Formal and informal sector employment in urban areas of Bolivia

Pradhan, M.P.; van Soest, A.H.O.

Publication date:
1993

Citation for published version (APA):
Pradhan, M. P., & van Soest, A. H. O. (1993). *Formal and informal sector employment in urban areas of Bolivia*. (CentER Discussion Paper; Vol. 1993-11). Unknown Publisher.

CentER for Economic Research

Discussion paper

[redacted]

Tilburg University

No. 9311

FORMAL AND INFORMAL SECTOR EMPLOYMENT IN URBAN AREAS OF BOLIVIA

by Menno Pradhan
and Arthur van Soest

February 1993

ISSN 0924-7815

December 1992

Formal and Informal Sector Employment in Urban Areas of Bolivia

Menno Pradhan

and

Arthur van Soest

This paper benefitted from comments by Joop Hartog, Francois Laisney, Thierry Magnac, Bertrand Melenberg, Ruud Picavet, an anonymous CentER referee 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.

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

Abstract

Earnings and labour market participation in urban areas of Bolivia are analyzed, using household level survey data, drawn in 1989. We distinguish between non-participation, formal sector work, and informal sector work, and estimate separate wage equations for the informal and formal sector. We account for endogeneity of household savings, and this significantly improves the results. Two types of models are analyzed: In the first, the informal sector is seen as a buffer zone between formal sector and non-participation (generalizing ordered probit), and in the second (multinominal logit), no ordering among sectors is imposed. We find that accounting for selectivity substantially affects the estimates of the wage equations. The direction of the selectivity effect is the same according to both types of models, but its magnitude varies, in particular for the informal sector. Some other results are quite robust: Wages are higher in larger local labour markets, and in both sectors females of ethnic minorities are generally underpaid, while for males the difference is insignificant.

1.Introduction

Urban labour markets in developing countries can be characterized by the coexistence of a regulated, formal sector and an unregulated, informal sector. The latter sector is dominated by one person firms and small enterprises that employ few apprentices or hired labourers. The slowdown of economic growth in developing countries in the 1980's and the resulting increase in the fraction of the labour force engaged in informal sector activities has stimulated economic research in this field. Most analytical studies have analyzed the relationship between the formal and informal sector within the context of labour market segmentation and dual labour markets (cf, e.g., Magnac, 1991).

This paper analyzes earnings and labour market participation for the urban areas in Bolivia, using household survey data from 1989. Three possible labour market states are distinguished: Not working, working in the informal sector, and working in the formal sector. The exact distinction between informal and formal sector work is arbitrary to some extent (cf Hart, 1985). Definitions emphasizing the economic distinction between the two sectors try to capture the unobserved variable "degree of formality" through measurable variables such as "worker's status" (i.e. wage workers versus self-employed) or "number of persons employed by the firm". The choice of definition depends on the objectives of the research and the nature of the data available. The ILO mission to Kenya in 1972 (cf 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.

In section 2, we briefly describe some characteristics of the Bolivian economy and the survey data used in the analysis. We compare two definitions of formal and informal sector jobs which can be used with the data at hand. In the rest of this study, we use the definition of Magnac (1991) based on worker's status.

In section 3, we introduce and estimate two models that are used to analyze earnings in the formal and informal sector. Three labour market states are distinguished: working in the formal sector, working in the informal sector and not working. We thus do not distinguish between, for example, those looking for a job and those not looking for a job. We use the term non-participants for all non-workers, including the involuntary unemployed. The models differ in the way they treat the sector participation decision.

In the first model working in the informal sector is viewed as intermediary status between not working and working in the formal sector. The nature of certain formal sector jobs may be such that they require a certain level of education and skills. Moreover, the regulation of the formal sector poses higher barriers to entry. Given that most formal work involves an explicit contract, employers may be hesitant to employ individuals who do not have an established education or employment record. These individuals may be better off to seek work in the informal sector, where earnings are generally more directly related to their labour inputs. Moreover, the informal sector can be seen as a buffer for workers laid off in the formal sector, who would otherwise be unemployed. In this case, the informal sector functions as a way to smooth income and consumption during a period of economic slowdown. In the second model we do not impose any prior ordering among sectors. The three sectors are modelled in a symmetric fashion. Gindling (1991) uses this specification to analyze the labour market in San Jose, Costa Rica.

Hourly earnings in both sectors are analyzed by means of joint models of sector assignment and wage determination. Heckman and Hotz (1986) point out that standard, ordinary least squares based, tests for differences in expected wages are infeasible, because of endogenous selection into sectors. In the case of two sectors two step methods can be used to correct for selectivity bias. See e.g. Hartog and Oosterbeek (1989) (for a developed country) and Van der Gaag and Vijverberg (1988) (for a developing country) for public-private sector wage differentials. In our models, three labour market states are distinguished, and correcting for selectivity bias can be achieved either by generalizing the two step procedure or by applying maximum likelihood. We shall consider both estimation procedures. The two sets of parameter estimates can be compared to test for misspecification. Results are discussed and evaluated in section 4. Some concluding remarks and suggestions for further research are mentioned in section 5.

2. Economic background and data

The country of focus is Bolivia. The Bolivian economy deteriorated rapidly from 1978 to 1985. By 1985, inflation had reached a level of 12,000 percent at annual rate (see ILO, 1991) and growth rates had been negative over the past five years. Unemployment increased from 9.7% in 1981 to 18.0% in 1985 (cf ILO, 1991, persons aged 10 years and over). Foreign lending, that took place in the sixties and seventies, had left the country with a huge debt burden. The government that took power in 1985 reacted to the situation with a "New Economic Policy" which was meant to stabilize prices and reduce the role of the state. The package of policy measures included a devaluation of the local currency as well as a reduction of government expenditures. The government also instituted a far reaching liberalization program of markets. Prices and interest rates were freed and public sector wages were renegotiated. Restrictions on the hiring and firing of employees were eased. The policy was successful in that it stabilized prices: over the period 1986 to 1989 consumer prices increased by an average of only 16 percent per year. However, economic growth failed to pick up, it reached a level of 2.5 percent over the period from 1986 to 1989, still less than the growth in population. The unemployment rate hardly changed (19.0% in 1990). During the period of economic slowdown between 1976 and 1987, the size of the informal sector in urban areas is estimated to have grown from 43 to 55 percent of the labour force (see Velasco et al., 1989).

The research will be based on data of the second round of the Bolivian household survey (Enquesta Intergrada de Hogares), drawn in 1989. The survey uses a random sample of the urban population and is administered yearly by the Bolivian National Bureau of Statistics (Instituto National de Estadistica) with technical assistance of the World Bank. The 1989 survey covers 7264 households in 8 urban centers. Household survey data, in contrast to firm level data, are particularly appropriate for measuring activity in the informal sector since they are drawn from the entire urban population. Firm level data often do not include non-listed firms (micro-enterprises), of which the bulk of the informal sector consists. The survey collects a measure for household consumption and, for every family member separately, detailed information on labour supply, earnings, education, health, fertility and migration. The labour section of the survey is extensive. It provides information on occupation, earnings, hours worked and search behaviour.

The information collected in the survey allows for different definitions of the formal and informal sector. Table 1 presents a cross tabulation using two different definitions, for the working population aged 19 to 65. The first definition uses the size of the enterprise as the primary indicator of formality. If the size of the enterprise is less than 6, the work is classified as informal. However, as is common, independent professionals, such as lawyers and doctors, are classified with the formal sector workers. Household workers and family workers are left unclassified. The second definition is based on the worker's status and corresponds to the definition used by Magnac (1991): Wage workers and independent professionals are classified as formal and self-employed[1] workers as informal. Others, that is employers, home and family workers, are left unclassified. The change of definition does not affect the sector of classification for 80 percent of the workers (if we include non-classified). Of all workers, 4 percent are employers, and are therefore not classified according to Magnac's definition.

The research focuses on labour supply behaviour of individuals from 19 to 65 years of age. Table 2 provides summary statistics for the individuals that were included in the estimation. The table uses the definition of formality based on the worker's status. Throughout the paper this definition will be maintained. Formal workers have a higher average education level than informal workers. This difference is the greatest among females. Striking is the high percentage of individuals reporting to

[1]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.

have completed university training.

Table 1. Comparison of two definitions of formality
(percentage of working individuals)

Definition according to size Definition according to worker's status	formal	informal	not classified	total
formal	34	16	0	50
informal	0	39	0	39
not classified	2	2	7	11
total	36	57	7	100

The variables "econ act" and "unemployment" characterize the local labour market, and vary across the (eight) urban areas from which the survey is drawn. The "econ act" variable denotes the number of people (males and females) working or looking for work in the (random) sample. It relates to the size of the local labour market and may pick up scale effects. It varies is between 695 for Potosi to 2486 for St. Cruz. The second one is the local unemployment rate (males and females jointly). The lowest unemployment rate is found in Trinidad (0.053) and the highest in Oruro (0.107). The variable "ethnic" is a dummy variable obtained from the language question: If the respondent commonly speaks another language than Spanish, the variable is set to one. Ethnic workers are overrepresented in the informal sector. The variable "Net dissavings" is defined as family expenditures minus family earnings. The reason for using this specific other family income measure will be discussed below.

Hourly earnings are higher for males than for females. For both sexes, average informal earnings are higher than average formal earnings. For females, the standard deviation of the informal sector earnings is more than twice that of the formal. For males also, the variance of formal sector hourly earnings is smaller than for informal sector earnings, but the difference is not as large as one might expect. Both for males and females, the average number of hours worked per week is higher in the informal sector than in the formal.

Problems in measurement exist for both formal and informal earnings. For formal earnings, non-wage income is not included into the reported wage. Of the males working in the formal sector, 57 percent reported to have received other benefits besides the regular wage during the past year. For females the percentage is 72. The survey does not collect a monetary equivalent for these benefits. Also, 48 percent of the males in the formal sector and 64 percent of the females report to be enroled in some sort of social security. The costs of this insurance may be partly paid by the employer. None of the informal workers is enroled in a social security program. For informal earnings, there may be a problem in measuring these net of costs. There is only one question included in the questionnaire to measure earnings of the self-employed. To get a more precise measure of their earnings, more questions would be preferable. Also the treatment of taxes in the survey is not clear. From the information that we have available we cannot make firm statements on the effect of these problems.

To get an idea to what extent the data represent national figures, the information contained in the survey can be compared with official estimates from other sources (ILO, 1991). In the survey, 78 percent of all males and 47 percent of all females have a paid job. The activity rates, i.e. the fraction of people working or looking for work, amounts to 84 percent of males and 51 percent of females is either working or searching a job. Official estimates report activity rates of 93.4 and 26.5 for males and females in the same age group, nationwide in 1990 (i.e. for both rural and urban areas). Official estimates report an average of 44.8 hours of work per week for employees (i.e. a subset of all formal sector workers). The official estimates differ substantially from the ones obtained from the (urban areas only) EIH data. It is not clear to us why the quality of the official estimates would be "better" than of the ones obtained from the EIH survey. At the time of the survey there was no other reliable employment survey in operation. The official estimates are probably also based on the EIH survey, and the difference could be caused by the use of different definitions or by the guestimate that is included in the official estimate for the rural areas.

Table 2. Descriptive statistics
(formal and informal sector definition based on worker's status, means and sample fractions; standard deviation in parentheses)

	male formal	inform	not working	female formal	inform	not-working
highest level of education[2] attended:						
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	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	35.9	39.7	38.6	33.7	39.3	36.8
net dissavings[3]	8.24	-31.6	322.8	98.98	-43.5	141.5
	(959)	(1009)	(634)	(1229)	(865)	(1065)
hourly earnings[4]	2.39	2.58		1.94	2.01	
(primary activity[5])	(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

The estimations are based on subsamples of 6349 out of 7937 males, and 7293 out

[2]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, probably, a large percentage of the "other" category is military training.

