	(0.048)	(0.080)	(0.110)	(0.117)	(0.046)	(0.091)	(0.102)	(0.127)
other	0.211*	− 0.404*	− 0.209*	− 0.042	0.211*	− 0.537*	− 0.210*	− 0.042
	(0.067)	(0.106)	(0.098)	(0.074)	(0.064)	(0.118)	(0.098)	(0.080)
yrs incompl educ	0.061*	0.046*	0.088*	− 0.002	0.062	0.017	0.094*	0.039
	(0.012)	(0.018)	(0.017)	(0.019)	(0.011)	(0.018)	(0.017)	(0.021)
econ active/10	0.172*	0.136*	0.176*	0.151*	0.169*	0.118*	0.177*	0.117*
	(0.023)	(0.034)	(0.032)	(0.035)	(0.022)	(0.036)	(0.032)	(0.038)
unemployment∗10	− 0.664*	− 0.889*	− 0.704*	− 0.854*	− 0.676*	− 1.218*	− 0.727*	− 0.904*
	(0.090)	(0.133)	(0.121)	(0.149)	(0.087)	(0.145)	(0.121)	(0.157)
ethnic	− 0.061	− 0.073	− 0.183*	− 0.145*	− 0.066*	− 0.028	− 0.194*	− 0.311*
	(0.034)	(0.045)	(0.055)	(0.048)	(0.033)	(0.048)	(0.054)	(0.053)
σ_i	0.905*	0.986*	0.684*	0.954*	0.878*	1.110*	0.689*	1.138*
	(0.012)	(0.026)	(0.011)	(0.017)	(0.011)	(0.033)	(0.013)	(0.039)
σ_{3i}	0.704*	0.653*	− 0.080	0.268*	0.655*	− 0.923*	− 0.138*	0.803*
	(0.023)	(0.040)	(0.076)	(0.062)	(0.022)	(0.048)	(0.068)	(0.070)

[a] See Table A.1 (Appendix A) for the definitions of explanatory variables.

sector wage variance is larger than that of the formal sector wage. This corresponds to the notion that the formal sector is more regulated, leading to a smaller earnings dispersion. The covariances determine the differences between the potential wage distribution for the whole population, and the actual wage distribution among those selected in the sector concerned. It is the latter which is reflected in the sample, and the former which we try to analyze. For males, the positive coefficients for the formal sector imply negative selectivity effects. This is not what one would expect. For males in the informal sector, the selectivity effects are positive according to both models.[3] This means that, for given observed characteristics, those with highest unobserved informal sector skills indeed work in the informal sector. For females, the multinomial logit model leads to exactly the opposite results: the selection effect is positive for the formal, and negative for the informal sector. The difference between the two models seems substantial here: according to the multinomial logit model, the selectivity effects are much stronger than according to the ordered probit model. The correlation coefficients for the informal sector are surprisingly high, explaining the huge selection effects, particularly in the multinomial logit model.

Education level is incorporated as follows. First, dummies are used to indicate the highest level of courses attended. Seven levels are distinguished. Second, for those who did not complete the course, we used the deviation between the level attained and the level if completed, expressed in years. This deviation is zero if the course is finished and negative otherwise. Those who followed training that was not classified, are included in the estimation. For them, the dummy variable *other* is set to one and *yrs incompl educ* (years of incomplete education) to zero.

For a discussion of the effect of education on the wage offers we turn to Table 2. For each level of (completed) education, we have calculated the predicted log wage. For the formal sector, the wage pattern as a function of education level is robust with respect to the chosen model. The relative increase in wage offer from lowest (*basic*) to highest (*univ*) (= university) type of education is 0.84 for males and 1.34 for females. For the informal sector the estimates vary substantially with the choice of the selection model. In particular, the estimated average wage levels are quite different according to the two models. This may be explained by the problems with semiparametric identification of the constant term (cf. the discussion in Section 3). Returns to education are lower in the informal sector than in the formal sector, as we would expect under the segmented labour market hypothesis.

Figure 1 shows the estimated population distribution of potential earnings in both sectors according to both models. There is substantial overlap in the distribution of earnings in informal and formal sector, caused by variation in

[3] Because the probability of non-participation is small for males, this statement is valid for the ordered probit model as well.

