

FORMAL MEASURES OF THE INFORMAL SECTOR WAGE GAP  
IN MEXICO, EL SALVADOR, AND PERU

March 28, 1995

Douglas Marcouiller  
Department of Economics, Boston College

Veronica Ruiz de Castilla  
Department of Economics, University of Texas at Austin

Christopher Woodruff  
Graduate School of International Relations and Pacific Studies  
University of California, San Diego

ABSTRACT: Using comparable micro-level data from three countries, we ask what type of person works in the informal sector and whether informal workers earn lower wages than observationally equivalent workers in the formal sector. The characteristics of informal workers are similar across countries. Surprisingly, when we control for these personal characteristics, we find a significant wage premium associated with formal employment in El Salvador and Peru but a premium associated with work in the *informal* sector in Mexico. A model of endogenous selection offers little help in explaining the differences in wage patterns. The research casts doubt on the received wisdom that the informal sector, always and everywhere, is a poorly-paid but easily-entered refuge for those who have no other employment opportunities.

We would like to thank Daniel T. Slesnick for encouraging this project. We are also grateful to those who facilitated access to the data: Agustín Escobar at CIESAS de Occidente in Mexico, Mauricio Alens of the Ministerio de Planificación of El Salvador, and the Poverty Analysis and Policy Division of the Population and Human Resources Department of the World Bank.

At the end of a meeting in Ciudad Nezahualcóyotl near Mexico City, several microentrepreneurs were asked why they chose informal employment. One answered, "When I sew uniforms at home, I can also watch over the kids and my house." A second said, "I hate having to answer to a boss; I like making my own decisions." A third said, "I make more money in my shop than I would in a factory."<sup>1</sup> Other workers, of course, might be informal simply because they could not find formal jobs.

The diversity of experiences of economic informality is reflected in the economic literature, parts of which emphasize the choices which workers make while other parts underscore the constraints on choice.<sup>2</sup> Some allege that the "modern" sector is incapable of generating sufficient employment. The small firms of the informal sector then offer second-best opportunities in easily-entered, competitive markets (Tokman, 1989). Others see the informal sector as that part of the economy which is "unregulated by the institutions of society, in a legal and social environment in which similar activities are regulated" (Castells and Portes, 1989, p.12). For better (de Soto, 1989) or for worse (Roberts, 1990), people freely shift operations between the formal sector, where licensing is enforced and worker entitlements are protected, and the unregulated informal sector (Marcouiller and Young, 1995).

The first of these strands of thought intertwines naturally with the analysis of labor market segmentation. Theories of segmentation generally make two assertions. The first is that "rewards in different economic sectors may differ for workers of equal potential productivity" (Magnac, 1991, p.165), or, more specifically, "that there is a distinct low-wage (secondary) labor market in which there are no returns to schooling and workers do not receive on-the-job training" (Dickens and Lang, 1985, p.792). The second assertion is that, "because of institutional barriers to occupational mobility between sectors, a worker in the lower sector has

---

<sup>1</sup> Interview by Douglas Marcouiller, July, 1992.

<sup>2</sup> The informal sector literature is usually traced to ILO (1972) and Hart (1973). Contemporary surveys of the extensive literature can be found in Lubell (1991), Pérez Sáinz (1991), Tokman (1992), and Rakowski (1994). Early models of wage differentials are found in Harris and Todaro (1970), Fields (1975), and Mazumdar (1983).

less than full access to a job in the upper sector held by an observationally identical worker” (Gindling, 1991, p.585). The image of workers queued for high-wage formal jobs fits this frame.

What do the data show? Using individual- and household-level data from surveys of urban areas in three different countries, we ask what type of person works in the informal sector and whether informal workers earn lower wages than observationally equivalent workers in the formal sector. We find significant wage premia associated with work in the formal sector in El Salvador and Peru. In Mexico, on the other hand, a premium is associated with work in the *informal* sector. Our paper does not formally test the segmentation hypothesis.<sup>3</sup> However, the evidence we offer on wage differentials does challenge the widespread notion that informal employment is, by its nature, the last resort of those who have no other choice.

Papers using recent large-scale household surveys from different countries in a completely consistent approach to the analysis of the informal sector are few and far between. Our work is close in spirit to Gindling’s analysis of informal, private formal, and public labor markets in Costa Rica (1991). Terrell (1989) estimates wage regressions for Guatemala City. Tannen (1991) and Telles (1993) study the wage structures of parts of Brazil. Heckman and Hotz (1986) explore a more general notion of primary and secondary labor markets in Panama. A recent working paper by Funkhouser (1994) presents interesting comparative analysis of household data from the five Spanish-speaking Central American countries.