[3]household level variable

[4]in Bolivianos; At the time of the survey 1 Boliviano was worth approximately 0.37 U.S. dollar. Taxes and premiums are relatively low in Bolivia. The survey questionnaire is not clear on information on whether the figures are gross or net of taxes.

[5]The survey does collect information on the secondary activity. However, the information for the second activity is not collected with the same detail as for the first. For example, the question on working hours does not specify whether average or actual hours are requested. Also, we do not observe a great number of individuals participating in both sectors. In the dataset 2.5 % percent of the males and 1 % of the females participated both in the formal and informal sector. In this study the information on any second activity is ignored.

of 9028 females. From the original sample, 855 males and 864 females were excluded because these individuals are not considered part of the labour force. They were, at the time of the survey, not working because of health problems or because they attended full-time education. The decision to go to school is thus assumed to be taken prior to the labour supply decision. 455 males and 648 females were excluded because according to the definition of formal and informal sector that is used in this study, they could not be classified. These individuals are home or family workers and employers. 19 males and 9 females were excluded because the information on earnings or hours was either missing or rather unrealistic. Individuals reporting to have worked more than 112 hours a week were excluded. We excluded 10 males and 5 females because the education level was not reported. 9 males and 8 females were excluded from the sample because reported total household expenditures were very different from reported total household income. Earnings were not reported for 226 males and 175 working females. Finally 14 males and 19 females were excluded because the error in the net dissavings equation could not be predicted.

3. Models

The models used in the analysis simultaneously deal with the questions of sector selection and hourly earnings in both sectors. The systems consist of two reduced form equations explaining the selection mechanism and two wage equations. The selection equations differ for the two models. The number of hours worked is not considered, wages are obtained by dividing total earnings by the number of hours worked. The model is short run, in the sense that the decisions to augment human capital are assumed to be taken prior to the labour supply decisions. The main objective of the model is to get better insight into the determinants driving participation, sector choice and earnings in both sectors.

The participation part of the model allows for three possible labour market states: working in the formal sector, working in the informal sector, and non-participation. Note that the selection equation explains someone's primary activity only. For the 2.5% of males and 1% of all females who work in both sectors, their secondary activity is ignored. In the first model selection is modelled using an ordered probit specification while in the second a multinominal logit specification is used. The models are equivalent in the number and choice of explanatory variables. The selection part of the model is reduced form, in the sense that the wage rate is not included as an explanatory variable. Wage effects are thus indirectly reflected by explanatory variables such as age and education level. The model is also reduced form in the sense that we consider someone's actual state only. Information on preferred labour market state and job search activities is not taken into account. We thus do not disentangle effects through preferences from those through rationing, costs of search, etc.

Ordered probit selection model

In the first model sector choice and participation are modelled using an ordered probit specification. The model assumes there is some ordering among the three labour market statuses. The ordering we assume is: participation in the formal sector - participation in the informal sector - non-participation. The interpretation is that there is a latent variable which indicates the formality of employment. A high level of this variable implies participation in the formal sector, a low value indicates that the lowest level of formality: non- participation. Non-participation includes being engaged in household production which is also associated with a low level of formality. An intermediate value implies participation in the informal sector.

We do not explicitly specify an underlying structural model leading to this ordering, but several interpretations could give rise to it. For example, it can be assumed that the formal sector always leads to a higher utility (based on potential wage and individual preference) than the informal sector and that the entry restrictions decrease monotonically over the three sectors. In the latter case, the informal sector can be viewed as a buffer zone between formal employment and non-participation. The empirical results must show whether this interpretation is reasonable. The issue is also raised by, for example, Todaro (1989, p. 268).

The formal representation of the model is:

$$Y = Z\delta_1 + \epsilon_3 \qquad \begin{array}{ll} Y < \alpha_1 & \textit{if working in formal sector} \\ \alpha_1 < Y < \alpha_2 & \textit{if working in informal sector} \\ \alpha_2 < Y & \textit{if non-participant} \end{array} \qquad (1)$$

Here Y is a latent variable expressing the inverse of the "degree of formality". ϵ_3 is distributed normal with mean zero and unit variance and is independent of Z. The subindex indicating the individual is suppressed. Not participating is associated with the lowest degree of formality and working in the formal sector with the highest. Z is a vector of individual, household and regional characteristics. By means of normalization, it does not contain a constant term.

In the standard version of the model stated above, α_1 and α_2 are assumed constant across the sample, i.e. do not depend on Z. Normalization requires only one of the α's to be constant. Keeping the other α constant unnecessarily restricts the way in which the probability of participation is allowed to depend on the explanatory variables. It implies that the choice between the three sectors is determined by the single index $Z\delta_1$. The probability to participate in the informal sector would then depend on exogenous variables only by virtue of the non-linearity of the distribution function of ϵ_3. This is a general drawback of the ordered probit model in its standard form. In order to allow for more flexibility, we therefore have adopted the following specification of α_2.

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

According to (1) and (2), the probability of formal employment is still determined by $Z\delta_1$ only. The choice between informal employment and non-participation however, depends on both $Z\delta_1$ and $Z\delta_2$. Incorporating the exponential term in $\alpha_2(Z)$ is for convenience only, to guarantee that $\alpha_2 > \alpha_1$, but is not necessary for identification. A minor drawback is that when α_2 is modelled as a constant and α_1 is α_2 minus an exponential term the specification of the model changes. In other words, α_1 and α_2 are do not enter symmetric.

<u>Multinominal logit selection model</u>

In the second model, sector choice and participation are modelled using a multinominal logit specification. No a priori ordering among sectors 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.

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

with $\eta_i \sim$ Extreme value type I and η_1,η_2,η_3 independent. The subscript i for Z is suppressed as we do not include sector characteristics among the explanatory variables. Alternative i is chosen if the utility associated with that alternative exceeds the utility of all other alternatives. Of course, since actual state 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, the value of Y_1 will be small. The Y_i's thus also reflect rationing and not just preferences. Normalization requires one of the δ_i to be constant (we choose $\delta_3=0$). Define

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

As is shown by Domencich and Mc Fadden (1975) the probability of participation in sector i equals

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

Where F is the distribution function of η_i^*. The vector Z contains a set of explanatory variables that could affect the sector choice. Since the system is reduced form, i.e. both wage rates are eliminated from the selection equations, all variables included in the wage equation are also included in Z. In addition, Z contains taste shifters capturing preferences for a sector that do not result from differences in potential wage rates.

Wage equations

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

$$\ln(w_i) = X_i\beta_i + \epsilon_i \qquad i=1(formal),\ 2(informal) \qquad (6)$$

with X containing a set of explanatory variables including personal characteristics (human capital variables), as well as variables describing the condition of the labour market by urban area. ϵ_i is an error term.

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. Without loss of generality, the variance of ϵ_3 is set equal to 1. The covariance between ϵ_1 and ϵ_2 cannot be identified, since we observe a wage in at most one sector.

In the multinominal logit case, we follow the approach suggested by Lee (1982) (see also Maddala, 1983, p. 273). Consider the transformation

$$\epsilon_{3i}^* = \Phi^{-1}F_i(\eta_i^*) = J_i(\eta_i^*) \qquad (7)$$

with Φ^{-1} the inverse normal distribution function. Since J is strictly increasing, the distribution of ϵ_{3i}^* is standard normal. Alternative i is chosen if $\epsilon_{3i}^* < J_i(Z\delta_i)$. Now assume that, for i=1,2, $(\epsilon_i,\epsilon_{3i}^*)$, has a bivariate normal distribution with mean (0,0) and covariance matrix Σ_i, with $\Sigma_i(2,2)=1$. These assumptions make it convenient to estimate the multinominal logit model and wage equations simultaneously with a two step method or with maximum likelihood, as will be seen below.

In both models, the fact that Z contains variables which are not included in the wage equation guarantees that identification of the model is not just based on arbitrary assumptions about linearity or non-linearity of the distribution of the errors, which do not have an economic interpretation. The exact lists of regressors contained in X and Z can be read from the tables with estimation results (tables A2 and A3 in the appendix).

Endogeneity of net savings

To allow for an income effect on labour supply, some measure for full income or income excluding earnings should be included in the vector Z. In order to be consistent with a life cycle framework (cf, e.g., Blundell and Walker, 1986), the measure should be corrected for savings. In a two stage budgeting framework, the agent sequentially decides on the allocation of total expenditures on consumption and leisure over time, and the within period allocation of leisure and consumption. The within period allocation, which is the subject of this paper, is thus conditional on full expenditures, or, equivalently, on net dissavings, defined as household consumption expenditures minus total household wage income.

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, very well be endogenous to the labour supply decision as a result of unobserved heterogeneity. We add a reduced form equation for net dissavings to account for this potential endogeneity:

$$netdissav_j = Q_j\xi + u_j \qquad u_j \sim N(0,\sigma^2) \tag{8}$$

with

$netdissav_j$ = net dissavings of household j

Q_j = vector of household specific variables influencing net dissavings.

Endogeneity is present if u_j is correlated with the error terms in the selection model. Q includes a measure of non-labour income which is not included in Z. This guarantees identification of these correlations.

In the ordered probit model, it seems natural to assume that u and the errors in the selection part of the model are jointly normally distributed. This implies that the conditional expectation of the latter three given u, is a linear function of u. In the multinominal logit model, we make a similar assumption by choosing the simultaneous distribution as follows:

$$Y_i = Z\delta_i + \gamma_i\mu + \eta_i^* \qquad \begin{array}{ll} \max\{Y_1,Y_2,Y_3\}=Y_1 & \textit{if working in formal sector} \\ \max\{Y_1,Y_2,Y_3\}=Y_2 & \textit{if working in informal sector} \\ \max\{Y_1,Y_2,Y_3\}=Y_3 & \textit{if non-participant} \end{array} \tag{9}$$

with $\eta_i^* = \eta_i - \gamma_i u \sim$ EV(I) and $\eta_1^*, \eta_2^*, \eta_3^*$ independent, conditional on u.

Normalization requires $\gamma_3=0$. In the case where $\gamma=(\gamma_1,\gamma_2)$ equals 0 the specification does not change. If $\gamma \neq 0$ the marginal distribution of (η_1,η_2,η_3) is no longer GEV and the model stated in (3) is misspecified. Again the exogeneity test boils down to testing the significance of γ.

Estimation

We first discuss the case that Z is independent of all errors in the model, and then explain how to adjust the estimators if this is not the case, i.e. if net dissavings are not exogenous.

If Z is exogenous, the models can be estimated by full information maximum likelihood, or by using a consistent two step estimator. The latter is computationally the least demanding. The first step consists of estimating the selection submodels (ordered probit or multinominal logit) by maximum likelihood. This yields consistent estimates of δ_1, δ_2 and (in the probit case) α_1. Using these estimates, the wage equations can be estimated by ordinary least squares including correction terms to account for endogenous selection. The wage equations for the formal and informal sector with correction terms are stated below. For the ordered probit specification they are (cf. Idson and Feaster, 1990)

$$\ln(w_1) = X_1\beta_1 - \sigma_{13}\frac{\phi(\hat{\alpha}_1 - Z\hat{\delta}_1)}{\Phi(\hat{\alpha}_1 - Z\hat{\delta}_1)} + \epsilon_1^* \qquad \textit{(formal sector)} \tag{10}$$

and

$$\ln(w_2) = X_2\beta_2 + \sigma_{23}\frac{\phi(\hat{\alpha}_1 - Z\hat{\delta}_1) - \phi(\hat{\alpha}_2 - Z\hat{\delta}_1)}{\Phi(\hat{\alpha}_2 - Z\hat{\delta}_1) - \Phi(\hat{\alpha}_1 - Z\hat{\delta}_1)} + \epsilon_2^* \qquad \textit{(informal sector)} \tag{11}$$

where $\hat{\alpha}_2 = \exp(Z\hat{\delta}_2)$.