Table 2
Effect of education on predicted log wages (*age* = 35, *city* = La Paz, *ethnic* = 0)

Education level	Corresp. years of education	Males		Females	
		Ord probit	Mult logit	Ord probit	Mult logit
Formal					
basic	5	0.638	0.603	– 0.300	– 0.390
		(0.048)	(0.047)	(0.144)	(0.130)
inter	8	0.720	0.685	– 0.080	– 0.163
		(0.052)	(0.050)	(0.137)	(0.127)
medio	12	0.883	0.851	0.223	0.151
		(0.042)	(0.040)	(0.110)	(0.101)
midtech	13	1.078	1.044	0.513	0.464
		(0.078)	(0.076)	(0.100)	(0.095)
hightech	15	1.220	1.186	0.729	0.689
		(0.097)	(0.095)	(0.124)	(0.122)
normal	17	0.717	0.709	0.672	0.655
		(0.079)	(0.076)	(0.059)	(0.058)
univ	20	1.465	1.457	1.012	0.981
		(0.046)	(0.044)	(0.072)	(0.069)
Informal					
basic	5	– 0.142	– 0.677	0.099	0.828
		(0.067)	(0.083)	(0.069)	(0.099)
inter	8	– 0.090	– 0.600	0.356	1.104
		(0.073)	(0.091)	(0.081)	(0.107)
medio	12	0.142	– 0.473	0.460	1.357
		(0.060)	(0.084)	(0.070)	(0.111)
midtech	13	0.234	– 0.462	0.666	1.830
		(0.138)	(0.150)	(0.119)	(0.164)
hightech	15	0.229	– 0.534	0.912	2.204
		(0.170)	(0.178)	(0.242)	(0.271)
normal	17	– 0.561	– 1.287	0.224	2.036
		(0.178)	(0.190)	(0.166)	(0.203)
univ	20	0.458	– 0.450	0.817	2.236
		(0.092)	(0.124)	(0.119)	(0.178)

[a] Standard errors in parentheses.

observed and unobserved characteristics. For males, formal sector wages are on average higher than informal sector wages. Examining the expected wage offer (not accounting for the unobserved characteristics) for all individuals separately, one finds that for all males in the sample the expected wage in the formal sector exceeds that in the informal sector, according to both models. In this respect, the distribution in the whole population differs from the distributions in the selected

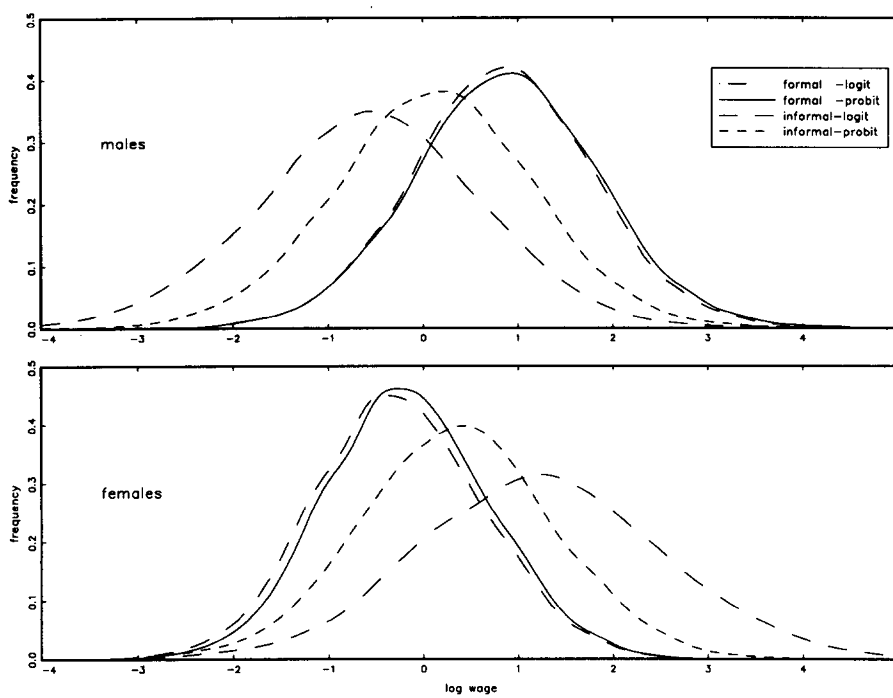

Fig. 1. Estimated distribution of wage offers.

