Documento de Trabajo No. 01/00
Marzo 2000

**Wage Differentials Between the Formal and the
Informal Sector in Urban Bolivia**

por
Trine Monsted

Wage Differentials Between the Formal and the

Informal Sector in Urban Bolivia

Trine Monsted[1]
ILSEC-UCB

La Paz, Bolivia

(February 2000)

Summary: Ibis paper analyses wage differentials between the formal and informal sectors in urban Bolivia using household survey data. As m other studies the wage differential between the formal and the informal sector is found to be quite large. The wage gap is estimated by a human capital model correcting for selectivity bias using Heckman's two-step procedure. A surprising result is that migrants tend not be employed in the formal sector. Another noticeable result is that the probability of entering the formal sector was a hump shaped function of altitude probably because the government cities (La Paz and Sucre) are located at mid altitudes. The wage generating processes for the two sectors are quite different, and the results indicate that productivity is relatively more important for wages in the informal sector than in the formal sector. It is also found that returns to education are generally higher for formal sector workers. In a non-linear specification of returns to schooling the conclusion is that basic education is not rewarded in the labour market, therefore seriously questioning the quality of the educational system. An Oaxaca decomposition of the wage differential between the two sectors shows that there are differences in characteristics between formal and informal sector workers, but these characteristics are also rewarded differently in the two sectors. The analysis also confirms that education and experience are the main factors contributing to the wage differentials between sectors.

1.   Introduction

Traditional human capital theory predicts that education and training are the main determinants of earnings and thereby of poverty, which explains their importance as policy variables. A higher level of schooling and experience implies a higher level of income (e.g. Andersen, 1998 and Jensen et al, 1997). Bolivia is the poorest country in South America, thus poverty alleviation is a serious consideration, which is why earnings are in focus. In addition the existence of dual labour markets makes it important to determine the distinguishing features of these markets.

The purpose of this paper is to examine wage differentials in Bolivia especially between the formal and informal labour markets. This is important because different mechanisms may be working at the two markets, which implies that policies have to be constructed particularly to the market in question. The data comes from the National Statistical Office (INE) and are collected in 8 rounds between 1989 and 1995 in the main cities in Bolivia.

The paper is organized as follows. First, some data issues and definitions of the relevant variables will be presented. Next, some descriptive statistics will provide an overview of the labour markets in urban Bolivia including the wage differential between the formal and informal labour markets, which is the focus of this paper. Third, a human capital model correcting for sample selection bias will be estimated assuming both linear and non-linear returns to education. Fourth, an Oaxaca decomposition of the wage gap will be calculated and the last section concludes.

2.   The formal and informal labour markets in Bolivia

The division of the labour market into a formal and an informal market is significant in most developing countries, where persons who cannot find employment in the formal labour market are forced to employ themselves in other ways. Ibis is also called hidden unemployment, as the people employed in the informal market would be unemployed in an economy with only a formal market.

Normally it is observed that the productivity level at the informal market is lower than that prevailing at the formal market, which can be explained by several factors. First, differences in technology is the main reason for variance in productivity. The formal market normally has access to more advanced equipment than the informal market has, which implies a higher productivity for workers employed in the formal labour market. Second, people working in the informal market may have fewer skills and lower abilities implying lower productivity for those workers. Since it is expected that productivity will be lower in the informal sector, wages will also be expected to be lower. Ibis analysis tries to identify the factors that determine participation in the two sectors and the distinguishing features of the wage generating processes. These processes are expected to differ which is why the division is important.

2.1   Data and definitions

The data comes from 8 rounds of household surveys (Encuesta Integrada de Hogares) collected by the national statistical office (INE) in Bolivia. Unavailability makes it impossible to use round 2 in this paper. The questions have been posed to people living in the

largest cities in Bolivia, and the subsequent analysis therefore applies only to the urban areas. Panel data models would have used the information in the sample more efficiently since time series information is included in these models, but the present dataset does not allow for such an analysis. Ibis is because the sample is not identical between rounds. Therefore, the model presented will only use cross sectional information.

The working population is defined as all respondents aged between 12 and 66 having a strictly positive income. This population is divided into two groups, those employed in the formal market and those employed in the informal market. The dataset does not contain any explicit variable describing whether a person belongs to the formal or the informal sector, but it is possible to construct such a variable from a combination of some of the other variables. When asked about their occupational status seven different categories were possible: worker, employed, employer, professional independent, self-employed, domestic worker and family worker. A rough division of the two last categories as the informal market and the five first as the formal may be appropriate, but the data allow for some further adjustments. These lead to the following definition of the informal sector. A person employed as a domestic worker or family worker having a positive salary will always be included in the sample of informal market workers. In addition the type of firm in combination with the form of payment, availability of social security, membership of a union[2] and the status of the job as permanent or temporary have been used. If a person not employed as a domestic worker or family worker, is employed outside the public sector and gets his or her salary as in-kind compensation or per job accomplished, is not a member of a union or has access to a social security scheme and has a temporary employment status, this person will be included in the informal sector. Ibis division leads to the following distribution of formal and informal sector workers in the different rounds.

Table 1: Participation Rates

Round	Participation Rate %	Formal Sector	Informal Sector
1(1989)	46,37	2527 (45,47%)	3031 (54,53%)
3 (1990)	43,40	5364 (62,5 1%)	3217 (37,49%)
4(1991)	45,77	5103 (59,81%)	3429 (40,19%)
5 (1992)	45,26	4888 (61,37%)	3077 (38,63%)
6(1993)	47,52	3654 56,27%	2840 (43,73%)
7 (1994)	48,37	6978 (76,35%)	2161 (23,65%)
8(1995)	49,33	6799 (77,78%)	1942 (22,22%)

Source: Author's calculations

As can be seen. from table 1, the participation rate averages 46%, that is, 46% of the population between 12 and 66 years are actually working. Of these, some are employed at the formal market and others at the informal market. There is an increasing share of employed at the formal labour market. Ibis pattern could be completely random though. First, the sample

---

is not identical over time; every year different people have been asked different questions. Second, in some rounds it has not been possible to distinguish between people being members of a union or not which again may have led to a downward bias in the share of employed at the informal markets. Finally, although the definition of the informal market has been approximately the same across rounds, a lower share of formal sector workers might occur if one used other division criteria. Ibis paper will continue with this distinction though, since there is no obvious way to improve the division criteria from the data at hand.