For the multinominal logit specification we get (cf. Lee, 1982):

$$\ln(w_i) = X_i\beta_i - \sigma_{i3}\frac{\phi(J_i(Z\hat{\delta}_i))}{F_i(Z\hat{\delta}_i)} + \epsilon_i^* \qquad \begin{array}{l} i=1 \text{ if working in formal sector} \\ i=2 \text{ if working in informal sector}\end{array} \tag{12}$$

Here ϕ and Φ denote the standard normal density and distribution function, respectively. The errors ϵ_1^* and ϵ_2^* now have zero conditional mean, given that the corresponding wage rate is observed. The equations can be estimated by OLS, using only those observations for which the relevant wage rate is observed. Including these correction terms is a straightforward extension of Heckman's (1974) two stage method. Ordinary least squares yields consistent estimates but, unless $\sigma_{i3}=0$, standard errors computed in the usual way are estimated inconsistently, because of the heteroscedasticity of ϵ^* and because the parameters in the correction terms are replaced by their estimates.

The likelihood function for the full information maximum likelihood estimator is given below. For workers the contribution to the likelihood consists of the marginal density of the wage times the conditional probability of participation in the individual's actual sector. For non-workers, the likelihood contribution equals the marginal probability of non-participation. All marginal and conditional probabilities can easily be computed, only requiring univariate normal distribution functions. In case of the multinominal logit model, this is the motivation for using the transformation given in (7).

Log-likelihood ordered probit model:

$$\sum_{formal}\ln\left\{f_1(\epsilon_1)\int_{-\infty}^{\alpha_1} f_{31}(\epsilon_3|\epsilon_1)d\epsilon_3\right\} + \sum_{informal}\ln\left\{f_2(\epsilon_2)\int_{\alpha_1}^{\alpha_2(Z)} f_{32}(\epsilon_3|\epsilon_2)d\epsilon_3\right\} + \sum_{\substack{non\\participant}}\ln\left\{\int_{\alpha_2(Z)}^{\infty} f_3(\epsilon_3)d\epsilon_3\right\} \tag{13}$$

Log-likelihood multinominal logit model:

$$\sum_{formal}\ln\left\{f_1(\epsilon_1)\int_{-\infty}^{J_1(Z\delta_1)} f_{31}(\epsilon_3|\epsilon_1)\,d\epsilon_3\right\} + \sum_{informal}\ln\left\{f_2(\epsilon_2)\int_{-\infty}^{J_2(Z\delta_2)} f_{32}(\epsilon_3|\epsilon_2)\,d\epsilon_3\right\} + \sum_{\substack{non\\participant}}\ln(P_3) \tag{14}$$

Here f_1 and f_2 are the marginal density functions for the wage equations. f_{31} and f_{32} are conditional (normal) density functions of ϵ_3, given ϵ_1 and ϵ_2, respectively. J_1, J_2

and P_3 are defined in (5) and (7).

Estimation in case of endogenous net savings
The more general model allowing for correlation between u in (8) and the errors in the selection equations, can be estimated as above, adding one preliminary step: First, the auxiliary equation (8) is estimated with OLS. Then the model itself is estimated conditional on the residual in (8). For the ordered probit selection model, the approach is the same as in Smith and Blundell (1986). The conditional distribution of ϵ_3 is again normal, and its mean is a linear function of u. Inclusion of an unbiased estimator of u, the OLS-residual $\hat{u}$, in Z corrects for the non-zero conditional expectation of ϵ_3. Significance of the estimated coefficient of $\hat{u}$ implies rejection of exogeneity of net savings.[6]

A similar procedure can be followed for the multinominal logit case, based on the linearity assumption implied by the specification given in (9). The OLS-residual from (8) is now added to the regressors in both selection equations in (3), and thus replace the unobserved error terms in (9). A joint test for the significance of the two corresponding coefficients indicates whether endogeneity plays a role. In both models, the ML-standard errors in the second step are inconsistent in case of endogeneity, because the parameters in (8) are replaced by their estimates. Correction is possible following the procedure given by Newey (1984).

4. Estimation Results

Preliminary OLS-estimates of the dissavings equation (8) were used for both models. Results are mentioned in table A1 in the appendix. As expected, the impact of other household income is positive and strongly significant.

Having obtained the residuals of (8), we then computed for both models first the two step estimates. These were used as starting values for maximum likelihood (ML).[7] The four sets of estimates are presented in the appendix (tables A2 and A3). Standard errors of the ML estimates have been corrected for the error due to replacing the parameters of (8) in the correction term by their estimates, following

---

[6] Using the terminology of Gourieroux and Monfort (1989), this estimation procedure can be interpreted as a quasi-generalized ML-estimator, since the unknown parameters in (8) have been replaced by their estimates.

[7] The terminology we use in the sequel ignores the preliminary step. The two step estimator thus is actually a three step estimator, where OLS of (8) is the first step. Similarly, what we call maximum likelihood is not ML of the full model, but the two step estimator with OLS on (8) in the first step, and quasi generalized ML on the rest of the model in the second step.

the procedures of Newey (1984). The difference between corrected and uncorrected standard errors appeared to be quite small. At first sight, the parameter estimates of the ML method and the two step method do not look very different. A formal specification test based upon the differences will be discussed below. Our discussion of the estimation results will be based upon the, more efficient, maximum likelihood estimates.

In the (extended) ordered probit model and the multinominal logit model, the number of slope parameters is the same. The models are nonnested. A comparison of likelihood values suggests that for males, the ordered probit model does better, whereas for females, the multinominal logit model outperforms ordered probit[8]. For both sexes however, multinominal logit outperforms ordered probit in the percentage of correctly predicted observations: multinominal logit predicts the labour market state of 59.0% of females and of 62.4% 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 was clearly rejected using a likelihood ratio test.

Interpreting the individual parameter estimates in the selection part of the model is not very useful. First, human capital variables appear in the vector of explanatory variables in different places. We have included both age and age squared, and a number of education level variables. Second, selection probabilities are often determined by more than one linear combination of the regressors. To overcome this problem, we have computed the effects of marginal changes of some of the characteristics on the state probabilities, for the average male and female. The computations are based on the estimates in tables A1 and A2; standard errors are computed using the delta method. The results for the ordered probit model and the multinominal logit model are presented in table 3. The signs and significance levels of the effects according to the two models are similar in most cases, for both males and females.

====================================
TABLE 3 INSERTED APPROXIMATELY HERE
====================================

The presence of young children in the household (YOUNG) significantly decreases the probability that the female works in the formal market and significantly (at 5% level) increases her probability of non-participation. For males, YOUNG is not

---

[8]This is not based on a formal test, since the models are non-nested. However,since the densities of the wages are defined with respect to the same measure, a comparison of the likelihood values is still feasible. A large likelihood then indicates that the model is closer to the empirical distribution in the data, in the Kullbach-Leibler sense. (c.f., e.g., Gourieroux and Monfort (1989))

significant. The presence of other prime age individuals, reflected in PRIME, implies the possibility that the male or female concerned is not the main earner in the household. This presence significantly reduces the female's informal employment probability, and increases her probability of non-participation. The latter is also the case for males, but then at the cost of the formal employment probability. The negative impact of PRIME on the participation decision is to be expected, but it is not clear why this variable would affect the choice between formal and informal employment. The presence of household members older than 65 (OLD) is significant in one case only.

For females, being married strongly reduces both the formal and the informal employment probability. The ceteris paribus effects are about -0.15 (with standard error 0.01) and -0.10 (st. error 0.01), respectively. For males, being married significantly reduces the probability of non-participation and increases the probability of participation in the formal sector. The effects are less pronounced than for females.

Per capita net dissavings has, for both sexes, a significantly positive effect on the probability of non-participation. If we interpret the model as a pure labour supply model, in which the participation decision just reflects preferences and constraints do not exist, this implies that leisure is a normal good. An increase of net per capita dissavings by 1% (i.e. 13 bolivianos) increases the probability of non-participation by 0.9% for males and 0.7% for females. For females, both the formal as the informal sector participation probabilities are negatively affected by an increase in net dissavings. For males, this only is the case for the formal sector.

The probability of informal employment increases significantly with age. This may reflect a cohort effect rather than a pure age effect. Surprisingly, the female's probability of not participating decreases significantly with age, and so does the male's probability of formal employment. Note however that all these effects are marginal effects at the average age level (37 for males and 35 for females). Due to the presence of the quadratic age terms, marginal effects at other age levels may be quite different. Note also that the age effects are a combination of direct and indirect effects: Because the wage rate is not explicitly included in the selection equations, the effects include those through the wage rate.

"Economic Activity" (EA) and the unemployment rate are regional variables that characterise the individual's local labour market. EA is used as a proxy for the size of the local labour market. A larger labour market increases significantly the probabilities that females work in the informal sector at the cost of non-participation. For males there is no significant effect. A high unemployment rate (given as a fraction *10) significantly increases the non-participation probability, for males as well as females. This may reflect a discouraged worker effect, but may also reflect an indirect effect through the wage: As we shall see below (table 4), the unemployment rate has a negative impact on formal as well as informal wages. For females, the impact on the formal employment probability is significant and much stronger than that on informal employment, while the effect on the informal wage is larger than on the formal wage. This suggests that the discouraged worker effect in the formal sector is substantial. For males, we find similar effects of unemployment on formal and informal wages, but now the impact on the formal employment probability is insignificant, whereas the effect on informal participation is significantly negative for the multinominal logit model. This does not support the combination of the discouraged worker hypothesis and the idea that, if unemployment is high, formal jobs will be harder to find than informal jobs.

For males and females belonging to 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 their wages. For females, ETHNIC has a significantly negative effect on wages in both sectors. The relative magnitude of the effect alters with the model chosen. The unambiguously positive effect of ETHNIC on the probability of informal work for both models suggests that also for females the effect is mainly direct.

In order to understand the effects of education, we have drawn figures 1 through 6. The 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 finish the course, we use 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 individuals who followed a training that was not classified are included in the estimation. For them, the dummy variable OTHER is set to one and "years of incompl educ" to zero. In the graphs OTHER is not included since we have no indication to which level it corresponds. In the figures, the levels are linked to the average numbers of years needed to complete them.

The first two figures refer to the participation probabilities. The patterns according to the two models are virtually identical. In general, they are as we expected: the probability of formal employment increases with education level, and the probability of informal employment decreases. The strong positive impact of 'normal' training on the formal employment probability is remarkable. This type of education includes teachers training college which is typically used in the formal sector.

Table 4 presents the ML estimates for the hourly earnings equations. The estimates appear to be sensitive to the choice of the sector selection model, particularly for the age pattern and the constant term in the informal sector. For most other explanatory variables the estimated coefficients are comparable in sign and significance. For females, the coefficients on the age variables are significant only for the formal sector. The top lies at about 46 years of age. "Economic activity" is positive and significant in all cases. The 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 figures 3 and 4, we have drawn the expected wage rate in the two sectors, conditional on participation in that sector as a function of education. We have thus included the expectation correction terms which also appear in the two step estimation method. Again, the two models yield the same picture. This is no surprise, since the conditional pattern of wage rates more or less directly reflects the observed pattern in the data and is not sensitive with respect to the model which is chosen. OLS (i.e. not correcting for selectivity bias) would basically yield the same results.

This however is no longer the case for the unconditional pattern, i.e. the estimated systematic part of the wage equation, not taking into account the errors. These are depicted in figures 5 and 6. For the formal sector the structural wage patterns are robust with respect to the model that is chosen. For females, the small selection effect causes the structural wages to be very similar to the conditional wages. For males, the structural wages are higher than the conditional. This is opposite from what one would expect. For the informal sector, conditional wages of males exceed structural wages. For females however, the opposite is the case. The difference between the two models seems substantial here: According to the multinominal logit model, the selectivity effects are much stronger than according to the ordered probit model.