Note: Wage offers include random draws of the error terms. Smoothing is done using a Gaussian kernel with the normal reference rule bandwidth.

subpopulations. In particular, the selection effect in the informal sector is quite strong: only those with relatively high potential informal sector earnings will indeed work in the informal sector. This is why the sample average of earnings of informal sector workers, exceeds that of formal sector earnings. This finding yields some support for the Fields (1975) model: on average, males would be better of in the formal sector.

For females however, we find opposite results. According to the multinomial logit model, expected earnings are higher in the informal sector for all females in the sample. Using ordered probit, we find that for 9 percent of the sample the predicted wage offer in the formal sector is higher. These are exclusively females with normal (mainly teachers) or university training. These results indicate that we cannot explain sector participation on the basis of expected wages and restrictions to entry in the formal sector only. Account must be taken of differences in preferences for the two sectors.

4.2. Selection equations

Interpreting separate parameter estimates in the selection equations is not very useful. Human capital variables appear in various ways (*age* and *age square*, for example), and selection probabilities are often determined by more than one linear combination of the regressors. In Table 3, we therefore present effects of marginal changes of some of the characteristics on the state probabilities, for the average male and female. Signs and significance levels of the effects according to the two models are similar in most cases.

The effect of family composition is captured trough *young*, *prime* and *old*. The presence of young children in the household (*young*) significantly decreases the probability that the female works in the formal sector, and increases her probability of non-participation. The presence of other prime age individuals (*prime*) implies that the male or female concerned may not be the main earner in the household. This significantly reduces the female's informal employment probability, and increases her probability of non-participation. The latter is also true for males, at the cost of the formal employment probability. A negative impact of *prime* on participation was expected. It is not clear why *prime* affects the choice between formal and informal employment. The presence of family members older than 65 (*old*) is significant in one case only: It increases the male's probability of informal employment.

Individual characteristics are *married*, *age* and *ethnic*. For females, being married strongly and significantly reduces employment probabilities in both sectors. For males, being married reduces the probability of non-participation and increases that of formal sector employment. The probability of informal employment increases significantly with age for both males and females. In return, the female's probability of not participating decreases with age, and so does the male's probability of formal employment. This may reflect a cohort effect rather than a pure age effect. Since we included *age* and *age square*, marginal effects at other age levels may be different. Note that age effects are a combination of direct and indirect effects: they include those through the wage rate. For ethnic minorities (*ethnic* = 1), the probability of informal employment is significantly larger than for others. For males, this must be a direct effect since *ethnic* hardly affects wages. The relative magnitude of the effect alters with the model chosen.

Education level enters through various variables, and it is not possible to present a single partial derivative. Effects of education on the participation probabilities according to the two models are virtually identical. In general, they are as expected: probabilities of formal and informal employment increase and decrease with education level, respectively. We find a strong positive impact of 'normal' training (*normal*) on the formal employment probability. This type of education includes teachers training college, typically used in the formal sector.

Table 3
Predicted partial derivatives of probability of participation for an average individual for ordered probit and multinomial logit selection model[a]