2.2 Earnings

Earnings are measured in the surveys as the income a particular person is receiving. In table 2 the wage gap between the two sectors is obvious.

Table 2: Wage Gap Between the Formal and the Informal Sectors in Bolivia

Round	Formal Sector		Informal Sector		T-statistic[a]
	Mean	Std. Dev.	Mean	Std. Dcv.	
1 (1989)	2,25	0,06	1,76	0,07	277,11
3 (1990)	2,94	0,14	1,70	0,06	476,78
4 (1991)	3,20	0,08	2,08	0,07	666,16
5 (1992)	3,60	0,08	2,40	0,08	651,82
6 (1993)	4,33	0,11	2,65	0,07	709,83
7 (1994)	4,77	0,08	2,12	0,06	1421,0
8 (1995)	5,39	0,11	2,45	0,09	1079,1

a) H_0: Formal wage = Informal wage
Source: Author's calculations

In all years, the wage for formal sector employees has been higher than for those employed in the informal sector, and the gap has increased substantially during the period under investigation. Ibis gap between wages in the two sectors reflects the lower productivity in the informal labour market, but also other effects like generally lower education levels in the informal sector, a higher proportion of people not having Spanish as their mother tongue etc. The rest of this paper will be concerned with an analysis of this wage gap in order to identify the distinguishing features of the wage generating processes in the two markets.

3. The Wage Gap

Ibis section will try to identify some of the factors influencing the difference in wages between the formal and informal labour markets in Bolivia. Several individual characteristics may turn out to be significant explanations of the wage gap. First, human capital theory predicts that education and experience are the main determinants of earnings, as also presented by Mincer (1974). Other characteristics like ethnicity, region of residence, and gender are also included here as determinants of earnings. In section 4 a human capital model correcting for selectivity bias will be estimated, taking ah these aspects into account simultaneously.

3.1 Education and experience

As mentioned above education and experience are important determinants of earnings. A high level of education generally leads to higher earnings, as does experience. Education can

be measured either by completed level or by years of education. Experience is measured by on the job experience and general potential experience at the labour market.[3]

Table 3: Education Levels Across Sectors

Round	Formal Sector			Informal Sector			T-statistic[a]
	N	Mean	Std. Dev.	N	Mean	Std. Dev.	
1 (1989)	2527	10,24	0,085	3031	6,61	0,074	1701,6
3 (1990)	5151	9,38	0,058	3129	8,41	0,064	709,3
4 (1991)	4735	12,16	0,083	3288	9,40	0,073	1537,9
5 (1992)	4589	14,09	0,13	2975	9,70	0,092	1600,2
6 (1993)	3315	13,90	0,15	2748	9,85	0,091	1238,9
7 (1994)	6978	10,46	0,071	2161	7,82	0,09	1412,5
8 (1995)	6799	13,48	0,10	1942	8,73	0,11	1804,5

a) H_0: Mean education in formal sector = Mean education in informal sector
Source: Author's calculations

The educational differences between the two sectors are illustrated in table 3, measured as years of education.

Education is thus unevenly distributed between the two sectors. In all rounds, years of education for informal sector workers are significantly lower than that of formal sector workers. Earnings are positively related to education (e.g. Wood and Patrinos, 1994). And since education is lower in the informal sector, earnings in this sector are also lower.

3.2   Ethnicity

In most countries with different ethnic groups; indigenous people usually earn less than no indigenous. Using language as a proxy for ethnicity,[4] the following results emerge.

Table 4 Ethnicity and Wages

4.1 Round 1, 1989

	Indigenous	Non-Indigenous
Informal	1,57	1,96
Formal	1,69	2,50

4.2 Round 3, 1990

	Indigenous	Non-Indigenous
Informal	1,55	1,77
Formal	2,09	3,46

4.3 Round 6, 1993

	Indigenous	Non-Indigenous
Informal	2,24	2,97
Formal		

4.4 Round 7, 1994

	Indigenous	Non-Indigenous
Informal	1,97	2,22
Formal	4,17	5,20

Mean wage for each sector, conditional on ethnicity
Source: Author's calculations

---

[3]   Experience is calculated as age - years of education - 6, where 6 is the age where schooling normally starts. Unemployment is the reason why this is only potential experience. See Oaxaca (1973) for further problems with using this measure.

[4]   People of ethnic origin are defined as those speaking an indigenous language.

Indigenous people tend to earn less than non-indigenous regardless of sector choice and the difference is significant.[5] Ibis could indicate that indigenous are either less productive or possess fewer skills than non-indigenous, or discrimination against indigenous people.

3.3   Altitude

An earlier study of the Bolivian labour market (Andersen, 1999) shows that altitude has significant explanatory power for wages. This paper will examine this result in order to see whether it still holds after taking the dual labour markets into account.

Figure 1: Altitude and Wages

Cities: 1=Cobija (0.221), 2=Trinidad (0.236), 3=Santa Cruz (0.416), 4=Tarija (1.866), 5=Cochabamba (2.558), 6=Sucre (2.790), 7=La Paz (3.640), 8=Oruro (3.709), 9=E1 Alto (3.848), 1O=Potosi (4.070). Height above sea level (in km) is mentioned in parentheses.
Source: Author's calculations

As is visible from figure 1, there are quite large differences between the formal and informal sector wages, also in the different regions. Wages are generally falling with altitude, which could be a reflection of the lower productivity in the cities higher aboye sea level than others.[6]

3.4   Gender

Normally women earn less than men, which is also illustrated in figure 2. It is seen that the difference between genders in the informal sector is even bigger than the difference between genders in the formal sector.