The difference between conditional and unconditional distribution is largely determined by the estimates of the correlation coefficients. The ordered probit model for males yields estimates of the correlation coefficients of 0.7, indicating negative interdependence between the choice for the formal sector and both wage rates. Thus in the formal sector, the conditional expectation given participation will be smaller than the unconditional one. For the informal sector, the selection works in the opposite way, which is what we one would expect. Note that this does not yield a complete view of the selection mechanism, since we have conditioned on the observed explanatory variables, and only consider the male with average characteristics. For females, the ordered probit model yields much smaller estimates for the correlation coefficients, and as a consequence the effects of selection on the wage distribution are small.

The multinominal logit model yields results which for the formal sector are similar to those of ordered probit: The correlation coefficients are of similar magnitude and can be interpreted in the same way. The correlation coefficients for the informal sector are surprisingly high, explaining the huge selection effects. For males, the estimate is -0.9, indicating that only those with quite high informal sector wage will be in the informal sector. For females, the estimate of 0.8 implies quite the opposite. The differences seem less realistic than those of the ordered probit model.

=====================================

TABLE 4 INSERTED APPROXIMATELY HERE

=====================================

Thus the conclusions from the figures correspond to the estimates of the covariance structure of the error terms. The fact that the two models yield rather different results suggests that the conclusions are sensitive to the details of the chosen specification. The reason for this may be the quality of regressors appearing in the selection equations, but excluded from the wage equations. These regressors guarantee non-parametric identification of the model.

Going back to figures 5 and 6, we find that the return to education is higher in the formal sector than in the informal sector, as we would expect under the segmented labour market hypothesis. For males, expected earnings in the formal sector exceed those in the informal sector. This was not the case for the sample averages, nor for the conditional distributions in figures 3 and 4, and can thus be ascribed to the selection process: Only those with relatively high potential informal sector earnings will indeed work in the informal sector. For females however, we find the opposite result. One of the reasons might be the measurement of informal sector earnings: These may exclude costs of operating in the informal market, or might include the contribution of assisting children or other household members. Also a possible difference in tax treatment for the two types of earnings could cause a measurement error. Finally note that the ML-estimates of the variances of the wage equations imply that the informal sector wage variance is always 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 effects of the other, non-education, variables on the expected log wages in the two sectors are presented in table 5. First, we have calculated the expected wage for a standard individual. We have chosen this individual to be 35 years of age, living in La Paz and Spanish speaking (i.e. not belonging to an ethnic group). The next rows in the table present deviations from this standard individual. In the formal sector the two models yield similar results with respect to the relative changes in the wage. A young male, chosen to be 20 years of age, earns about 30 percent less that the standard male. For females the effect is stronger, the loss in earnings is about 38 percent. And older male, who is 50 years old, earns about 35

percent more than an the standard male. For females the effect is only around 8 percent. This suggest that for females the earnings profile in the formal sector is relatively steep during the first years and becomes flatter when the individual becomes older. Living in St Cruz has a positive effect on the wage rate. This is due to the lower unemployment rate (20 percent less) and the larger labour market prevailing in this city. For both males and females the increase in earnings is about 19 percent. Belonging to an ethnic group has a much stronger effect for females than for males. The loss in earnings is about 6 percent for males and about 17 percent for females.
For the informal sector the relative effects differ rather strongly with the choice of the selection model. Some general conclusions can still be drawn: According to both models the earnings profile for females is less steep for females than for males. The increase in earnings as a result of living in St Cruz is larger for the informal sector than for the formal sector. This indicates that this sector is more sensitive to local labour market conditions. The effect of belonging to an ethnic group is similar to the formal sector.

Both specifications were tested formally using (generalized) Hausman specification tests (cf. Hausman, 1978), based upon comparison of the (efficient) maximum likelihood estimates with the (consistent but inefficient) two step estimates. Computational details are given in the appendix.[9] The tests are generalized in the sense that we do not have a specific alternative in mind for which ML is inconsistent while the two step estimator remains consistent. This could imply that the tests have little power, but will obviously not affect the size. The results are presented in table 6. The model specification is rejected for all models. Given the large number of observations and the previous experience in the literature with these kind of models this is no big surprise. For example, the model of Magnac (1991) fails to pass similar specification tests. The ranking of the values of the test statistic leads to the same conclusions as the likelihood values: For males, the ordered probit version of the model performs better, for females the multinominal logit model.

---

[9] Due to the preliminary estimation step, our ML-procedure is not fully efficient. Given the negligible difference between uncorrected ML-standard errors and standard errors corrected for estimating the parameters in (8), we decided to ignore this and to treat the OLS-residuals from (8) as observed error terms. We did correct the two step method standard errors for the wage equation parameters for heteroscedasticity and imputing an estimated regressor (following Newey, 1984).

Table 6. Hausman tests

	Males	Females
Ordered probit	363	285
Multinominal logit	465	121

Degrees of freedom: 70
Critical value at 5 percent level: 90

5. Conclusions and future research

Earnings and labour market participation in urban areas of Bolivia were analyzed using two different methods to model participation. The first, the ordered probit model, explicitly models the informal sector as a buffer sector, in between non-participation and the formal sector. The second, the multinominal logit model, does not impose any a priori ordering of sectors. The comparison of the two models is based on the 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 becomes a double index model (two linear combinations of regressors play a role), and has similar flexibility as a multinominal model without a priori ordering. Formal tests show that this generalization is indeed an improvement: The special case of a constant threshold is strongly rejected for both sexes, using e.g. likelihood ratio or Wald tests.

Our 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 that we use, i.e. net dissavings. This appears to be an improvement also: Formal tests confirm that endogeneity is significantly present, and 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 equation change substantially. Allowing for endogeneity, we find that leisure is a normal good, in the sense that 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.

We find that accounting for selectivity substantially affects the wage equation estimates. The direction in which conditional and unconditional wage distributions

differ is the same according to both models. In the informal sector however, the magnitude of the difference substantially depends on the model which is chosen.
The two models are non-nested and it is difficult to make an unambiguous choice. Comparing likelihoods and considering specification tests, we tend to prefer the ordered probit for males and the multinominal logit model for females. The estimated returns to education in the wage equations however are comparable across models. In accordance with economic theory, returns to education in the formal sector exceed those in the informal sector. According to both models and for both sexes and sectors, we find that wages 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 seems to be insignificant.

In general, it then seems reasonable to conclude that the impact of most explanatory variables is robust with respect to the specification choice, but not so 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 1. The derivation of the Hausman test statistic

The Hausman test is based on the statistic $(\beta_{ml}-\beta_{ts})'\hat{V}\,(\beta_{ml}-\beta_{ts})$. With β_{ml} the maximum likelihood estimator of β and β_{ts} the two step estimator. The calculation of $\hat{V}\,(\beta_{ml}-\beta_{ts})$ is not trivial. Although $\hat{V}\,(\beta_{ml})-\hat{V}(\beta_{ts})$ is a consistent estimator of the asymptotic variance covariance matrix of $\sqrt{T}(\beta_{ml}-\beta_{ts})$ it is not guaranteed that in practice the result will be positive definite. We therefore use a different method that guarantees positive definiteness (see Newey 1984). Define β_i,f_i to be the parameter vector (Kx1) and the first moment matrix (KxN) of the ith step respectively. Let $\beta=(\beta_1,\beta_2)$ and $f=(f_1,f_2)$. It can be shown that

$$\sqrt{T}(\beta_{ml}-\beta_{ts})=\sqrt{T}(\beta_{ml}-\beta)-\sqrt{T}(\beta_{ts}-\beta)\xrightarrow{d}N(0,F)$$

with $\hat{F}=[\hat{H}_{ml}^{-1}\hat{G}_{ml}\hat{H}_{ml}^{-1'}+\hat{H}_{ts}^{-1}\hat{G}_{ts}\hat{H}_{ts}^{-1'}-\hat{H}_{ml}^{-1}\hat{G}_{ml/ts}\hat{H}_{ts}^{-1'}-\hat{H}_{ts}^{-1}\hat{G}_{ts/ml}\hat{H}_{ml}^{-1'}]$

and

$\hat{H}_{ml}$ $=E[\partial L/\partial\beta\partial\beta']$,the hessian matrix of the maximum likelihood estimator

$\hat{G}_{ml}$ $=E[(\partial L/\partial\beta)(\partial L/\partial\beta)']$,the cross product of the first derivatives of the ML estimator

$\hat{H}_{ts}$ =a block matrix $\begin{bmatrix} A & 0 \\ C & D \end{bmatrix}$ with

$A=E[\partial f_1/\partial\beta_1]$, the hessian of the first step with respect to β_1

$B=E[\partial f_2/\partial\beta_1]$, the derivative of the moment conditions in the second step with respect to β_1

$C=E[\partial f_2/\partial\beta_2]$, the hessian of the second step with respect to β_2

$\hat{G}_{ts}$ $=E[f\,f']$ the cross product of the moment conditions in the two step method.

$\hat{G}_{ml/ts}$ $=E[(\partial L/\partial\beta)f']$

$\hat{G}_{ts/ml}$ $=E[f(\partial L/\partial\beta)']$

References

BLUNDELL, R. (1990), "Evaluating Structural Microeconometric Models of Labour Supply", Working Paper, University College London and Institute for Fiscal Studies.

BLUDELL, R. AND I. WALKER (1986), "A Life Cycle Consistent Empirical Model of Labour Supply using Cross Section Data", Review of Economic Studies,53,539-58.

DOMENCICH, T., AND D. MC FADDEN (1975), Urban Travel Demand: A behaviorial Analysis, Amsterdam: North Holland.

GAAG, J. VAN DER 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.

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, vol 39, no 3 pp 585-606.

GOURIEROUX, C. AND A. MONFORT (1989), Statistiques et Modeles Econometriques, Paris: Economica.

HART, K. (1985), "The Informal Economy", Cambridge Antropology, vol 10, no 2 ,pp54-58.

HARTOG, J. AND H. OOSTERBEEK (1989), "Public and Private Sector Wages in the Netherlands or: The Vital Role of Unobserved Heterogeneity", Research Memorandum no. 8926, University of Amsterdam.

HAUSMAN, J.A. (1978), "Specification Tests in Econometrics", Econometrica, vol. 46, pp 1251-1272.

HECKMAN, J. (1974), "Shadow Prices, Market Wages and Labor Supply", Econometrica, vol. 47, pp 153-161.

HECKMAN, J. AND V.J. HOTZ (1986), "The Sources of Inequality for Males in Panama's Labor Market", The Journal of Human Recources, vol. 21, no 4, pp 507-542.

IDSON, T.L. AND D.J. FEASTER (1990), "A Slectivity Model of Employer Size Wage Differentials", Journal of Labor Economics, vol. 8, no 1, pp 99-122.

INTERNATIONAL LABOUR OFFICE. (1991), Yearbook of Labour Statistics, 50 ed., Geneva.

LEE, L.F. (1982), "Some Approaches to the Correction of Selectivity Bias" Review of Economic Studies, vol. 49, pp 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, 1 ed., Cambridge University Press, Cambridge.

MAGNAC, TH. (1991), "Segmented or Competitive Labor Markets", Econometrica, vol. 59, pp 165-187.

MEDINA, S.D. (1986), La Economia Informal en Bolivia, 1 ed., EDOBOL, La

Paz.

NEWEY, K.N. (1984), "A Method of Moment Interpretation of Sequential Estimators", Economic Letters, vol. 14, pp 201-206.

NEWMAN J.L. AND P.J. GERTLER (1992), "Family productivity, Labor Supply and Welfare in a low Income Country", working paper, World Bank (PHRWH), Washington D.C..

SMITH, R.J. AND R. BLUNDELL (1986), "An exogeneity test for a simultaneous equation Tobit model with an application to labor supply", Econometrica, vol. 54, pp 679-686.