The first section of this paper describes our data, our definitions, and the characteristics of workers and jobs in the formal and informal sectors. In the second section we estimate separate wage regressions for men and women in each sector, finding that the wage structures differ significantly, that statistically significant wage premia are associated with formal employment in Peru and El Salvador, and, unexpectedly, that a premium is associated with informal work in Mexico. In the third section, we examine the possible impact of endogenous sample selection on the wage gaps. The fourth section summarizes and concludes.

---

<sup>3</sup> The complexity of such tests is discussed in Heckman and Hotz (1986) and Magnac (1991).

## Section 1. Overview, Data and Definitions

We use data from three large government-sponsored household surveys.<sup>4</sup> Six thousand households were surveyed in El Salvador in 1990, eleven thousand in Mexico in 1990, and five thousand in Peru in 1985.<sup>5</sup>

Two different operational definitions of the informal sector are common in empirical work. Analysts who emphasize the generation of employment generally focus on firm size as the defining criterion (Mezzera 1990). Those who emphasize social regulation focus on coverage by the social security system (Roberts, 1990). We have worked with both definitions. Our “firm size criterion” counts as informal those workers who are self-employed or who operate or are employed in firms with five or fewer employees (four or fewer in the case of El Salvador), with the exception of professionals,<sup>6</sup> who are considered to be formal. Our “benefits criterion” or “social security criterion” counts as formal those who receive social security coverage or pay social security taxes as a condition of employment, regardless of occupational status. Following common practice, we define domestic service as a separate sector.<sup>7</sup> We also must exclude from the wage regressions those people who work in family enterprises without pay.

---

<sup>4</sup> Mexican data from the Encuesta Nacional de Empleo Urbano, 3rd quarter, 1990 (INEGI, 1990). For El Salvador we use the Encuesta de Hogares de Propósitos Múltiples (MIPLAN, 1990). Peruvian data are from the Encuesta Nacional de Hogares sobre Medición de Niveles de Vida, 1985-1986 (World Bank, INE, 1986).

<sup>5</sup> The Salvadoran and Peruvian data cover all urban areas. The Mexican data cover Mexico City, Guadalajara, and Monterrey which, according to the 1990 census, account for 33% of Mexico's urban population. All available data are used in the analysis presented in the body of this paper, which was repeated for metropolitan San Salvador and Lima alone with qualitatively similar results.

<sup>6</sup> Occupational groups 0/1 and 2 of the 1968 International Standard Classification of Occupations and corresponding categories in the Mexican classification scheme.

<sup>7</sup> The rationale for excluding paid domestic servants is that they receive an unknown portion of their pay in kind (Tokman, 1989, p. 1069; see also Telles, 1993, p.234).

Each of the economies has a large informal sector. Even excluding domestic servants and unpaid family workers, roughly half the urban workforce is informal in Peru and El Salvador, and Mexico's informal sector is only slightly smaller (Table 1).<sup>8</sup> The size criterion is the more restrictive of the two. Almost all of those who are informal by size (92%, 97%, and 89% in Mexico, El Salvador, and Peru, respectively) also lack benefits; the converse is not true.

Table 1. URBAN LABOR FORCE BY SECTOR

	<u>El Salvador</u>	<u>Peru</u>	<u>Mexico</u>
<u>By Size of Firm:</u>			
Formal	52.1%	42.5%	69.2%
Informal	47.9%	57.5%	30.8%
<u>By Social Security Coverage</u>			
Formal	37.2%	38.2%	56.8%
Informal	62.8%	61.8%	43.2%

The gender composition of the paid labor force of each country is shown in Table 2, along with the percentage of male and female workers who are informal. In El Salvador and Peru, but not in Mexico, women are more likely than men to be informal.<sup>9</sup>