	Males			Females		
	Formal	Informal	Not working	Formal	Informal	Not working
Ordered probit						
young	– 0.005	0.004	0.001	– 0.009*	– 0.005	0.014*
	(0.004)	(0.004)	(0.003	(0.003)	(0.003)	(0.004)
prime	– 0.012*	– 0.004	0.016*	– 0.002	– 0.015*	0.017*
	(0.003)	(0.004)	(0.003	(0.003)	(0.003)	(0.004)
old	– 0.028	0.025	0.003	– 0.0075	0.02	0.007
	(0.015)	(0.016)	(0.013)	(0.012)	(0.013)	(0.017)
married	0.091*	0.019	– 0.110*	– 0.152*	– 0.115*	0.267*
	(0.017)	(0.018)	(0.013)	(0.010)	(0.012)	(0.015)
netdissav pc/1000	– 0.624*	– 0.189	0.814*	– 0.205*	– 0.250*	0.455*
	(0.176)	(0.197)	(0.135)	(0.097)	(0.107)	(0.141)
*age**100	– 0.478*	0.534*	– 0.057	– 0.046	0.547*	– 0.501*
	(0.070)	(0.073)	(0.057)	(0.051)	(0.054)	(0.068)
econ act/10	– 0.007	0.003	0.004	0.004	0.026*	– 0.029*
	(0.011)	(0.010)	(0.008)	(0.007)	(0.007)	(0.009)
*unemployment**10	– 0.005	– 0.060	0.065*	– 0.144*	– 0.007	0.151*
	(0.041)	(0.040)	(0.030)	(0.027)	(0.029)	(0.038)
ethnic	– 0.062*	0.049*	0.013	– 0.056*	0.062*	– 0.006
	(0.014)	(0.013)	(0.010)	(0.011)	(0.009)	(0.013)
Multinomial logit						
young	– 0.003	0.004	– 0.001	– 0.010*	– 0.004	0.014*
	(0.004)	(0.003)	(0.002)	(0.003)	(0.004)	(0.004)
prime	– 0.015*	– 0.003	0.018*	– 0.004	– 0.014*	0.018*
	(0.003)	(0.003)	(0.002)	(0.003)	(0.004)	(0.004)
old	– 0.027	0.032*	– 0.006	– 0.003	– 0.001	0.004
	(0.016)	(0.013)	(0.011)	(0.011)	(0.017)	(0.019)
married	0.103*	– 0.026	– 0.077*	– 0.154*	– 0.094*	0.248*
	(0.017)	(0.015)	(0.010)	(0.010)	(0.013)	(0.016)
netdissav pc/1000	– 0.703*	0.129	0.574*	– 0.207*	– 0.360*	0.566*
	(0.185)	(0.169)	(0.108)	(0.097)	(0.144)	(0.155)
*age**100	– 0.585*	0.630*	– 0.045	– 0.061	0.669*	– 0.608*
	(0.073)	(0.069)	(0.046)	(0.052)	(0.067)	(0.072)
econ active/10	– 0.002	0.001	0.002	0.007	0.027*	– 0.034*
	(0.011)	(0.010)	(0.007)	(0.007)	(0.009)	(0.010)
*unemployment**10	0.023	– 0.111*	0.088*	– 0.134*	0.000	0.134*
	(0.042)	(0.039)	(0.026)	(0.026)	(0.037)	(0.041)
ethnic	– 0.065*	0.058*	0.007	– 0.054*	0.098*	– 0.043*
	(0.014)	(0.013)	(0.009)	(0.011)	(0.012)	(0.014)

[a] Standard errors in parentheses; *: significance at 5% level.

Per capita net dissavings (*netdissav pc*) has, for both sexes, a significantly positive effect on the probability of non-participation. Interpreted in labour supply terms, this implies that leisure is a normal good. An increase of net per capita dissavings by 1 percent increases the probability of non-participation by 0.9 percent for males and 0.7 percent for females. For females, both formal and informal sector participation probabilities are negatively affected. For males, this only is the case for the formal sector.

Economic activity (*econ act*) and the unemployment rate (*unemployment*) characterise the individual's local labour market. A larger labour market significantly increases the probability that a female works in the informal sector, at the cost of non-participation. A high unemployment rate increases non-participation, for males and females. This may reflect a discouraged worker effect, or, to some extent, an indirect wage effect. For females in particular, results suggest that the discouraged worker effect in the formal sector is substantial.