After having identified some explanations of the wage gap, a model incorporating sector choice and the possible differences in the determination of wages in the two sectors, will be presented in the next section.

---

[5]   Even though results are only presented using rounds 1 and 8, the general picture carries over to the other rounds
[6]   Using a t-test shows significant differences in ah cases. Results are not reported here, but are available upon request. 6 Even though results are only presented using rounds 1 and 8, the general picture carries over to the other rounds.

Figure 2: Gender and Wages

Source: Author's calculations.
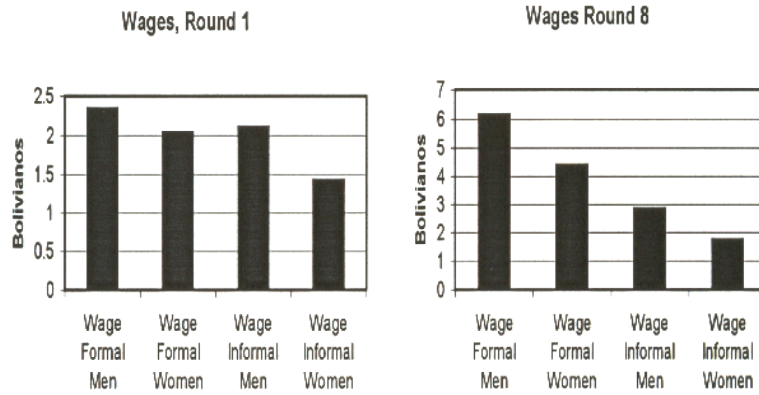
4. Wage Estimation

This section will look more closely at the earnings differential between the formal and the informal sectors in urban Bolivia. Estimation of a human capital model with correction for sample selection bias will be carried out. Then returns to education under different assumptions about the functional relationship will be tested and, lastly, a decomposition of the wage differential will be presented.

4. *The Model*

The main point of interest is estimation of a wage equation of the type proposed by Mincer (1974). However, in a sample like the one at hand, there are a division of individuals into two different sectors, and this division may not be random. This possible non-randomness of the sample implies that OLS may not be consistent, and the solution to this problem is to apply a 2-step procedure proposed by Heckman (1976, 1979).

The first step is to estimate a probit selection equation, which determines the probability of entering the formal sector [7] (one minus this probability is thus the probability of entering the informal sector) given some individual characteristics. The equation will look like

$$F_i \quad = \quad \beta`Y_i \quad + \quad v_i,$$

where F_i is a binary choice variable, which takes the value 1 if the worker is employed in the formal sector, and O if not. Y_i is a vector of individual characteristics expected to influence the choice of sector. The method of estimation is maximum likelihood.

From this equation it is possible to calculate the probability of being in the formal sector given that you are a member of the urban labour force. This probability is called lambda, and when it is included in the earnings equations in the next step, it corrects for a possible sample selection bias.

---

[7] It could have been interesting to investigate whether self-selection has an effect on participation in the labour force in general. However, this *is* outside the scope *of this* paper, and for now it is assumed that labour force participation is random.

An earnings equation will be estimated separately for each market because it is not the same process that generates wages in the two sectors, i.e. the determinants may differ significantly between sectors.

Additionally, for the sake of identification of the empirical model it is important not to include all variables from the selection equations in the wage equations. Ibis problem is easily solved, more variables are relevant for earnings than for sector choice and some of the proxies in the selection equations are exchanged with their true variables, i.e. experience.

The dependent variable is log wages in each sector, and the estimated equations will have the following form.

$$Ln(W_i^f) = \alpha + \beta^f X_i + u_i^f$$
$$Ln(W_i^i) = \gamma + \beta^i X_i + w_i^i,$$

where the superscript f denotes the formal sector and i the informal sector. X is a vector of individual characteristics expected to influence the income generating process in each sector, and u and w are residuals. The reason for taking the logarithm to wages is that the distribution of earnings is skew, and better results emerge when the logarithm is used. The method of estimation is thus OLS, which produces unbiased and consistent estimates of the coefficients. (See Heckman 1976, 1979 or Greene, 1997 for more technical details).

4.2 The selection equation

The results of estimation are presented in the table below. The dependent variable is a binary choice variable with the value 1 when an individual belongs to the formal sector and zero if the individual belongs to the informal sector. The point of departure was to include the following explanatory variables in the selection equation: years of education, age, age squared (both as a proxy for experience), altitude, altitude squared (both as a proxy for region /productivity), whether the individual is a household head, female, married and migrant. In the table only the significant coefficients are reported (t-statistics in parentheses).

As is visible from the table, education increases the probability of employment in the formal sector significantly. Ibis is expected since the tasks performed by the formal sector are in general more complex and therefore require more skilled labour. Experience as approximated by age seems to have a hump shaped relation with the estimated probability. Until the age of 49[8] the probability of being employed formally is increasing with age, thereafter the probability is decreasing. One possible interpretation is that the formal sector mainly employs individuals in their most productive years and the informal sector is stuck with the young and inexperienced or the old, both groups having a lower productivity than those employed in the other sector.

One very significant finding is that the relation between the probability of being employed in the formal sector and altitude of residence is hump shaped. Ibis means that the probability of formal employment is increasing for residents in the lowlands and valleys, but

---

[8]   This is an outweighed average of the optimum from the seven estimated equations. he lowest age is 44 years in round 1 and the highest is 52 years in round 6.

is decreasing in the highlands.[9] This is probably because the two government cities (Sucre and La Paz, cities 6 and 7) contain most of the public sector jobs, which account for a large share of formal jobs. It is also consistent with the result that indigenous people have a higher probability of informal employment than non-indigenous, because the density of the indigenous population is increasing with altitude (Wood and Patrinos, 1994).