TODARO, M.P. (1989), Economic Development in the Third World, New York: Longman.

VELASCO, A.P., R.C. SAINZ, S.E. PABON, Y H.L. CORDOVA (1989), Informalidad e Ilegalidad: Una Falsa Indentidad, ed., CEDLA, La Paz.

Table A1. Results first step net dissavings regression

Dependent variable: Net dissavings
R^2 : 0.03

	estimate	std err
INTERCEP	-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 FEM	-17.15	11.33
PRIME MALE	-38.38	12.33
OLD	-15.55	38.38
AGE HEAD OF HH	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
OTRO	-115.89	46.69
YRS INCOMPL EDUC	21.91	9.56

Table A2. Ordered probit estimates (ML std errors corrected, two step std errors not corrected)

MALES:	ML		two step		FEMALES: ML		two step	
	coeff	std. error	coeff	std. error	coeff	std. error	coeff	std. error
α_1	-0.159	0.218	-0.079	0.229	-1.961	0.281	-2.008	0.278
δ_2:cnst	-1.919	0.322	-1.843	0.318	-1.177	0.262	-1.137	0.261
young	0.007	0.013	0.000	0.013	0.007	0.011	0.007	0.011
prime	-0.050	0.015	-0.051	0.015	-0.039	0.012	-0.044	0.012
old	0.059	0.065	0.093	0.059	0.004	0.050	0.000	0.050
married	0.996	0.186	1.000	0.193	0.023	0.138	-0.014	0.137
net dsav pc /1000	-2.389	0.726	-2.179	0.718	-0.279	0.410	-0.311	0.416
ehat-nets	-0.100	0.727	-0.431	0.719	-0.137	0.410	-0.084	0.416
mar*age	-0.018	0.004	-0.017	0.005	0.000	0.003	0.000	0.003
age	0.093	0.014	0.095	0.013	0.033	0.011	0.033	0.011
age square /100	-0.085	0.016	-0.089	0.016	-0.020	0.012	-0.020	0.012
inter	0.084	0.063	0.075	0.062	-0.092	0.057	-0.089	0.058
medio	-0.047	0.062	-0.062	0.061	-0.393	0.057	-0.388	0.058
midtech	-0.262	0.154	-0.293	0.151	-0.981	0.133	-0.979	0.133
hightech	-0.199	0.182	-0.241	0.181	-1.077	0.231	-1.071	0.229
normal	-0.484	0.226	-0.531	0.223	-1.625	0.167	-1.626	0.167
university	-0.593	0.103	-0.644	0.100	-1.254	0.134	-1.253	0.134
other	-0.228	0.127	-0.219	0.132	0.016	0.065	0.013	0.065
yrs incompl educ	-0.033	0.019	-0.046	0.019	-0.079	0.014	-0.079	0.014
econ activity /10	-0.001	0.036	-0.010	0.036	0.070	0.028	0.066	0.028
unemployment *10	-0.308	0.143	-0.384	0.138	0.319	0.115	0.303	0.115
ethnic	0.096	0.048	0.087	0.047	0.321	0.039	0.318	0.039
δ_1:young	0.013	0.009	-0.003	0.011	0.043	0.013	0.045	0.014
prime	0.031	0.009	0.066	0.011	0.011	0.012	0.008	0.012
old	0.071	0.038	0.077	0.052	0.023	0.053	0.019	0.054
married	0.061	0.108	0.114	0.136	1.098	0.138	1.085	0.137
net dsav pc /1000	1.587	0.454	2.478	0.552	0.942	0.446	0.885	0.442
ehat-nets	-0.673	0.453	-2.183	0.551	-0.544	0.451	-0.555	0.446
mar*age	-0.008	0.003	-0.007	0.004	-0.011	0.004	-0.010	0.004
age	-0.035	0.010	-0.033	0.010	-0.124	0.013	-0.126	0.012
age square /100	0.072	0.012	0.065	0.013	0.181	0.016	0.184	0.016
inter	-0.060	0.052	-0.076	0.053	-0.139	0.074	-0.137	0.074
medio	-0.111	0.047	-0.154	0.048	-0.325	0.063	-0.319	0.063
midtech	-0.321	0.092	-0.380	0.096	-0.964	0.088	-0.958	0.088
hightech	-0.349	0.117	-0.447	0.117	-1.196	0.129	-1.188	0.129
normal	-0.998	0.106	-1.120	0.103	-1.977	0.078	-1.973	0.078
university	-0.598	0.063	-0.706	0.066	-1.424	0.082	-1.418	0.081
other	-0.305	0.091	-0.236	0.095	0.118	0.091	0.118	0.091
yrs incompl educ	-0.037	0.013	-0.062	0.014	-0.109	0.017	-0.109	0.017
econ activity /10	0.017	0.027	-0.004	0.027	-0.017	0.032	-0.021	0.032
unemployment *10	0.012	0.105	-0.035	0.107	0.660	0.126	0.654	0.125
ethnic	0.158	0.036	0.154	0.037	0.256	0.049	0.253	0.049
β_1:cnst	0.042	0.174	0.317	0.252	-1.777	0.288	-1.892	0.301
age	0.026	0.008	0.019	0.009	0.081	0.012	0.083	0.013
age square /100	-0.007	0.010	0.005	0.013	-0.088	0.017	-0.092	0.018
inter	0.082	0.049	0.069	0.044	0.220	0.088	0.227	0.090
medio	0.245	0.041	0.225	0.038	0.523	0.075	0.539	0.076
midtech	0.440	0.078	0.385	0.073	0.813	0.104	0.854	0.100
hightech	0.582	0.100	0.532	0.084	1.028	0.141	1.079	0.124
normal	0.078	0.083	-0.050	0.095	0.972	0.121	1.045	0.116
university	0.827	0.048	0.748	0.061	1.311	0.110	1.370	0.106
other	0.211	0.067	0.146	0.073	-0.209	0.098	-0.214	0.118
yrs incompl educ	0.061	0.012	0.057	0.011	0.088	0.017	0.092	0.017
econ activity /10	0.172	0.023	0.180	0.020	0.176	0.032	0.176	0.030
unemployment *10	-0.664	0.090	-0.665	0.082	-0.704	0.121	-0.723	0.118
ethnic	-0.061	0.034	-0.037	0.031	-0.183	0.055	-0.194	0.050
β_2:cnst	0.221	0.278	0.327	0.252	0.273	0.331	0.169	0.331
age	-0.005	0.012	-0.017	0.013	0.009	0.014	0.013	0.014
age square /100	0.030	0.015	0.047	0.016	0.000	0.017	-0.005	0.017
inter	0.052	0.061	0.052	0.053	0.256	0.071	0.257	0.064
medio	0.284	0.054	0.291	0.048	0.360	0.064	0.364	0.060
midtech	0.377	0.136	0.384	0.120	0.566	0.121	0.589	0.136
hightech	0.371	0.169	0.363	0.144	0.813	0.245	0.844	0.231
normal	-0.419	0.166	-0.545	0.177	0.125	0.175	0.199	0.189
university	0.600	0.080	0.565	0.082	0.718	0.117	0.760	0.143
other	-0.404	0.106	-0.401	0.107	-0.042	0.074	-0.046	0.072
yrs incompl educ	0.046	0.018	0.037	0.016	-0.002	0.019	-0.001	0.018
econ activity /10	0.136	0.034	0.128	0.031	0.151	0.035	0.153	0.035
unemployment *10	-0.889	0.133	-0.890	0.119	-0.854	0.149	-0.866	0.141
ethnic	-0.073	0.045	-0.063	0.041	-0.145	0.048	-0.145	0.045
sigmaf	0.905	0.012			0.684	0.011		
sigmai	0.986	0.026			0.954	0.017		
sigf	0.704	0.023	0.949	0.149	-0.080	0.076	-0.140	0.075
sigi	0.653	0.040	0.765	0.090	0.268	0.062	0.216	0.069
Log Likelihood	11601	nr obs:6349			Log Likelihood:10160		nr obs:7293	

Table A3. Multinominal logit Estimates (ML std errors corrected, two step std errors not corrected)

MALES:	ML		two step		FEMALES ML		two step	
	coeff	std. error	coeff	std. error	coeff	std. error	coeff	std. error
δ_1:cnst	-1.079	0.519	-1.075	0.527	-4.269	0.547	-4.424	0.559
young	0.003	0.025	0.030	0.027	-0.101	0.026	-0.102	0.027
prime	-0.187	0.023	-0.221	0.025	-0.064	0.023	-0.059	0.025
old	0.006	0.116	-0.051	0.122	-0.030	0.102	-0.025	0.104
married	0.511	0.278	0.487	0.307	-2.436	0.271	-2.381	0.285
net dsav pc /1000	-6.247	1.124	-7.310	1.136	-2.525	0.875	-2.242	0.906
ehat-nets	2.871	1.119	4.687	1.144	1.349	0.888	1.030	0.920
mar*age	0.009	0.007	0.007	0.008	0.023	0.008	0.021	0.008
age	0.208	0.023	0.213	0.023	0.311	0.025	0.321	0.026
age square /100	-0.295	0.029	-0.297	0.030	-0.436	0.033	-0.450	0.034
inter	0.260	0.140	0.298	0.141	0.272	0.149	0.265	0.145
medio	0.138	0.122	0.223	0.123	0.500	0.126	0.484	0.122
midtech	0.359	0.229	0.397	0.232	1.527	0.164	1.527	0.161
hightech	0.394	0.266	0.466	0.271	1.926	0.236	1.927	0.225
normal	1.446	0.301	1.572	0.289	3.294	0.158	3.293	0.158
university	0.566	0.154	0.725	0.156	2.385	0.157	2.353	0.157
other	0.107	0.216	0.025	0.209	-0.221	0.193	-0.234	0.189
yrs incompl educ	0.032	0.033	0.061	0.034	0.208	0.031	0.206	0.033
econ activity /10	-0.019	0.067	-0.007	0.067	0.110	0.061	0.109	0.062
unemployment *10	-0.735	0.263	-0.759	0.275	-1.248	0.241	-1.287	0.240
ethnic	-0.167	0.092	-0.177	0.091	-0.339	0.097	-0.347	0.094
δ_2:cnst	-3.093	0.566	-3.043	0.579	-4.496	0.432	-4.446	0.428
young	0.024	0.026	0.027	0.029	-0.039	0.018	-0.049	0.020
prime	-0.172	0.026	-0.178	0.028	-0.081	0.019	-0.097	0.021
old	0.161	0.118	0.146	0.133	-0.013	0.090	-0.018	0.093
married	1.114	0.295	1.488	0.330	-0.937	0.213	-1.082	0.224
net dsav pc /1000	-4.621	1.186	-5.519	1.243	-2.269	0.753	-2.312	0.755
ehat-nets	3.207	1.185	2.679	1.245	-0.236	0.757	0.574	0.767
mar*age	-0.014	0.007	-0.019	0.008	0.005	0.005	0.006	0.005
age	0.265	0.025	0.256	0.025	0.239	0.018	0.248	0.018
age square /100	-0.306	0.030	-0.292	0.031	-0.281	0.022	-0.294	0.022
inter	0.205	0.143	0.244	0.145	-0.034	0.091	-0.020	0.090
medio	-0.052	0.127	-0.018	0.128	-0.352	0.084	-0.355	0.083
midtech	-0.282	0.263	-0.313	0.263	-0.532	0.173	-0.531	0.173
hightech	-0.326	0.308	-0.323	0.310	-0.458	0.299	-0.454	0.290
normal	-0.358	0.350	-0.451	0.345	-0.448	0.209	-0.461	0.209
university	-0.642	0.178	-0.662	0.177	-0.530	0.172	-0.531	0.175
other	-0.580	0.226	-0.581	0.227	-0.109	0.104	-0.135	0.104
yrs incompl educ	-0.035	0.036	-0.044	0.037	-0.052	0.026	-0.050	0.026
econ activity /10	-0.012	0.073	-0.024	0.073	0.154	0.049	0.157	0.048
unemployment *10	-1.153	0.285	-1.179	0.297	-0.226	0.195	-0.292	0.193
ethnic	0.139	0.098	0.129	0.097	0.431	0.065	0.437	0.065
β_1:cnst	-0.058	0.168	-0.757	0.195	-1.887	0.278	-1.911	0.298
age	0.031	0.008	0.052	0.008	0.083	0.012	0.084	0.013
age square /100	-0.013	0.010	-0.046	0.011	-0.092	0.016	-0.093	0.018
inter	0.082	0.048	0.093	0.044	0.227	0.088	0.228	0.090
medio	0.248	0.040	0.268	0.038	0.540	0.073	0.543	0.076
midtech	0.441	0.076	0.513	0.071	0.853	0.100	0.860	0.098
hightech	0.582	0.099	0.679	0.081	1.079	0.137	1.087	0.122
normal	0.106	0.080	0.372	0.071	1.044	0.112	1.054	0.112
university	0.854	0.046	1.002	0.047	1.371	0.102	1.381	0.104
other	0.211	0.064	0.290	0.070	-0.210	0.098	-0.211	0.118
yrs incompl educ	0.062	0.011	0.082	0.010	0.094	0.017	0.095	0.017
econ activity /10	0.169	0.022	0.176	0.020	0.177	0.032	0.178	0.030
unemployment *10	-0.676	0.087	-0.632	0.082	-0.727	0.121	-0.727	0.118
ethnic	-0.066	0.033	-0.111	0.029	-0.194	0.054	-0.197	0.050
β_2:cnst	-1.961	0.302	1.357	0.469	1.874	0.387	2.419	0.479
age	0.095	0.013	0.011	0.015	-0.024	0.015	-0.039	0.016
age square /100	-0.094	0.015	-0.008	0.016	0.034	0.018	0.050	0.019
inter	0.077	0.065	0.083	0.054	0.275	0.076	0.286	0.063
medio	0.204	0.058	0.369	0.052	0.529	0.069	0.563	0.065
midtech	0.215	0.137	0.679	0.132	1.001	0.131	1.089	0.144
hightech	0.143	0.167	0.684	0.159	1.375	0.251	1.471	0.236
normal	-0.610	0.166	0.649	0.227	1.208	0.161	1.363	0.200
university	0.227	0.091	1.159	0.140	1.407	0.127	1.474	0.150
other	-0.537	0.118	-0.065	0.120	-0.042	0.080	-0.041	0.071
yrs incompl educ	0.017	0.018	0.081	0.018	0.039	0.021	0.045	0.019
econ activity /10	0.118	0.036	0.126	0.031	0.117	0.038	0.113	0.036
unemployment *10	-1.218	0.145	-0.814	0.131	-0.904	0.157	-0.922	0.138
ethnic	-0.028	0.048	-0.228	0.047	-0.311	0.053	-0.350	0.053
sigmaf	0.878	0.011			0.689	0.013		
sigmai	1.110	0.033			1.138	0.039		
sig13	0.655	0.022	0.115	0.081	-0.138	0.068	-0.146	0.070
sig23	-0.923	0.048	0.741	0.204	0.803	0.070	0.988	0.138
Log Likelihood:11635		nr obs:6349			Log Likelihood:10120		nr obs:7293	