Table 2. GENDER AND THE INFORMAL WORKFORCE

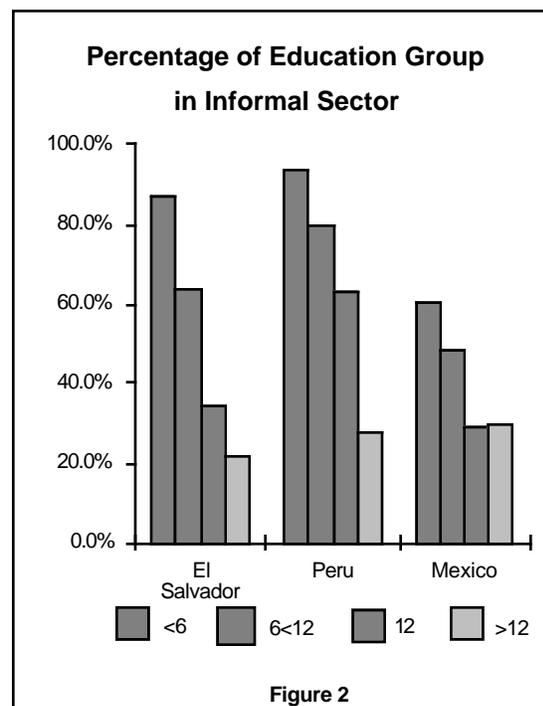
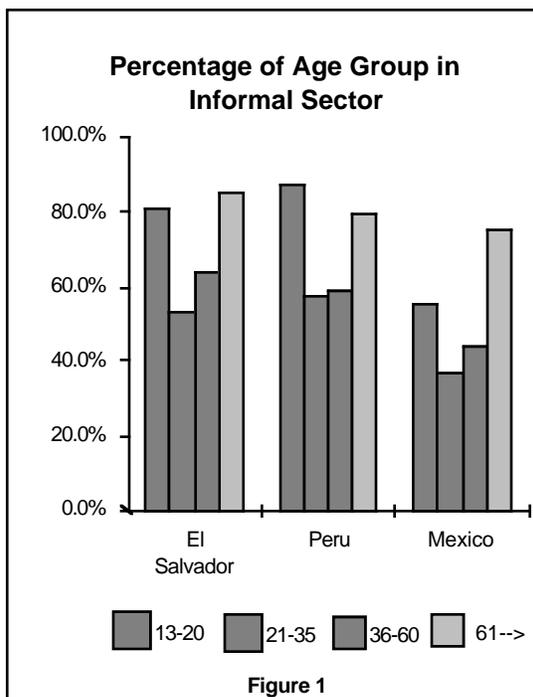
	<u>El Salvador</u>	<u>Peru</u>	<u>Mexico</u>
<u>Proportion of Paid Labor Force:</u>			
Males	59.9%	65.2%	69.9%
Females	40.1%	34.8%	30.1%
<u>Percent Informal by Size of Firm:</u>			
Males	41.2%	53.0%	33.4%
Females	57.8%	66.0%	24.8%
<u>Percent Informal by Benefits:</u>			
Males	60.0%	58.0%	46.3%
Females	67.0%	69.0%	36.1%

<sup>8</sup> Observations for Mexico and El Salvador are weighted by their survey expansion factors in Tables 1 and 2 and Figures 1 through 4. The size of the Mexican informal sector in Table 1 may be understated because the information is limited to the three largest cities. Roberts (1990) notes that in an earlier survey in Mexico, "the lowest percentage of informal employment is found in Mexico City...[while the] highest percentage of informality is reported for the non-metropolitan areas." (Roberts 1990, p. 6).