Table 5: Probit Estimation of the Probability of Being in the Formal Sector

	1989	1990	1991	1992	1993	1994	1995
Intercept	-4,31	-5,38	-5,42	-5,43	-5,50	-5,95	-5,92
	(-91,40)	(-176,68)	(-171,87)	(-163,32)	(-140,73)	(-280,66)	(-277,04)
Years of	0,027	0,0048	0,0030	0,00328	0,0030	0,00323	0,0028
Education	(20,99)	(6,76)	(4,15)	(6,03)	(4,38)	(8,83)	(7,52)
Age	0,008	0,021	0,0205	0,021	0,024	0,011	0,001
	(3,06)	(12,64)	(11,49)	(11,34)	(11,46)	(9,05)	(8,11)
Age Squared	-0,00009	-0,00021	-0,00021	-0,00021	-0,00023	-0,00011	-0,0001
	(-2,70)	(-10,26)	(-9,33)	(-9,22)	(-9,06)	(-7,38)	(-6,41)
Altitude	0,23	0,179	0,189	0,188	0,16	0,135	0,13
	(14,62)	(20,87)	(19,31)	(19,12)	(12,80)	(18,31)	(17,63)
Altitude	-0,043	-0,029	-0,03 1	-0,03 1	-0,022	-0,019	-0,018
Squared	(-11,82)	(-14,48)	(-13,97)	(-13,77)	(-7,92)	(-11,40)	(-10,82)
Head	-	-	0,0152	0,010	-	-	-
			(1,68)	(1,27)			
Married	-	0,034	0,043	0,049	0,053	0,024	0,018
		(4,66)	(5,59)	(6,12)	(5,68)	(4,55)	(3,35)
Female	-0,09	-	0,06	0,052	0,06	0,028	0,016
	(-8,12)		(6,87)	(6,64)	(8,01)	(6,29)	(3,47)
Migrant	-0,041	-0,025	-0,046	-0,010	-0,017	-	-0,018
	(-2,77)	(-2,90)	(-2,24)	(-1,11)	(-1,33)		(-1,33)
Ethnicity	-0,056	-0,008	-	-	-0,0 10	-0,017	-
	(4,64)	(-1,14)			(-1,26)	(-3,27)	
N	5558	8581	8532	7965	6494	9139	8741

Source: Author's calculations; tic significance level is 10%.

Being a household head is significantly increasing the estimated probability in two cases out of seven, thus this is not a very robust result.

The effect of being female is significantly positive, but vanes over time in magnitude. One might argue that the probability of being employed is lower for women than for men, but once they get into the labour force, they have a higher probability of entering the formal sector than their male counterparts. The reason why women are generally less likely to work than men is that women have other considerations before entering the labour market. They may have children requiring maternal care etc., which may make women more reluctant to enter the labour market as frequently as men do.

Being married increases the estimated probability as is often the case in such estimations. Ibis may be due to marriage being correlated with some other unobserved characteristics like e.g. culture.

Migration is a significant factor explaining employment in the two sectors. Migration is defined as a person who has moved in the last 5 years[10] and the relation is quite robust. A

---

[9]   The probability of employment in the formal sector is increasing for an individual living in areas below 3230 meters on average. After this point the probability is decreasing. This means that individuals living in La Paz, El Alto, Potosi and Oruro have a lower probability of entering tic formal than individuals living in cities situated lower.

[10]   In some rounds it was not possible to use this definition because only questions regarding changes of location during the past year were asked. This was tic case in round 4 and 8. However, tic results do not change in these cases.

recently migrated person has a higher probability of being employed in the informal sector, whereas a non-migrant is normally employed in the formal sector. There are several plausible reasons for this. One factor could be that migrants may come from rural areas, where schooling is generally lower and qualifications in general are worse. Another reason is the fact that migrants may have moved location because their earnings in the previous location were insufficient. Since earnings and qualifications in general are related positively (Wood and Patrinos, 1994), the reason for low earnings stays with the individual even after moving. Therefore the probability of employment in the formal sector decreases with migration. A third reason may be that when a person migrates, he or she does not have any connections in the new location which could mean that it gets harder to obtain employment in the formal sector.

4.2   The earnings equations

As a starting point the following explanatory variables are tested, in order to see how well they explain wages. Education, general experience and experience in current job, both squared to allow for a hump shaped profile of experience. The variable controlling for selection bias, lambda, is also included, in order to se whether there is positive self-selection in to the formal sector and negative self-selection into the informal sector. The sign of the coefficient to lambda measures the extent to which mean incomes in the sector in question are lower or higher than the population mean. Therefore a negative coefficient in the formal sector indicates that on average persons employed in this sector earn more than the total population. On the other hand, for the informal sector a positive coefficient to lambda would indicate the same. In this case there is positive self-selection into each sector, ah individuals choose the sector where they do best compared to the average. If the coefficient to lambda turns out to be insignificant, it indicates that no self-selection takes place. Altitude will be included because it has been shown to have significant explanatory power for wages (Andersen, 1999). The effect of being the main provider to a family, measured as the household head, gender, and economic sector will also be included. In addition language measures, degree of illiteracy and migration will be included such as to illustrate discrimination of ethnic minorities, illiterates and new residents.

The table shows the significant results from the estimation of wages for the formal and the informal sector respectively.

For the formal sector, it can be seen that years of education is a very significant factor explaining the wages in the sector. There are large returns to education in this sector of the economy. The return to general experience is hump shaped as expected in rnost cases. On the other hand returns to experience in current job seem to be almost linearly related to earnings.

The coefficient to lambda, the variable controlling for self-selection, is significantly negative in ah rounds, indicating positive self-selection into the formal sector. This is an important finding because it indicates that the formal sector is better at attracting individuals with the best unobserved characteristics such as ability, motivation etc. compared to the informal sector.

Altitude does not only have explanatory power for the probability of employment in the formal sector, it also explains wages very well. The higher an individual hives the less he/she earns. The relation is linear as opposed to the probit equation estimated above. This means that people employed in the formal sector in the highlands earn significantly less than their colleagues in the valleys and the lowlands.