Table 3. Predicted partial derivatives of probability of participation for an average individual for ordered probit and multinominal logit selection model (standard errors in parentheses; * denotes significance at 5% level).

ORDERED PROBIT	males			females		
	formal	informal	not working	formal	informal	not working
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	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)

MULTINOMINAL LOGIT		males			females	
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)

Table 4. Estimated coefficients for wage equations (maximum likelihood results)

	ordered probit				multinominal 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	-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_{λ}	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)

Table 5. Predicted log wages for standard individuals (basis is 35 years old, living in La Paz and does not belong to an ethnic group, young is 20 years old, old is 50 years old and standard errors are in parentheses)

	Males		Females	
	ord probit	mult logit	ord probit	mult logit
FORMAL				
basis	0.883	0.851	0.223	0.151
	(0.042)	(0.040)	(0.110)	(0.101)
young	0.545	0.496	-0.262	-0.338
	(0.051)	(0.050)	(0.115)	(0.105)
old	1.189	1.147	0.312	0.224
	(0.047)	(0.045)	(0.135)	(0.122)
St. Cruz	1.050	1.020	0.399	0.332
	(0.039)	(0.037)	(0.104)	(0.094)
ethnic	0.822	0.785	0.040	-0.043
	(0.046)	(0.045)	(0.133)	(0.123)
INFORMAL				
basis	0.142	-0.473	0.460	1.357
	(0.060)	(0.084)	(0.070)	(0.111)
young	-0.025	-1.128	0.331	1.447
	(0.077)	(0.107)	(0.098)	(0.145)
old	0.445	-0.241	0.589	1.419
	(0.061)	(0.085)	(0.079)	(0.115)
St. Cruz	0.350	-0.198	0.663	1.566
	(0.055)	(0.078)	(0.063)	(0.108)
ethnic	0.069	-0.501	0.314	1.046
	(0.060)	(0.081)	(0.074)	(0.104)

figure 1. Predicted Probabilty of Participation in Sectors (ordered probit)

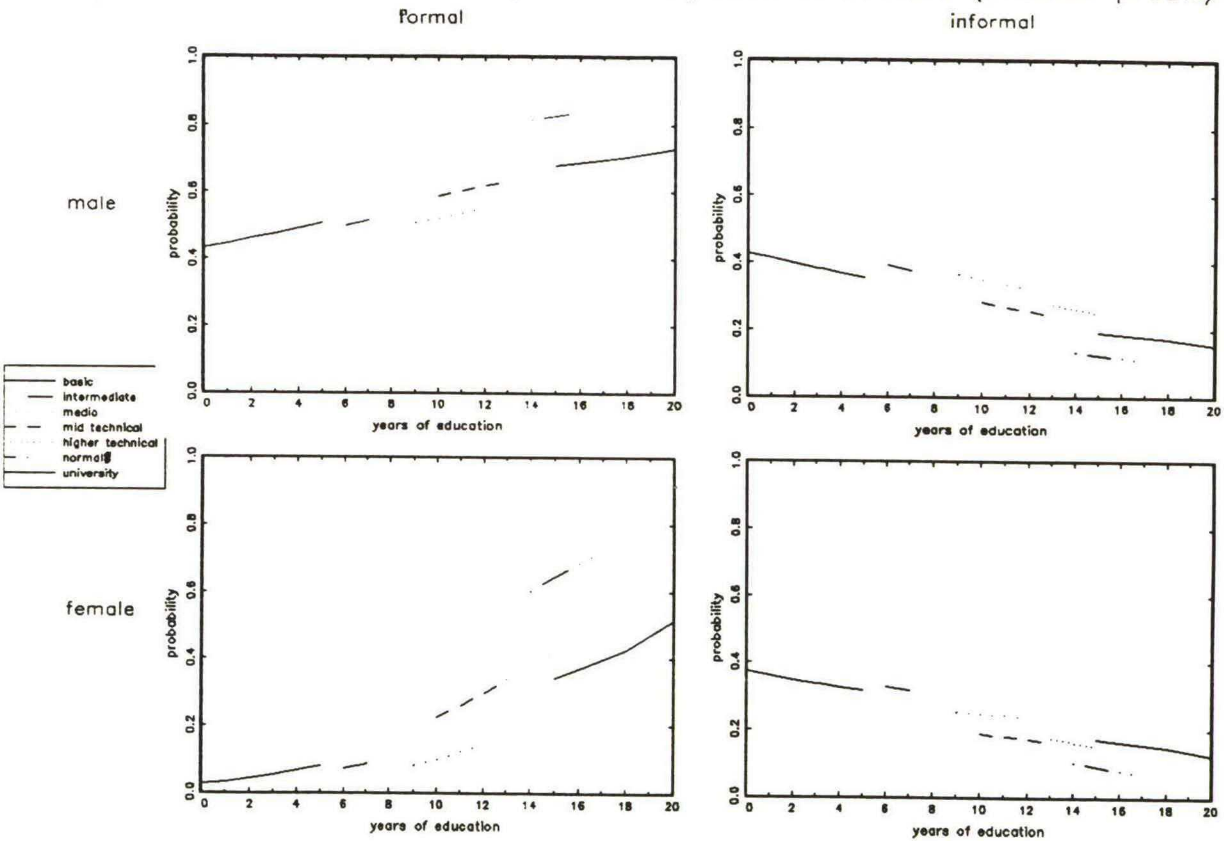


Figure 2. Predicted Probabilty of participation in sectors (mn logit)

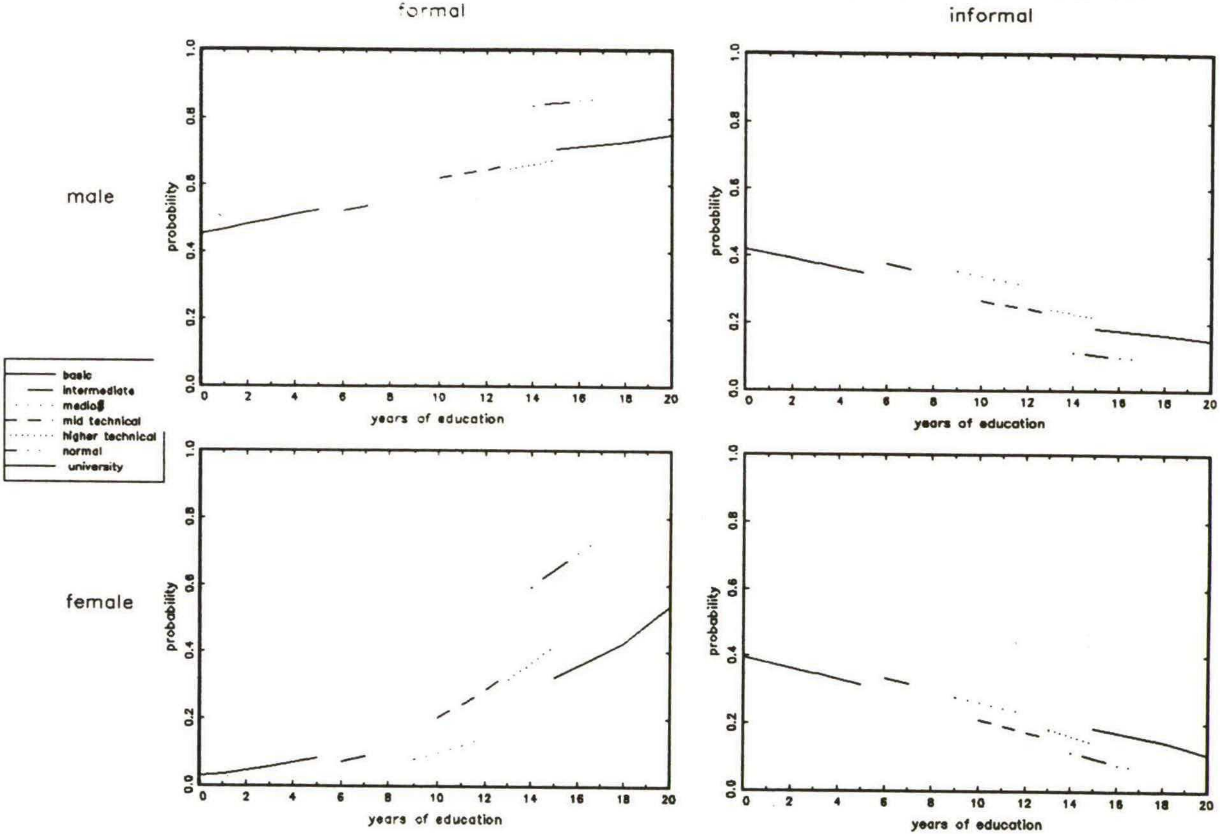


Figure 3. Pred cond log wage as a function of education (ordered probit)