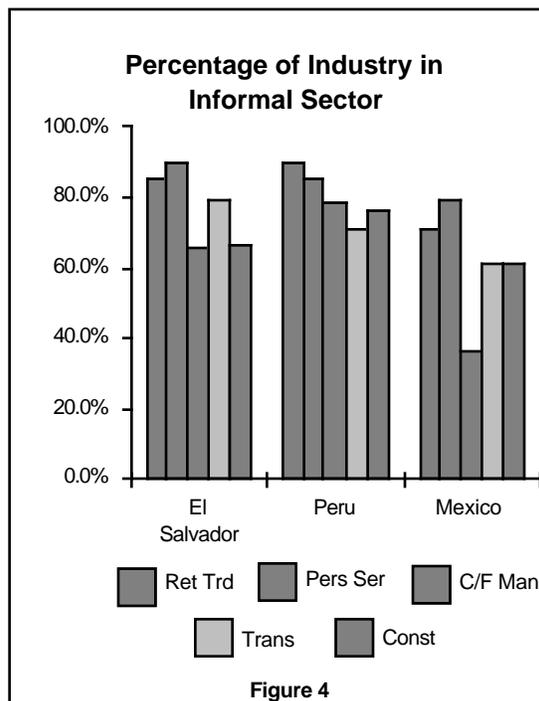
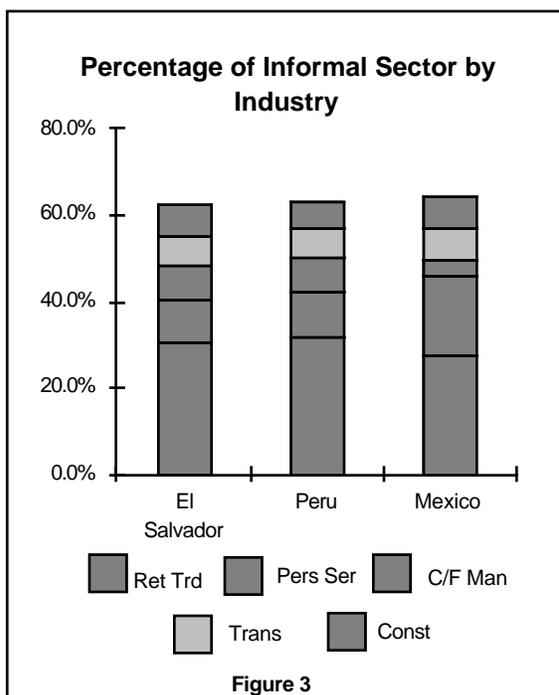
<sup>9</sup> Neither the lower female participation nor the tendency of females to be formal in Mexico is explained by the fact that the Mexican data do not include smaller urban areas. The 1990 population census indicates that females are a lower percentage of the non-agricultural labor force in Mexico as a whole than they are in the three cities we use. Selby et al (1990) find that females are less likely to be informal than their male counterparts using data from ten smaller cities in Mexico.

There are marked differences between sectors in the age and schooling of workers. The differences are displayed in Figures 1 through 4, using the social security criterion of informality. In Figure 1, we divide all paid workers into four age groups: 13 through 20 years of age, 21 through 35, 36 through 60, and above 60. The oldest and the youngest workers are especially likely to be informal. However, the different age groups have different roles within the informal sector. The youngest workers are likely to be wage workers in small firms, while older workers are more likely to be self-employed or the owners of small businesses. In El Salvador, for example, 15% of informal sector workers aged 13-20 are owners or self-employed, compared to 46% of informal sector workers aged 21 through 35 and 72% of informal workers older than 35 years.

In Figure 2, we show the distribution of schooling using four educational groups--less than six years of primary school, at least six years but less than complete secondary school, complete secondary, and some post-secondary schooling. The propensity to be informal declines with education in all three countries.



The sectors differ not only in the personal characteristics of the workers but also in the characteristics of jobs--particularly in the industrial composition of employment and in the hours distribution. Five two-digit industrial classification groups together account for roughly two-thirds of informal employment in each country: retail trade, construction, clothing and footwear manufacturing, transportation and personal services. Figure 3 shows the distribution of the informal work force across the five categories, and Figure 4 shows the proportion of the workers in each industry who are informal. Under the social security criterion, 63% of employment is informal in El Salvador, 62% in Peru, and 43% in Mexico. Informal employment is more common than average in retail trade, construction, transportation and personal services in each of the three countries. Only the propensity of clothing and footwear manufacturing to be informal varies notably across the countries.



Finally, we note that the distribution of hours worked per week differs across sectors. The tails of the hours distribution are fatter in the informal sector than in the formal. In Peru,

23% (7%) of informal women (men) work less than 20 hours, compared with only 8% (4%) of formal women (men). Likewise, 14% (17%) of informal women (men) report working more than 65 hours per week, compared to only 2% (8%) of formal women (men). The standard deviation of mean weekly hours is 10.6 for Mexican formal workers but 15.9 for Mexican informals, 12.2 for Salvadoran formal workers but 19.9 for informals, 15.7 for Peruvian formal workers but 21.0 for informals.

## Section 2. Sectoral Wage Differentials

Controlling for gender, age, and schooling, do the returns to labor differ systematically between the formal and informal sectors for observationally identical workers?<sup>10</sup> Judging by a Chow test for differences between estimated standard earnings functions, we find in each country that the process of wage determination does differ significantly between the sectors. We find statistically significant formal wage premia in El Salvador and Peru. However, we find in the Mexican case that it is the *informal* workers who enjoy the unexplained premium.