Being illiterate affects earnings significantly in only one year. This may be due to the fact that education measures approximately the same effect causing multicollinearity in the model.

Being a woman significantly increases the probability of getting employment in the formal sector, but once a female is employed, she earns significantly less than her male counterpart.

Household heads generally earn more than people who are not the main providers in a family.

Table 6: Formal Sector Estimations

	1989	1990	1991	1992	1993	1994	1995
Intercept	-	-	1,017	0,6	1,102	1,041	-
			(2,8)	(1,677)	(3,375)	(1,917)	
Years of	0,05618	0,08814	0,0963	0,07736	0,09551	0,09973	0,106
education	(17,735)	(27,973)	(16,976)	(19,536)	(23,969)	(27,726)	(41,182)
Experience	0,03777	0,02827	0,01123	0,01887	0,007796	0,02174	0,03
	(8,955)	(8,445)	(4,66)	(3,397)	(3,931)	(5,213)	(8,168)
Experience	-0,00057	-0,00039	-	-0,00026	-	-0,00021	-0,0003
squared	(-6,523)	(-5,841)		(-2,339)		(-2,577)	(-3,818)
Job experience	0,000 19	0,000037	0,00021	0,000123	-	0,000226	0,000203
	(3,846)	(2,622)	(3,821)	(2,545)		(5,123)	(4,222)
Job exp.	-	-0,00000	-	-	-	-	-
Squared		(-1,9)					
Lambda	-1,073	-1,099	-2,712	-1,420	-1,988	-2.135	-0,451
	(-11,901)	(-10,833)	(-4,236)	(-2,269)	(-3,239)	(-2,7)	(-6,221)
Altitude	-0,111	-0,17	-0,162	-0,156	-0,151	-0,124	-0,123
	(-10,222)	(-19,084)	(-7,858)	(-8,148)	(-7,628)	(4,945)	(-14,002)
Ethnicity[11]	-0,202	-0,07219	-	-	-0,129	-0,0709	-
	(-6,134)	(-2,557)			(-4,073)	(-2,72)	
Illiteracy	-0,293	-	-	-	-	-	-
	(-2,712)						
Woman	-	-	-0,241	-0,163	-0,401	-0,224	-0,136
			(-6,552)	(4,118)	(-12,667)	(-6,811)	(4,198)
Head	0,210	0,282	-	0,143	-	0,115	0,08958
	(6,436)	(11,020)		(3,772)		(3,587)	(2,623)
Public	0,138	-	-	-	0,174	-	-
Admin.	(3,736)				(3,903)		
R²	0,293	0,405	0,252	0,349	0,355	0,449	0,808
F-statistic	153,856	388,388	97,686	112,922	25,946	287,341	1564,501
H₀ : All coef = 0	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)
N	2527	5364	5103	4888	3654	6978	6799

Source: Author's calculations; tic significance level is 10%.

Employment in the public sector increases earnings in some cases and when the variable is significant it is positive. However the result does not seem particularly robust, therefore no firm conclusion can be reached here. The choice of sector seems not to be of importance for earnings in the formal sector which is opposite of the findings for the informal sector as will be illustrated next.

Another result that emerges is that migration is not significantly influencing earnings in the formal sector. The reason may be that migration influences the sector choice, but once employment in a sector is obtained, whether an individual has lived in the area for a long

---

[11] This variables cannot be constructed iii 1991, 1992 and 1995 which is why tic variable is not included in these years.

time or not is no longer of importance when explaining the wage setting process between the two sectors.

Table 7: Informal Sector Estimations

	1989	1990	1991	1992	1993	1994	1995
Intercept	-	-	-	-	-	2,41	1,117
						(3,289)	(1,41)
Years of	0,03 172	0,08759	0,09425	0,08117	0,105	0,0574	0,05737
education	(4,243)	(23,594)	(25,352)	(25,872)	(31,208)	(9,792)	(8,534)
Experience	0,02899	0,02529	0,03692	0,02865	0,0408	0,01539	0,0049
	(4,064)	(6,578)	(9,260)	(7,401)	(11,028)	(2,686)	(1,74)
Experience	-0,00038	-0,00037	-0,00055	-0,0004	-0,00059	-0,00028	-
squared	(-2,399)	(-4,518)	(-6,408)	(-4,402)	(-7,211)	(-2,808)	
Job experience	0,001	0,000128	0,000696	0,000894	0,00435	0,0006	0,00049
	(3,059)	(6,285)	(4,814)	(5,449)	(4,647)	(3,429)	(4,576)
Job exp.	-0,00000	-0,00000	-0,00000	-	-0,00001	-0,00000	-
Squared	(-2,13)	(-4,545)	(-2,997)		(-3,091)	(-2,047)	
Lambda	-0,467	-1,883	-1,052	-0,491	-1,093	-3,432	-1,09
	(-2,903)	(-22,418)	(-11,188)	(-5,243)	(-11,029)	(-3,310)	(-0,97)
Altitude	-0,144	-0,16	-0,153	-0,176	-0,127	-0,232	-0,166
	(-7,734)	(-18,22)	(-17,092)	(-19,378)	(-12,095)	(-7,05)	(-4,566)
Ethnicity	-	-	-	-	-0,103	-	-
					(-3,189)		
Woman	-0,237	0,154	-0,172	-0,147	-0,101	-0,24	-
	(-2,599)	(4,285)	(-4,289)	(-3,987)	(-2,734)	(-4,825)	
Head	-	0,149	0,08737	0,231	0,08525	0,08205	0,157
		(4,174)	(2,273)	(6,782)	(2,24)	(1,992)	(3,424)
Migrant	-	-	0,165	-	0,132	-	-
			(2,110)		(3,024)		
Trade	0,279	-	-0,224	-0,211	-0,133	-0,112	-0,129
	(2,219)		(-5,399)	(-5,070)	(-3,322)	(-2,554)	(-2,298)
Manufacturing	-	0,153	-0,154	-0,166	-0,116	-	-0,181
		(4,052)	(-4,4)	(-4,736)	(-3,224)		(-2,843)
Utility	-	0,474	0,255	-	-	-	-
		(2,746)	(1,958)				
Construction	-	0,267	-	0,138	0,125	0,146	-
		(6,082)		(3,269)	(2,853)	(3,941)	
Hotel	-	-	-0,33	-	-0,329	-	-
			(-4,8 14)		(-2,755)		
Agriculture	-	-	-0,472	-0,176	-0,271	-	-
			(-5,041)	(-1,923)	(-2,772)		
Transport	-	-	-0,177	-	0,316	-	0,781
			(-3,948)		(3,236)		(4,096)
Mining	-	-	-	-	0,197	0,335	-
					(2,051)	(3,783)	
R^2	0,194	0,309	0,437	0,545	0,355	0,264	0,202
F-statistic	13,310	115,972	131,591	230,974	235,302	48,731	35,111
H_0 : All coef = 0	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)	(0,000)
N	3031	3217	3429	3077	2840	2161	1942