4.3. Specification tests

The models are non-nested. In the (extended) ordered probit model and the multinomial logit model, the number of slope parameters is the same. Vuong (1989) developed a test for the null hypothesis that the two models are equally close, in the Kullback–Leibner sense, to the true data generating process against the alternative that one of the models is closer. This test is quite general, since it does not assume that either of the models represents the truth.[4] For males the null hypothesis cannot be rejected at a 5 percent level. The likelihood of the ordered probit model is larger than that of the multinomial logit model, but the difference is too small to reject equality of expected log likelihoods. For females, the null hypothesis is rejected in favour of the multinomial logit model. For both sexes, multinomial logit outperforms ordered probit in the percentage of correctly predicted observations: multinomial logit predicts the labour market state of 59.0 percent of females and of 62.4 percent of males correctly. For the ordered probit the percentages are 56.1 and 59.2, respectively. The special case of constant α_2 of the ordered probit model is clearly rejected by a likelihood ratio test.

Model specifications were separately tested using (generalized[5]) Hausman specification tests (Hausman, 1978), based upon comparison of (efficient) ML

[4]Performing the test is easy here, since we have strictly non-nested models (Vuong, 1989, p. 317). We ignored the preliminary step of estimating the dissavings equation. Since estimates of ML standard errors which account for this preliminary step, are virtually identical to standard ML standard errors ignoring the first step uncertainty, we do not think that this leads to biased results.

[5]The tests are generalized in the sense that we have no specific alternative in mind for which ML is inconsistent while the two step estimator remains consistent. This may affect the power, but does not affect the size.

estimates with (consistent but inefficient) two-step estimates. Computational details are available upon request.[6] All model specifications are rejected at the 1 percent level. Given the large number of observations and the previous experience in the literature with these kind of models, this is no surprise. For example, the model of Magnac (1991) fails to pass similar tests. The values of the test statistic lead to the same conclusions as the likelihood values: for males, the ordered probit version of the model performs better, for females the multinomial logit model.

5. Conclusions

We have analyzed earnings and labour market participation in urban areas of Bolivia using two different methods to model participation. The first, the ordered probit model, explicitly models the informal sector as a buffer sector, between non-participation and the formal sector. The second, the multinomial logit model, does not impose any ordering. We have compared the two models using estimation results, specification tests and comparative statics.

We have generalized the standard version of the ordered probit model, allowing one of the thresholds to vary with characteristics. The model thus gets the same flexibility as a multinomial model without ordering. This generalization appears to be an improvement: the special case of a constant threshold is strongly rejected for both sexes, using e.g. likelihood ratio or Wald tests.

The models are consistent with a life cycle framework in the treatment of income other than earnings. Moreover, we allow for endogeneity of the other income measure, net dissavings. This appears to be an improvement also: LM tests confirm that endogeneity is significant. Estimating the model without allowing for endogeneity yields some quite different results: the impact of unearned income changes sign, and the conclusions about selectivity in the wage equations change substantially. Allowing for endogeneity, we find that leisure is a normal good: the income effect on non-participation is clearly positive. For males, the income effect on formal sector participation is clearly negative, whereas it is insignificant for the informal sector. For females, both income effects of participation are significantly negative. Although the difference between the two models is in the way sector participation is modelled, the estimated probabilities correspond reasonably well, and depend on the regressors in a similar way.

---

[6] Again, we ignored the preliminary OLS estimation of the net dissavings equation. Given the negligible differences between uncorrected and corrected ML-standard errors, we do not think that this is important. Two step method standard errors for the wage equation are corrected for heteroscedasticity and imputing the estimated Mill's ratio following Newey (1984).

Contrary to Gindling (1991) we find that accounting for selectivity substantially affects the wage equation estimates. The direction of the selectivity effects is the same according to both models. In the informal sector, however, the magnitude of the selectivity effect substantially depends on the model which is chosen. As a consequence, the two models lead to substantially different predictions of potential informal sector earnings of nonparticipants and formal sector employees. Using specification tests for non-nested models, we tend to prefer the multinomial logit model for females. For males, we find slightly more support for the ordered probit model, but the difference is insignificant.