The form of our earnings function is quite standard. For male workers, we regress the log of the hourly wage<sup>11</sup> on years of schooling, a dummy for completion of secondary school, experience and experience-squared (as represented by the proxy age-schooling-6), six industry dummies, and a dummy for residence in San Salvador, Lima, Guadalajara, or Monterrey. Since the variable, age-schooling-6, is likely to be less adequate a proxy for the experience of women workers, for women we substitute age and age-squared for the experience variables. The means of these independent variables are presented in Table 3.<sup>12</sup>

Earnings functions were estimated separately for men and for women in each of the sectors under each of the operational definitions of informality. Results under the social security criterion are presented here; results by firm size are discussed in the Appendix. The

---

<sup>10</sup> For the moment we ignore the endogenous sample selection.

<sup>11</sup> Calculated as  $(12 \times \text{monthly earnings}) / (52 \times \text{weekly hours})$ .

<sup>12</sup> Our data do not include job tenure. Domestic servants and unpaid workers have been excluded from the regressions.

coefficients for both men and women have the expected signs (Tables 4 and 5). The returns to schooling are substantial in both sectors. Although the high returns to schooling in both sectors do not alone disprove the segmentation hypothesis, high returns in the informal sector do run counter to the description of a secondary sector given by Dickens and Lang (1985).<sup>13</sup>

Table 3. MEANS OF INDEPENDENT VARIABLES

	FEMALE WORKERS					
	El Salvador		Peru		Mexico	
	Formal	Informal	Formal	Informal	Formal	Informal
Years Schooling	11.64	4.85	13.46	8.51	10.82	8.76
High School Completed	.673	.107	.948	.481	.568	.307
Age	33.3	39.8	33.5	35.5	29.7	34.4
Agricultural Ind.	.01	.04	.002	.04	.001	.004
Manufacturing	.20	.25	.13	.18	.28	.18
Construction	.01	.002	.00	.00	.01	.006
Commerce	.15	.57	.12	.55	.19	.53
Trans & Comm	.02	.005	.03	.01	.03	.01
Services	.54	.12	.62	.21	.42	.25
Sample Size	942	2255	465	981	2645	1393
	MALE WORKERS					
	El Salvador		Peru		Mexico	
	Formal	Informal	Formal	Informal	Formal	Informal
Years Schooling	9.61	5.50	12.02	9.52	9.92	8.44
High School Completed	.418	.105	.778	.627	.356	.222
Experience	20.8	26.3	20.3	20.4	17.6	20.9
Agricultural Ind.	.02	.20	.02	.08	.002	.012
Manufacturing	.20	.23	.20	.20	.38	.16
Construction	.09	.13	.04	.10	.06	.11
Commerce	.13	.18	.13	.30	.16	.32
Trans & Comm	.07	.10	.08	.11	.06	.09
Services	.40	.15	.38	.19	.28	.30
Sample Size	1738	3056	1176	1604	5580	4378

<sup>13</sup> We are not the first to find high returns to schooling in the informal sector. For Peru see Moock, Musgrove and Stelcner (1990) and King (1990). For Central America see Funkhouser (1994).

Table 4. FORMAL / INFORMAL SECTOR WAGE EQUATIONS  
FEMALE WORKERS

	<u>FORMAL</u>			<u>INFORMAL</u>		
	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>
Years Schooling	.0646 (11.06)	.0543 (3.95)	.0787 (16.58)	.0413 (5.92)	.0821 (4.65)	.0496 (5.31)
H.S. Diploma	.0299 (.55)	.1697 (.95)	.0404 (1.23)	.2534 (2.99)	.0094 (.08)	.1144 (1.48)
Age	.0389 (4.40)	.0801 (3.97)	.0430 (8.53)	.0421 (7.02)	.0916 (7.27)	.0577 (8.28)
Age Squared	-.00033 (2.86)	-.00084 (3.20)	-.00040 (5.71)	-.00043 (6.51)	-.00091 (5.98)	-.00062 (7.33)
Urban 1	.0620 (2.11)	.1219 (1.76)	.0074 (.29)	.1151 (2.85)	.0062 (.09)	.0920 (1.80)
Urban 2			.0524 (2.32)			.1831 (3.48)
Agricultural Ind.	-.3477 (2.34)	-.4011 (.58)	-.0815 (.26)	-.0804 (.30)	-.9883 (3.47)	.2971 (.84)
Manufacturing	-.2020 (3.09)	-.1700 (1.19)	-.2229 (4.94)	-.1338 (.52)	-.1317 (.54)	-.1127 (.73)
Construction	-.2178 (1.42)		-.1262 (1.34)	-.1878 (.41)		.3500 (1.17)
Commerce	-.2984 (4.59)	-.4319 (3.06)	-.2954 (6.39)	-.1828 (.71)	.1631 (.69)	-.0900 (.60)
Transport & Comm.	.2040 (1.65)	.2066 (1.02)	-.0485 (.75)	.6793 (1.89)	.5819 (1.31)	.3910 (1.50)
Services	.0380 (.66)	-.0589 (.51)	-.2136 (4.99)	-.1117 (.43)	-.2269 (.94)	.0161 (.11)
Constant	-.0294 (.17)	-.6218 (1.49)	-.4402 (4.24)	-.3618 (1.26)	-1.62 (4.47)	-.5693 (2.66)
Sample Size	942	465	2645	2255	981	1393
R2	.46	.18	.33	.09	.18	.14