Source: Author's calculations; the significance level is 10%.

In the informal sector, the picture is more varied. The returns to education are still very significantly positive, but in most cases less than for individuals employed in the formal sector.[12] This means that for people employed in the informal sector, one additional year of schooling would contribute significantly to earnings but less so than in the formal sector.

---

[12] The average return to schooling in tic formal sector is 0,08846 whereas it is only 0,0735 in tic informal sector.

Returns to general experience are hump shaped as in the formal sector, which means that previous experience increases earnings to a certain point beyond which experience decreases earnings. Ibis is probably related to the fact the elderly people are less productive than younger ones who have been working for some years. The returns to experience in current job are also of a hump shaped character, even more so than in the formal sector. Ibis means that earnings in the formal sector may be generated by tenure rather than qualifications, whereas tenure is less important in the informal sector and productivity declines after a certain time period in the same job.

The coefficient to the selection variable, lambda, is negative, which points to negative self-selection into the informal sector, because the formal sector was the reference category when estimating the probability of getting employment in each sector. Whereas the formal sector is best at attracting individuals with favourable unobserved characteristics, the informal sector has to employ whoever is left. This finding suggests that formal sector workers would earn more regardless of sector choice, implying that the same son of abilities are rewarded in both sectors.

Altitude is important for earnings; the higher an individual lives, the less he/she earns. On average the coefficients to altitude is higher for the informal sector, than for its formal counterpart.[13] Ibis means that earnings are more sensitive to altitude in the informal sector than in the formal sector and could be an indication of productivity being relatively more important for wages in the informal sector. In the formal sector on the other hand, there are indications that productivity is not the main determinant of earnings since returns to experience in current job is not hump shaped and marginal productivity is usually falling.

Ethnicity does not influence earnings as much as in the formal sector, another indication of earnings being determined by productivity in this sector to a higher degree than in the formal sector, assuming that indigenous people are as productive as non-indigenous. Thus, discrimination does not seerm to be a problem in the informal sector maybe because productivity is more important than other individual characteristics.

Being female still decreases earnings significantly, indicating discrimination in this case if it is assumed that women have the same productivity as men. This is not the case, however, if it is taken into account that informally employed women often fulfill other duties such as childcare while working. Poor families tend to get more children than non-poor (Wood and Patrinos, 1994), and since informal workers earn less than their formal sector counterparts they are likely to get more children too. Therefore this finding also suggests payment according to productivity.

Household heads earn more than other people employed in the informal sector.

Being a migrant does in some cases influence earnings, and in this case positively. This means that being a migrant is not only important for sector choice; when a migrant is employed in the informal sector he/she is rewarded with a higher wage than the average employee. The result should be drawn with caution however since the variable only turns out significant in two cases. Nevertheless this is more than in the case of the formal sector, and again productivity may be the explanation.

Informal earnings generally seem to be more sensitive to choice of line of business than formal sector earnings. Employment in trade, manufacturing, the hotel sector or agriculture significantly decreases earnings in some cases. On the other hand employment in

---

[13] The simple average is -0,142 for the formal sector and -0,165 for the informal sector.

utility, construction, mining and transport to a certain degree, increases informal sector earnings.

It seems that the equation has more explanatory power for the formal sector than for the informal sector judged by the higher R^2 in the former.[14]

So far the analysis has assumed that returns to schooling were related to earnings in a log-linear way. This assumption may show up to be too restrictive, however, since different levels of schooling may have different returns. This is a result of some levels of schooling being more efficient in raising an individual's marginal product than other levels. The next section elaborates this possibility in depth by allowing for different levels of schooling to have different returns.

4.3   Earnings Equations with Non-Linear Returns to Schooling

The non-linear returns to schooling are modeled with dummy variables indicating the highest level attained. Hence there is a dummy variable for each level of education rather than one variable measuring years of education. The models from the previous section are applied again, adjusted for these new educational variables. The reason for keeping the two sectors separated in this section as well is that different sectors may give different returns due to several factors: Imperfect markets, barriers to entry in a sector, or a specific level of education may be irrelevant in a particular sector. As is demonstrated in the previous section there are rather large differences between the two sectors when it comes to the wage-setting process and the factors influencing wages.

The regressions have no schooling as the reference category, and the main results are shown in table 8. The rest of the regression results do not differ much from the results presented above and are not reported here.