A number of conclusions appears to be robust with respect to the chosen model. The dispersion of potential earnings in the informal sector is much larger than in the formal sector (see Fig. 1). This is in accordance with the notion that the informal sector is strongly heterogeneous (Fields, 1990). Wages in both sectors are higher in larger local labour markets, and lower in areas with high unemployment. In both sectors also, females of ethnic minorities are generally underpaid, while for males this effect is insignificant. In accordance with economic theory, returns to education in the formal sector exceed those in the informal sector. Predicted wages are higher in the formal sector for males, which supports the view in theoretical models (Fields, 1975) that formal sector employment is preferred to informal sector employment, but that formal sector jobs are rationed. For females, however, the opposite holds. This could indicate that it is not sufficient to look at wages only to explain the participation decision of females. It may also indicate relatively more favourable labour market conditions for females in the informal sector. The result is not in line with Magnac (1991) who could not reject weakly competitive labour markets for married woman in Colombia. Further analysis which disentangles the effects through rationing and differences in preferences seems necessary here.

In general, it seems reasonable to conclude that the impact of most explanatory variables is robust with respect to the specification choice. This is not the case for the exact magnitude of selectivity effects, particularly in the informal sector. Looking at two models instead of just one might then be a first step towards a more robust view on the segmented labour market hypothesis.

Appendix A: Data source and definition of formality

The study is based on the 1989 Bolivia household survey. It contains a measure for household consumption and, for every family member, detailed information on labour supply, earnings, education, health, fertility and migration. The labour section is extensive, with information on occupation, earnings, hours worked and search behaviour.

Table A.1
Sample means by labour market state (sample standard deviations of other variables than dummies in parentheses)[a]

	Male			Female		
	Formal	Informal	Not working	Formal	Informal	Not working
Education level:						
basic	0.21	0.35	0.24	0.09	0.39	0.29
inter	0.14	0.19	0.13	0.08	0.16	0.15
medio	0.29	0.30	0.33	0.22	0.20	0.29
middle technical	0.04	0.03	0.04	0.09	0.03	0.05
higher technical	0.03	0.02	0.03	0.05	0.01	0.01
normal (teacher)	0.06	0.01	0.02	0.27	0.02	0.03
university	0.19	0.07	0.16	0.17	0.03	0.05
other	0.04	0.03	0.05	0.03	0.16	0.13
married (dummy)	0.79	0.85	0.59	0.55	0.71	0.80
ethnic	0.30	0.39	0.33	0.19	0.46	0.34
age (in years)	35.9	39.7	38.8	33.8	39.3	36.8
	(10.7)	(11.6)	(15.4)	(9.3)	(11.3)	(12.7)
per cap net	0.01	− 0.04	0.79	0.24	− 0.04	0.37
dissavings (* 10)	(2.20)	(1.88)	(1.71)	(2.79)	(1.65)	(1.99)
hourly earnings	2.39	2.58		1.94	2.01	
(primary activity)[b]	(4.4)	(4.5)		(1.8)	(4.6)	
hours worked per week	49.7	52.3		38.5	46.9	
(primary activity)	(16.8)	(19.2)		(16.5)	(24.7)	
Number of observations	3605	1863	881	1439	1972	3882

[a] *Explanation*:
Education levels: *basic*, *inter* and *medio* are subsequent courses of formal education. Vocational training is referred to as *technical* and includes industry, commercial and agriculture training. *normal* refers to training for teachers. *university* includes both public and private universities. For males, a large percentage of the *other* category is military training.
yrs incompl educ: if individual did not complete training: minus number of years before completion. Otherwise equal to zero.
ethnic: dummy variable obtained from the language question: if the respondent commonly speaks another language than Spanish, the variable is set to one.
per cap(ita) net dissavings: family expenditures minus family earnings divided by family size.
econ act: number of individuals (males and females) working or looking for work in the same urban area in the (random) sample. It relates to the size of the local labour market and may pick up scale effects. It varies between 695 for Potosi and 2468 for St. Cruz.
unemployment: the local unemployment rate (males and females jointly). It varies between 0.053 in Trinidad and 0.107 in Oruro.
[b] The survey collects limited information on the secondary activity. Secondary activity hours and earnings could not be computed in the same way as for the primary job. The number of people with primary and secondary activity in different sectors is small: 2.5 percent of males and 1 percent of females. We therefore ignore the secondary activity in this study.