Note: t-values in parentheses.

Table 5. FORMAL / INFORMAL SECTOR WAGE EQUATIONS  
MALE WORKERS

	<u>FORMAL</u>			<u>INFORMAL</u>		
	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>
Years Schooling	.0726 (17.73)	.1132 (14.58)	.0860 (24.82)	.0614 (10.69)	.1035 (8.41)	.0617 (13.29)
H.S. Diploma	.0543 (1.45)	-.1605 (2.16)	.1401 (4.99)	.2594 (3.85)	-.0438 (.52)	.2080 (4.98)
Experience	.0267 (10.59)	.0363 (6.73)	.0385 (21.13)	.0407 (14.52)	.0599 (13.07)	.0458 (21.76)
Experience Sq	-.00023 (5.37)	-.00038 (3.48)	-.00052 (14.21)	-.00051 (12.48)	-.00082 (10.32)	-.00064 (17.18)
Urban 1	.0469 (2.12)	.0759 (1.86)	.1592 (8.34)	.1909 (5.46)	.1907 (4.00)	.2193 (8.74)
Urban 2			.1058 (6.25)			.1643 (6.30)
Agricultural Ind.	-.3709 (4.36)	-.4141 (2.76)	-.4905 (3.05)	-.6307 (4.39)	-.5457 (3.29)	-.8583 (6.24)
Manufacturing	-.0662 (1.51)	-.1925 (2.90)	-.1174 (3.40)	-.2057 (1.45)	-.0712 (.48)	-.3111 (3.00)
Construction	-.1508 (2.88)	-.2108 (1.90)	-.0921 (2.07)	-.0862 (.60)	.1527 (.96)	-.3015 (2.86)
Commerce	-.2004 (4.25)	-.3201 (4.36)	-.1927 (5.21)	-.1577 (1.11)	.0369 (.25)	-.3476 (3.40)
Transport & Comm.	.0050 (.09)	-.0986 (1.16)	-.0411 (.92)	.1662 (1.14)	.2328 (1.48)	-.1139 (1.07)
Services	-.0240 (.59)	-.2003 (3.50)	-.2052 (5.89)	-.1284 (.90)	.0731 (.49)	-.2257 (2.21)
Constant	.5089 (8.31)	.3054 (2.65)	-.0698 (1.40)	.1383 (.91)	-.2913 (1.59)	.2743 (2.47)
Sample Size	1738	1176	5580	3056	1603	4378
R2	.38	.27	.34	.25	.21	.23

Note: t-values in parentheses.

Because all six regressions show the expected pattern, one wonders whether the processes of wage determination really differ at all between the sectors. However, as Table 6 shows, Chow tests reject in every case the hypothesis of equality of all coefficients across sectors.<sup>14</sup>

Table 6. CHOW TESTS FOR DIFFERENCES IN WAGE EQUATION COEFFICIENTS

	Females	Males
<b>El Salvador</b>		
F-Value	12.02**	16.69**
(Deg Freedom)	(12, 3173)	(12, 4770)
<b>Peru</b>		
F-Value	7.51**	3.02**
(Deg Freedom)	(11, 1424)	(12, 2757)
<b>Mexico</b>		
F-Value	3.58**	13.48**
(Deg Freedom)	(13, 4012)	(13, 9932)

Notes: \*\* indicates the F-value is significant at the .01 level.