Table 8: Non-Linear Returns to Education, Formal Sector

	1989	1990	1991	1992	1993	1994	1995
None	--	-	-	-	-	-	-
Primary	-0,605	0,07694	-0,677	-0,65	-0,951	-0,492	-0,791
	(-7,042)	(0,43)	(-6,716))	(-7,425)	(-8,947)	(-8,027)	(-11,50)
Intermediate	-0,475	0,313	-0,629	-0,555	-0,723	-0,42	-0,655
	(-5,757)	(1,766)	(-6,687)	(-6,586)	(-6,863)	(-7,126)	(-10,41)
Mediate	-0,231	0,572	-0,379	-0,329	-0,498	-0,238	-0,345
	(-3,316)	(3,36)	(4,42)	(4,58)	(-4,965)	(4,687)	(-6,485)
Normal	0,125	0,855	0,103	0,09507	0,09461	0,29	0,368
	(1,697)	(5,092)	(1,134)	(1,224)	(0,866)	(4,84)	(6,284)
University	0,47	1,362	0,407	0,449	0,41	0,605	0,8
	(7,05)	(8,154)	(4,627)	(5,998)	(4,037)	(11,239)	(15,031)
Technical	0,101	-	0,01268	0,121	-0,115	-	-
	(1,292)		(0,129)	(1,368)	(-1,032)		
Technical Mediate	-	0,797 (4,474)	-	-	-	0,18 (2,592)	0,361 (4,672)
Technical Superior	-	1,045 (5,864)	-	-	-	0,338 (4,30)	0,512 (5,922)

Source: Authors' calculations, sample sizes as in table 6

---

[14]   There are significant differences across gender and ethnicity in the earnings regressions. Further analysis of these differences shows that only deviations in the numencal value of the coefficients differ across groups, not signs, and/or relative importance.

It is seen that returns to education are not linear in years of schooling.[15] The liberalization largely ignores the negative returns for the low levels of education.

In the formal sector the returns to primary, intermediate and mediate levels of education are largely negative. This means that persons possessing relatively low levels of education earn less than those with no schooling at all. Individuals who have obtained a higher level of education on the other hand are secured positive returns according to these estimates, especially those having a university degree.

In the informal sector primary and intermediate education gives the same return (or less) than no schooling. Thus, there is no incentive to obtain low levels of education in either sector. Higher levels of education tend to yield (slightly) positive returns in the informal sector as well.

Basic school skills are in general not rewarded in either labour market. This can be either because the demand for skills is low or because the accumulated skills are not useful in the labour market. The relatively large differences in returns across levels of education indicate that the linear specification is not appropriate for measuring returns to education. Additionally it is worth noticing the low/negative returns to basic education. Ibis seriously questions the quality of schooling in Bolivia; the results indicate that the value of additional education is very low, in many cases even negative. For a country with relatively large problems with poverty, this is an area worth further investigation and investment. It is known that education decreases poverty, but if it is not worth going to school, people may choose working instead, thus not reducing poverty in the long run.

Table 9: Non-Linear Returns to Education, Informal Sector

	1989	1990	1991	1992	1993	1994	1995
None	--	-	-	-	-	-	-
Primary	-0,0315	-0,222	0,102	0,101	-0,769	-0,0075	-0,214
	(-0,217)	(-1,246)	(0,832)	(0,755)	(-6,231)	(-0,102)	(-2,24)
Intermediate	0,008597	-0,01	0,267	0,153	-0,699	0,02935	-0,071
	(0,6)	(-0,58)	(2,192)	(1,156)	(-5,833)	(0,394)	(-0,74)
Mediate	0,15	0,273	0, 478	0,414	-0,531	0,124	-0,041
	(1,151)	(1,643)	(4,121)	(3,286)	(4,617)	(1,687)	(-0,42)
Normal	-0,439	0,778	1,13	0,821	-0,0268	0,319	0, 07467
	(0,971)	(4,265)	(7,436)	(4,819)	(-0,16)	(1,269)	(0 334)
University	0,71	0,936	1,232	1,232	0,279	0,745	0,553
	(4,18)	(5,688)	(10,557)	(9,941)	(2,359)	(7,991)	(4,683)
Technical	0,233	-	1,042	1,017	0,156	-	-
	(1,215)		(8,45)	(7,719)	(1,274)		
Technical Mediate	-	0,505	-	-	-	0,253	0,07214
		(2,863)				(1,741)	(0,416)
Technical Superior	-	0,519	-	-	-	0,798	0,579
		(2,825)				(4,121)	(2,575)

Source: Authors' calculations, sample sizes as in table 7

4.4   Oaxaca Decomposition

The wage gap between the two sectors can be divided between differences in characteristics and differences in rewards to these characteristics. An Oaxaca decomposition makes this distinction possible (Oaxaca, 1973). Assuming that all workers would be paid like informal

---

[15]   One could also have searched for degree effect, i.e. if there are any differences between individuals who have completed a certain degree and others who have not. However this is outside tic scope of this paper.

sector workers on average, it is possible to construct a decomposition between how much of the wage differential is due to differences in characteristics and how much is due to differences in rewards on the two markets. Letting W_i be the informal sector wage and W_f be the formal sector wage, the following decomposition makes it possible to differentiate between the factors mentioned before:

$$Ln(W_f) - Ln(W_i) = \Delta \overline{Z}`\hat{\beta}_i - \overline{Z}_f`\Delta\hat{\beta},$$

where $\overline{Z}$ is a vector of average characteristics from each sample, A denotes differences, and $\hat{\beta}$ is a vector of estimated coefficients from the earnings equations in section 4.2. The first pan of the right hand side of the equation is thus the share of the wage gap that can be ascribed to differences in characteristics, whereas the second pan is due to differences in rewards. Therefore the interesting pan of this analysis is the differences in rewards; if the two sectors reward the same characteristics differently, it indicates that there are different wage setting mechanisms at the two markets. The results are as follows for round 1 and 7.[16]

Table 10: Oaxaca Decomposition, Round 1, 1989

	Average Characteristics		Decomposition	
	Formal	Informal	Differences in characteristics	Differences in rewards
Intercept	-	-	0	0
Years of education	10,24	6,61	0,115	0,2505
Experience	18,69	22,28	-0,104	0,164
Experience Squared	495,32	706,02	0,081	-0,091
Job Experience	287,69	115,13	0,173	0,233
Job Ex . Squared	198825	53464,7	-0,107	0,146
Lambda	0,46	0,51	0,023	-0,278
Altitude	2,46	2,02	-0,063	0,08 1
Woman	0,31	0,52	0,05	0,072
Head	0,61	0,46	0	0,127
Ethnicity	0,38	0,52	0	-0,077
Public Admin.	0,17	0	0	0,023
Analphabetic	0,02	0,11	0	-0,006
Trade	0,06	0,34	-0,079	-0,016
Ln (monthly wage)	5,52	5,19		
Difference		0,34	0,0898	0,1634