We use subsamples of 6349 out of 7937 males, and 7293 out of 9028 females. From the original sample, 855 males and 864 females were excluded because they were not working because of health problems or because they attended full-time education. 455 males and 648 females were excluded because, according to our definition of formal and informal sector, they could not be classified. These are home or family workers and employers. 19 males and 9 females were excluded because of missing or implausible information on earnings or hours worked. We excluded 10 males and 5 females because the education level was not reported, 9 males and 8 females because reported household expenditures were extremely different from reported household income, 226 working males and 175 working females because they did not report earnings, and 14 males and 19 females because of missing information on other household income.

The data allow for different definitions of the formal and informal sector. We have considered two definitions. The first uses the size of the enterprise as the primary indicator of formality: if this is less than 6, the work is classified as informal, if it is at least 6, the job is formal. However, we follow the common approach of classifying independent professionals, such as lawyers and doctors, in the formal sector. Household workers and family workers are left unclassified. The second definition is based on the worker's status and corresponds to Magnac (1991): wage workers and independent professionals are classified as formal and self-employed[7] workers as informal. Others (employers, home and family workers) are left unclassified. The two definitions lead to the same classification for 80 percent of all workers (if we include non-classified). Of all workers, 4 percent are employers, and are therefore not classified according to Magnac's definition. These workers are deleted from the sample.

Details on definitions of explanatory variables and statistics of the sample used for estimation, are presented in Table A.1.

Appendix B: Endogeneity of net savings

In the standard life cycle model, net dissavings and the error terms in the labour supply equations (due to future uncertainty only) are uncorrelated. Net savings may, however, be endogenous as a result of unobserved heterogeneity. We add a reduced form equation for net dissavings to account for this:

$$netdissav = Q\xi + u, \quad u \sim \mathrm{N}(0,\sigma^2) \tag{B.1}$$

[7] According to the survey an individual is engaged in self-employment if: the person worked without dependence on a boss, managing his own economic unit, with or without the help of family workers or unpaid apprentices, but without using more than two salaried workers.

Table B.1
OLS results net dissavings equation[a]

	estimate	std err
intercept	– 284.94	136.76
other income	0.26	0.02
La Paz	– 19.76	48.32
Cochabamba	2.07	49.09
Oruro	– 130.06	53.17
Potosi	– 81.44	61.24
Tarija	53.55	59.26
St Cruz	– 8.42	49.31
Trinidad, Cobija	– 82.44	52.61
young	– 1.16	8.64
prime female	– 17.15	11.33
prime male	– 38.38	12.33
old	– 15.55	38.38
age head of household	8.49	5.58
age square	– 0.03	0.06
*married * age*	95.80	33.12
ethnic	26.22	28.44
inter	71.98	36.46
medio	134.85	32.78
midtech	227.42	71.56
hightech	289.16	81.12
normal	102.69	57.63
university	220.36	41.16
other	– 115.89	46.69
yrs incompl educ	21.91	9.56

[a] Dependent variable: Household net dissavings, per capita. R^2: 0.03

Here *netdissav* is per capita net dissavings, and Q is a vector of household specific variables.[8] Endogeneity arises if u is correlated with the errors in the selection model.

In the ordered probit model, we assume that u and the error ε in the selection equation are bivariate normal. This implies that the conditional expectation of ε given u, is linear in u. In the multinomial logit model, we make a similar assumption, concerning the distribution of the error terms η_i in (3), conditional on u:

$$\bar{\eta}_i = \eta_i - \gamma_i u \sim \mathrm{E}\,V(I);\quad \bar{\eta}_1, \bar{\eta}_2, \bar{\eta}_3 \text{ independent.} \tag{B.2}$$

[8] Q_j includes household non-labour income. Since this is not in the vector Z of explanatory variables in the selection equations, identification of the model is guaranteed.

Normalization requires $\gamma_3 = 0$. If $\gamma = (\gamma_1, \gamma_2) = (0,0)$, the specification is as described in Section 3, and net savings are exogenous.

Both models can be consistently estimated if endogeneity is allowed for by adding an auxiliary estimation step (cf. Smith and Blundell (1986) for the ordered probit case). First, (B.1) is estimated by OLS. OLS residuals are then added to the regressors of the selection equation(s). Next, ML or two-step estimation is carried out. Testing for endogeneity boils down to testing the significance of the coefficient(s) of the OLS residuals. ML-standard errors in the second step are inconsistent in case of endogeneity, because the parameters in (B.1) are replaced by their estimates. Correction is possible following Newey (1984).