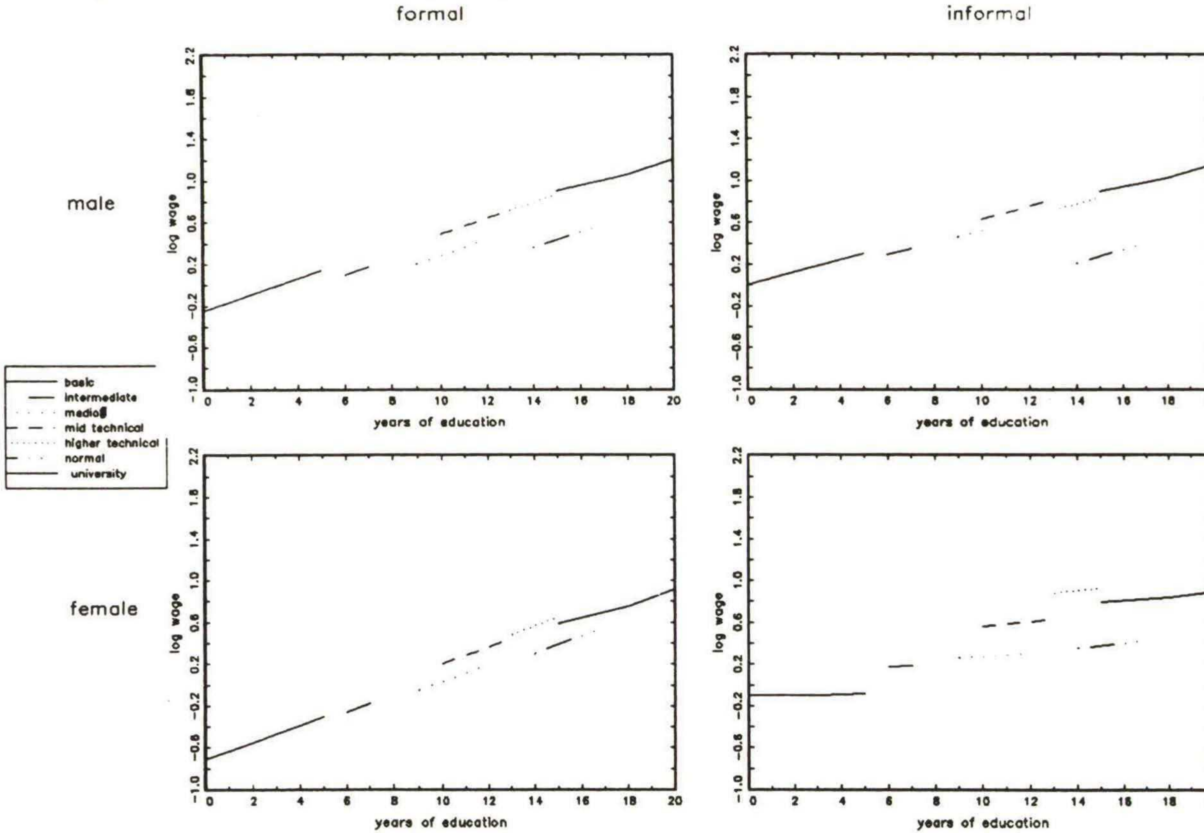


Figure 4. Pred cond log wage as a function of education (mn logit)!'

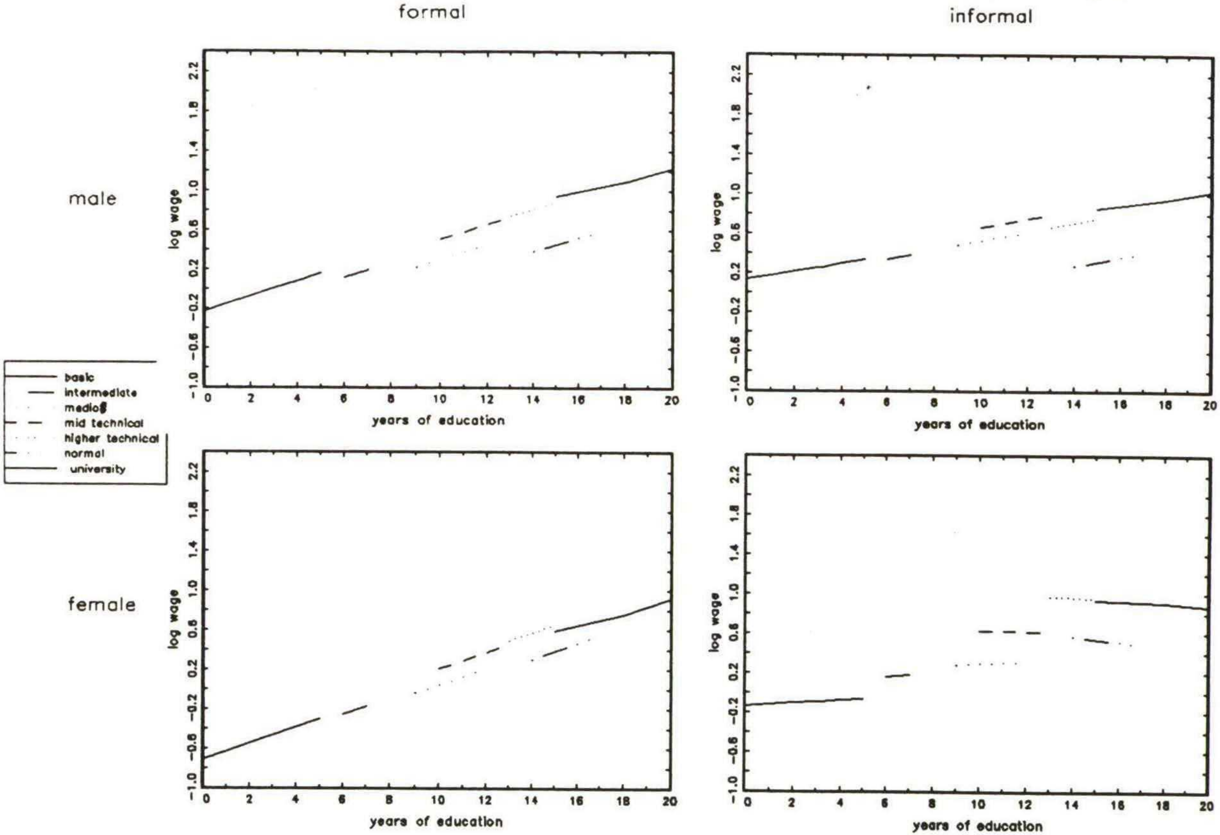


figure 5. Predicted log wage as a function of education (ordered probit)

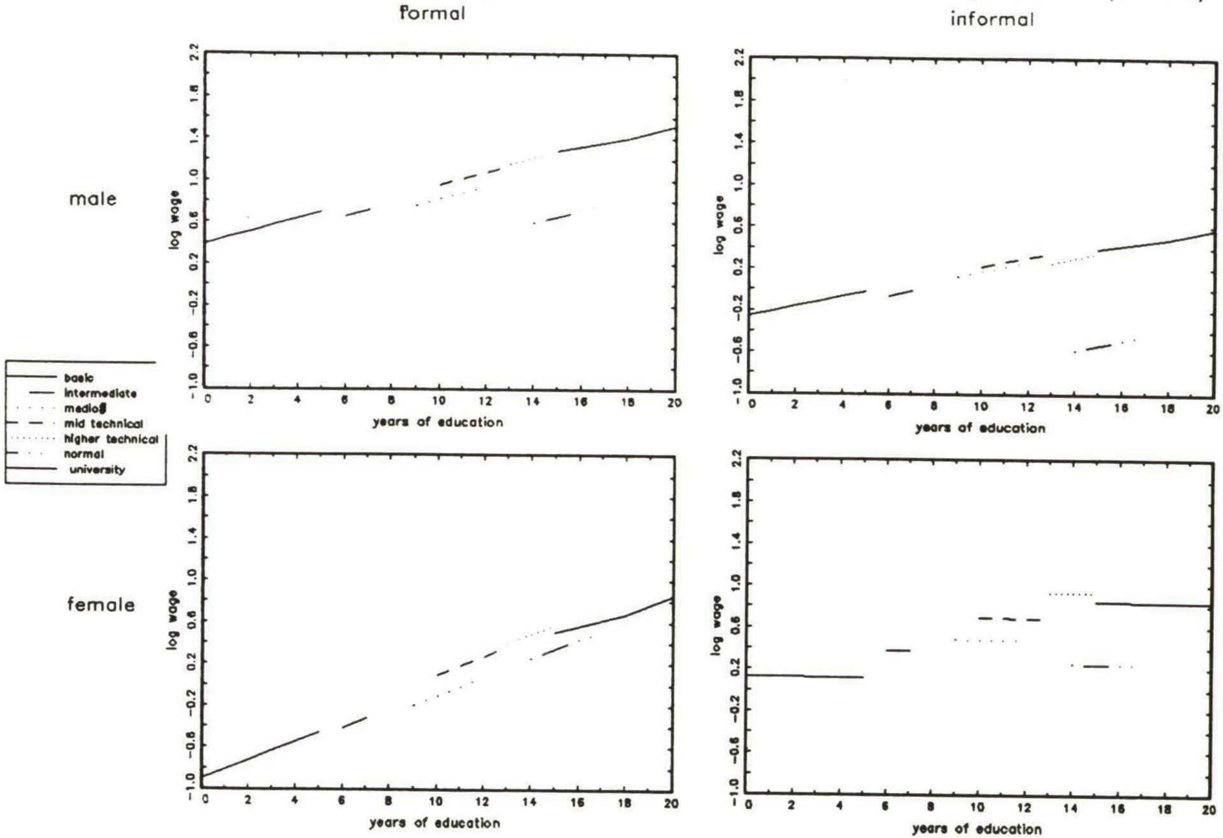
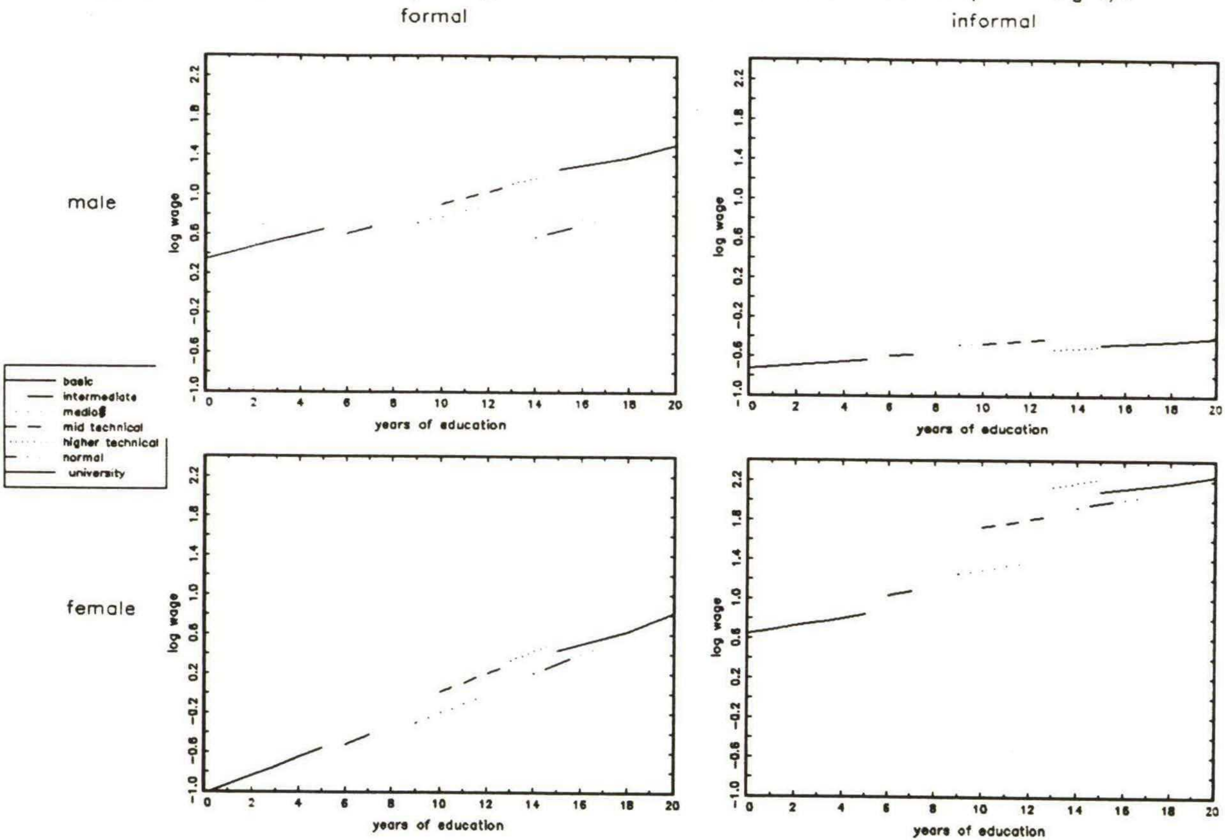


figure 6. Predicted log wage as a function of education (mn logit)!'