Note: Job experience is measured in weeks.
Source: Author's calculations

It is seen that for round 1, approximately 27% of the wage differential can be accounted for by differences in characteristics, whereas differences in rewards account for approximately 49% of the wage gap.[17] The differences in rewards and characteristics are primarily explained by returns to education and experience. This means that in the informal sector workers tend to have worse characteristics than workers employed in the formal sector. In addition formal sector workers are rewarded differently than informal sector workers after having controlled for differences in characteristics (the fourth column in table

---

[16] These two rounds were chosen because they represent different time periods and include approximately tic same variables. Especially it was judged that tic variable ethnicity, which is included in both samples, would make the comparison more appropriate between rounds 1 and 7 instead of rounds 1 and 8.

[17] There is also an unexplained part of the wage differential due to the fact that the earnings equations do not explain earnings perfectly, i.e. $R^2 < 1$ in tie equations. This unexplained part amounts to 24% in round 1 and 17% in round 7.

9), they are rewarded more for education and general experience and less for experience in current job, which is another conclusion reinforcing the results from the analysis above.

Round 7 shows slightly different results.

In 1994, *54%* of the wage differential can be accounted for by differences in characteristics, whereas differences in rewards only account for 29% of the wage gap. Differences in characteristics mainly consist of formal sector workers having more education and experience, both in general and job specific. Rewards in the two sectors depend primarily on education and general experience, self-selection and altitude. However, it is not possible to conclude that the decline in the pan of the wage differential, which is due to differences in rewards, is a phenomenon that will continue over time due to the large differences across rounds, which was illustrated above. The main conclusion in this section is that the results from the decomposition confirm the previous analysis, which contributes to the robustness of the results.

Table 11: Oaxaca Decomposition, Round 7, **1994**

| | Average Characteristics | | Decomposition | |
	Formal	Informal	Differences in characteristics	Differences in rewards
Intercept	-	-	0	-1,369
Years of education	10,46	7,82	0,152	0,443
Experience	20,49	14,96	0,085	0,130
Experience Squared	597,64	375,35	-0,062	0,040
Job Experience	257,24	91,85	0,100	-0,098
Job Ex . Squared	154172	39040	-0,03 5	0,047
Lambda	0,52	0,55	0,105	0,678
Altitude	2,57	2,39	-0,042	0,278
Woman	0,43	0,34	-0,02 1	0,007
Head	0,56	0,43	0,010	0,018
Ethnicity	0,42	0,40	0	-0,030
Public Admin.	0,05	0,27	-0,03 2	-0,007
Analphabetic	0,11	0,14	0,004	0,012
Trade	0,02	0,03	-0,003	-0,007
Ln (monthly wage)	6,22	5,73		
Difference		0,485	0,260	0,142

Note: Job experience is measured in weeks.
Source: Author's calculations

5. Conclusion

As in other studies the wage differential between the formal and the informal sector has been found to be quite large in urban Bolivia. This wage gap has been estimated by a human capital model correcting for selectivity bias using Heckman's two-step procedure. The probability of entering the formal sector was estimated, and the most surprising result was that migrants tend not be employed in this sector. Another noticeable result was that the probability of entering the formal sector was a hump shaped function of altitude.

Wages were then estimated using a Mincerian specification, and there were some differences between the formal and informal labour markets. Returns to education are generally higher for formal sector worker, whereas job specific experience did not exhibit the usual hump shape for these workers. Instead wages depend linearly on experience in current job, indicating that tenure is more important than productivity. Being female or indigenous decreases formal sector earnings, and altitude turned out to be a significant explanation of

wages as well. People living in the valleys and lowlands earn more in general than their counterparts in the highlands. Sector choice turned out to be less significant in the formal sector than in the informal sector.

In the non-linear specification of returns to education, it turns out that returns to schooling differ significantly between levels. Basic education is either not rewarded in the labour market at ah, or the returns are negative. Only returns for relatively high degrees are significantly positive, indicating that the quality of basic education in Bolivia is very low.

A decomposition of the wage differential into differences in characteristics and differences in rewards between the two sectors only confirm the first analysis, because education and experience are the main factors contributing to the differences between sectors.

References

Andersen, Lykke E. and Zafiris Tzannatos (1998), "Econometric Analysis of Formal and Informal Labor Markets in Thailand 1988-1996", *Working Paper, World Bank.*

Andersen, Lykke E. (1999), "Wage Differentials Between Bolivian Cities", *Working Paper 02/99, HSEC-UCB,* La Paz, Bolivia.

Greene, William H. (1997), *"Econometric Analysis",* Prentice Hall, Inc., New Jersey, USA.

Heckman, James J. (1976), "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models", *Annals of Economic and Social Measurement 5/4.*

Heckman, James J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica vol. 47, no. 1.*

1NE (Instituto Nacional de Estadística) *(1989-1995), "Encuesta Integrada des Hogares",* La Paz, Bolivia.

Jensen et al (1997), "Employment and Earnings in Zambia in 1993", *Working Paper, Centre for Labour Market and Social Research,* University of Aarhus, Denmark.

Mincer, J. (1974), "Schooling, Experience and Earnings", NBER.

Oaxaca, Ronald (1973), "Male-Female Wage Differentials in Urban Labor Markets", *International Economic Review, vol. 14 no. 3.*

Wood and Patrinos (1994), "Urban Bolivia", in *"Indigenous People and Poverty in Latin America, an Empirical Analysis",* (1994), edited by Psacharopoulos and Patrinos, World Bank Regional and Sectoral Studies.