Preliminary OLS-estimates of the dissavings equation (B.1) were used for both models. Results are mentioned in Table B.1. As expected, the impact of other household income is positive and strongly significant.

References

Blundell, R., 1990, Evaluating structural microeconometric models of labour supply, in: J. Laffont and C. Sims, eds., Advances in econometrics, Sixth World Congress (Cambridge University Press, Cambridge).

Blundell, R. and I. Walker, 1986, A life cycle consistent empirical model of labour supply using cross section data, Review of Economic Studies 53, 539–558.

Domencich, T. and D. McFadden, 1975, Urban travel demand: A behavioral analysis (North-Holland, Amsterdam).

Fields, G.S., 1975, Rural urban migration, urban unemployment and underemployment, and job-search activity in LDCs, Journal of Development Economics 2, no 2, 165–187.

Fields, G.S., 1990, Labour market modelling and the urban informal sector: Theory and evidence, in: D. Turnham, B. Salomé and A. Schwartz, eds, The informal sector revisited (OECD Development Centre, Paris) 49–69.

Gindling, T.H., 1991, Labor market segmentation and the determination of wages in the public, private-formal, and informal sectors in San José, Costa Rica (Economic Development and Cultural Change 39, no 3, 585–606).

Hart, K., 1985, The informal economy, Cambridge Anthropology 10, no 2, 54–58.

Hartog, J. and H. Oosterbeek, 1993, Public and private sector wages in the Netherlands, European Economic Review 37, no. 1, 97–114.

Hausman, J.A., 1978, Specification tests in econometrics, Econometrica 46, 1251–1272.

Heckman, J., 1990, Varieties of selection bias, American Economic Review 80, no 2, 313–318.

Heckman, J. and B.E. Honoré, 1990, The empirical content of the Roy model, Econometrica 58, no 5, 1121–1149.

Heckman, J. and V.J. Hotz, 1986, The sources of inequality for males in Panama's labor market, The Journal of Human Resources 21, no 4, 507–542.

International Labour Office, 1991, Yearbook of labour statistics, Vol. 50 (ILO, Geneva).

Lee, L.F., 1982, Some approaches to the correction of selectivity bias, Review of Economic Studies 49, 355–372.

Lubell, H., 1990, The informal sector in the 1980s and 1990s (Development Centre, O.E.C.D., Paris).

Maddala, G.S., 1983, Limited dependent and qualitative variables in econometrics (Cambridge University Press, Cambridge).

Magnac, Th., 1991, Segmented or competitive labor markets, Econometrica 59, 165–187.
Newey, K.N., 1984, A method of moment interpretation of sequential estimators, Economics Letters 14, 201–206.
Roy, A., 1951, Some thoughts on the distribution of earnings, Oxford Economic Papers 3, 135–146.
Smith, R.J. and R. Blundell, 1986, An exogeneity test for a simultaneous equation Tobit model with an application to labor supply, Econometrica 54, 679–686.
Thomas, J.J., 1992, Informal economic activity (Harvester Wheatsheaf, Hertfordshire).
Todaro, M.P., 1989, Economic development in the Third World (Longman, New York).
Unidad de Analisis de Politicas Economicas, 1991, Estadísticas economicas de Bolivia (UAPE, La Paz).
van der Gaag, J. and W. Vijverberg, 1988, A switching regression model for wage determinants in the public and private sectors of a developing country, The Review of Economics and Statistics 70, no. 2, 244–252.
Velasco, A.P., R.C. Sainz, S.E. Pabon, and H.L. Cordova, eds., 1989, Informalidad e ilegalidad: Una falsa indentidad (CEDLA, La Paz).
Vuong, Q.H., 1989, Likelihood ratio tests for model selection and non-nested hypotheses, Econometrica 57, 307–333.
World Bank, 1989, Bolivia: Updating economic memorandum, Report nr 7645-BO (World Bank, Washington, DC).