Discussion Paper Series, CentER, Tilburg University, The Netherlands:

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
9141	C. Fershtman and A. de Zeeuw	Capital Accumulation and Entry Deterrence: A Clarifying Note
9142	J.D. Angrist and G.W. Imbens	Sources of Identifying Information in Evaluation Models
9143	A.K. Bera and A. Ullah	Rao's Score Test in Econometrics
9144	B. Melenberg and A. van Soest	Parametric and Semi-Parametric Modelling of Vacation Expenditures
9145	G. Imbens and T. Lancaster	Efficient Estimation and Stratified Sampling
9146	Th. van de Klundert and S. Smulders	Reconstructing Growth Theory: A Survey
9147	J. Greenberg	On the Sensitivity of Von Neuman and Morgenstern Abstract Stable Sets: The Stable and the Individual Stable Bargaining Set
9148	S. van Wijnbergen	Trade Reform, Policy Uncertainty and the Current Account: A Non-Expected Utility Approach
9149	S. van Wijnbergen	Intertemporal Speculation, Shortages and the Political Economy of Price Reform
9150	G. Koop and M.F.J. Steel	A Decision Theoretic Analysis of the Unit Root Hypothesis Using Mixtures of Elliptical Models
9151	A.P. Barten	Consumer Allocation Models: Choice of Functional Form
9152	R.T. Baillie, T. Bollerslev and M.R. Redfearn	Bear Squeezes, Volatility Spillovers and Speculative Attacks in the Hyperinflation 1920s Foreign Exchange
9153	M.F.J. Steel	Bayesian Inference in Time Series
9154	A.K. Bera and S. Lee	Information Matrix Test, Parameter Heterogeneity and ARCH: A Synthesis
9155	F. de Jong	A Univariate Analysis of EMS Exchange Rates Using a Target
9156	B. le Blanc	Economies in Transition
9157	A.J.J. Talman	Intersection Theorems on the Unit Simplex and the Simplotope

No.	Author(s)	Title
9158	H. Bester	A Model of Price Advertising and Sales
9159	A. Özcam, G. Judge, A. Bera and T. Yancey	The Risk Properties of a Pre-Test Estimator for Zellner's Seemingly Unrelated Regression Model
9160	R.M.W.J. Beetsma	Bands and Statistical Properties of EMS Exchange Rates: A Monte Carlo Investigation of Three Target Zone Models Zone Model
9161	A.M. Lejour and H.A.A. Verbon	Centralized and Decentralized Decision Making on Social Insurance in an Integrated Market Multilateral Institutions
9162	S. Bhattacharya	Sovereign Debt, Creditor-Country Governments, and
9163	H. Bester, A. de Palma, W. Leininger, E.-L. von Thadden and J. Thomas	The Missing Equilibria in Hotelling's Location Game
9164	J. Greenberg	The Stable Value
9165	Q.H. Vuong and W. Wang	Selecting Estimated Models Using Chi-Square Statistics
9166	D.O. Stahl II	Evolution of $Smart_n$ Players
9167	D.O. Stahl II	Strategic Advertising and Pricing with Sequential Buyer Search
9168	T.E. Nijman and F.C. Palm	Recent Developments in Modeling Volatility in Financial Data
9169	G. Asheim	Individual and Collective Time Consistency
9170	H. Carlsson and E. van Damme	Equilibrium Selection in Stag Hunt Games
9201	M. Verbeek and Th. Nijman	Minimum MSE Estimation of a Regression Model with Fixed Effects from a Series of Cross Sections
9202	E. Bomhoff	Monetary Policy and Inflation
9203	J. Quiggin and P. Wakker	The Axiomatic Basis of Anticipated Utility; A Clarification
9204	Th. van de Klundert and S. Smulders	Strategies for Growth in a Macroeconomic Setting
9205	E. Siandra	Money and Specialization in Production
9206	W. Härdle	Applied Nonparametric Models
9207	M. Verbeek and Th. Nijman	Incomplete Panels and Selection Bias: A Survey
9208	W. Härdle and A.B. Tsybakov	How Sensitive Are Average Derivatives?

No.	Author(s)	Title
9209	S. Albæk and P.B. Overgaard	Upstream Pricing and Advertising Signal Downstream Demand
9210	M. Cripps and J. Thomas	Reputation and Commitment in Two-Person Repeated Games
9211	S. Albæk	Endogenous Timing in a Game with Incomplete Information
9212	T.J.A. Storcken and P.H.M. Ruys	Extensions of Choice Behaviour
9213	R.M.W.J. Beetsma and F. van der Ploeg	Exchange Rate Bands and Optimal Monetary Accommodation under a Dirty Float
9214	A. van Soest	Discrete Choice Models of Family Labour Supply
9215	W. Güth and K. Ritzberger	On Durable Goods Monopolies and the (Anti-) Coase-Conjecture
9216	A. Simonovits	Indexation of Pensions in Hungary: A Simple Cohort Model
9217	J.-L. Ferreira, I. Gilboa and M. Maschler	Credible Equilibria in Games with Utilities Changing during the Play
9218	P. Borm, H. Keiding, R. Mclean, S. Oortwijn and S. Tijs	The Compromise Value for NTU-Games
9219	J.L. Horowitz and W. Härdle	Testing a Parametric Model against a Semiparametric Alternative
9220	A.L. Bovenberg	Investment-Promoting Policies in Open Economies: The Importance of Intergenerational and International Distributional Effects
9221	S. Smulders and Th. van de Klundert	Monopolistic Competition, Product Variety and Growth: Chamberlin vs. Schumpeter
9222	H. Bester and E. Petrakis	Price Competition and Advertising in Oligopoly
9223	A. van den Nouweland, M. Maschler and S. Tijs	Monotonic Games are Spanning Network Games
9224	H. Suehiro	A "Mistaken Theories" Refinement
9225	H. Suehiro	Robust Selection of Equilibria
9226	D. Friedman	Economically Applicable Evolutionary Games
9227	E. Bomhoff	Four Econometric Fashions and the Kalman Filter Alternative - A Simulation Study

No.	Author(s)	Title
9228	P. Borm, G.-J. Otten and H. Peters	Core Implementation in Modified Strong and Coalition Proof Nash Equilibria
9229	H.G. Bloemen and A. Kapteyn	The Joint Estimation of a Non-Linear Labour Supply Function and a Wage Equation Using Simulated Response Probabilities
9230	R. Beetsma and F. van der Ploeg	Does Inequality Cause Inflation? - The Political Economy of Inflation, Taxation and Government Debt
9231	G. Almekinders and S. Eijffinger	Daily Bundesbank and Federal Reserve Interventions - Do they Affect the Level and Unexpected Volatility of the DM/$-Rate?
9232	F. Vella and M. Verbeek	Estimating the Impact of Endogenous Union Choice on Wages Using Panel Data
9233	P. de Bijl and S. Goyal	Technological Change in Markets with Network Externalities
9234	J. Angrist and G. Imbens	Average Causal Response with Variable Treatment Intensity
9235	L. Meijdam, M. van de Ven and H. Verbon	Strategic Decision Making and the Dynamics of Government Debt
9236	H. Houba and A. de Zeeuw	Strategic Bargaining for the Control of a Dynamic System in State-Space Form
9237	A. Cameron and P. Trivedi	Tests of Independence in Parametric Models: With Applications and Illustrations
9238	J.-S. Pischke	Individual Income, Incomplete Information, and Aggregate Consumption
9239	H. Bloemen	A Model of Labour Supply with Job Offer Restrictions
9240	F. Drost and Th. Nijman	Temporal Aggregation of GARCH Processes
9241	R. Gilles, P. Ruys and J. Shou	Coalition Formation in Large Network Economies
9242	P. Kort	The Effects of Marketable Pollution Permits on the Firm's Optimal Investment Policies
9243	A.L. Bovenberg and F. van der Ploeg	Environmental Policy, Public Finance and the Labour Market in a Second-Best World
9244	W.G. Gale and J.K. Scholz	IRAs and Household Saving
9245	A. Bera and P. Ng	Robust Tests for Heteroskedasticity and Autocorrelation Using Score Function
9246	R.T. Baillie, C.F. Chung and M.A. Tieslau	The Long Memory and Variability of Inflation: A Reappraisal of the Friedman Hypothesis

No.	Author(s)	Title
9247	M.A. Tieslau, P. Schmidt and R.T. Baillie	A Generalized Method of Moments Estimator for Long-Memory Processes
9248	K. Wärneryd	Partisanship as Information
9249	H. Huizinga	The Welfare Effects of Individual Retirement Accounts
9250	H.G. Bloemen	Job Search Theory, Labour Supply and Unemployment Duration
9251	S. Eijffinger and E. Schaling	Central Bank Independence: Searching for the Philosophers' Stone
9252	A.L. Bovenberg and R.A. de Mooij	Environmental Taxation and Labor-Market Distortions
9253	A. Lusardi	Permanent Income, Current Income and Consumption: Evidence from Panel Data
9254	R. Beetsma	Imperfect Credibility of the Band and Risk Premia in the European Monetary System
9301	N. Kahana and S. Nitzan	Credibility and Duration of Political Contests and the Extent of Rent Dissipation
9302	W. Güth and S. Nitzan	Are Moral Objections to Free Riding Evolutionarily Stable?
9303	D. Karotkin and S. Nitzan	Some Peculiarities of Group Decision Making in Teams
9304	A. Lusardi	Euler Equations in Micro Data: Merging Data from Two Samples
9305	W. Güth	A Simple Justification of Quantity Competition and the Cournot-Oligopoly Solution
9306	B. Peleg and S. Tijs	The Consistency Principle For Games in Strategic Form
9307	G. Imbens and T. Lancaster	Case Control Studies with Contaminated Controls
9308	T. Ellingsen and K. Wärneryd	Foreign Direct Investment and the Political Economy of Protection
9309	H. Bester	Price Commitment in Search Markets
9310	T. Callan and A. van Soest	Female Labour Supply in Farm Households: Farm and Off-Farm Participation
9311	M. Pradhan and A. van Soest	Formal and Informal Sector Employment in Urban Areas of Bolivia

P.O. BOX 90153 5000 LE TILBURG THE NETHERLANDS