The raw wage gap observed between workers in the formal and informal sectors can be decomposed into two parts, one explained by the difference between the sectors in the independent variables, the other not so explained:

$$\ln W_f - \ln W_i = (\bar{X}_f - \bar{X}_i)' \beta_f + \bar{X}_i' (\beta_f - \beta_i) = (\bar{X}_f - \bar{X}_i)' \beta_i + \bar{X}_f' (\beta_f - \beta_i)$$

Table 7 reports the raw formal log wage premium,  $\ln W_f - \ln W_i$ , the unexplained premium weighted by the average characteristics of the formal workers,  $\bar{X}_f' (\beta_f - \beta_i)$ , and the unexplained premium weighted by the average characteristics of the informal workers,  $\bar{X}_i' (\beta_f - \beta_i)$ .<sup>15</sup>

The premia are substantial in El Salvador. They are quite large for Peruvian women but also significant for Peruvian men. Unexpectedly, however, controlling for observable

<sup>14</sup> As always, the results are for the social security criterion. Firm size results in Appendix.

<sup>15</sup> Precise weighting rules are controversial. See Oaxaca and Ransom (1994).

characteristics completely washes out the small raw formal premium in Mexico; observationally identical workers earned more in Mexico in 1990 in the informal sector than in the formal sector.

Table 7. TOTAL AND UNEXPLAINED WAGE GAPS

	Females	Males
<b>El Salvador</b>		
Mean Ln Formal Wage	1.60	1.59
Mean Ln Informal Wage	0.64	0.88
Diff LnWage	0.96	0.71
Unexpl. (FW)	0.48**	0.26**
(t-value)	(19.7)	(16.2)
Unexpl. (IW)	0.38**	0.36**
(t-value)	(24.4)	(29.2)
<b>Peru</b>		
Mean Ln Formal Wage	1.90	1.92
Mean Ln Informal Wage	1.02	1.42
Diff LnWage	0.89	0.50
Unexpl. (FW)	0.31**	0.14**
(t-value)	(7.17)	(5.87)
Unexpl. (IW)	0.56**	0.12**
(t-value)	(19.1)	(6.06)
<b>Mexico</b>		
Mean Ln Formal Wage	1.12	1.20
Mean Ln Informal Wage	1.04	1.17
Diff LnWage	0.09	0.02
Unexpl. (FW)	-0.07**	-0.11**
(t-value)	(5.60)	(13.5)
Unexpl. (IW)	-0.09**	-0.13**
(t-value)	(5.40)	(14.5)

Notes: FW uses the mean characteristics of formal sector workers; IW uses the mean characteristics of informal sector workers. The t-values are calculated from pooled regression standard errors.

\*\* Indicates significant at the .01 level.

Three points might be added here. First, benefits obviously differ between workers who are covered by social security and those who are not. On this ground alone, one would expect informal wages to be higher than formal wages. This may help to explain the informal premium in Mexico; it cannot explain the formal premia in El Salvador and Peru.

Second, because of their numerical importance in the informal sector (see Table 8), we have included self-employed people and small entrepreneurs in all of the above analysis. However, we recognize that measuring the earnings of the self-employed is an especially difficult problem. Profits may be confused with returns to labor, and the return to the work of unpaid family members may be added into the measured income of the self-employed person.

Table 8. INFORMAL OWNERS AND SELF-EMPLOYED AS PERCENT OF ALL INFORMALS

<b>El Salvador</b>	Female	75.7%
	Male	40.7%
<b>Peru</b>	Female	75.0%
	Male	62.9%
<b>Mexico</b>	Female	45.6%
	Male	51.0%

We also note that students of wage determination in the United States have found a positive correlation between wage rates and firm size. This literature does not focus on truly small firms (by our firm size definition, informal firms have five or fewer workers), and its estimated elasticities of the wage with respect to firm size imply premia much smaller than those we find in Salvador and Peru (see, for examples, Brown and Medoff, 1989, p.1034). Nonetheless, the work on firm size in some ways parallels work on formal/informal wage differentials. If the categorical variables on firm size in our data permitted it, we would try to explore the parallels more closely.

### Section 3. Endogenous Sample Selection

We calculate wage premia on the basis of the estimated coefficients of the sectoral earnings functions. However, if workers freely choose between formal and informal employment, the wage and the sector of employment are simultaneously determined, and the same individual peculiarities which drive one result may also drive the other. One of Ciudad Nezahualc6yotl's microentrepreneurs said that he earned more in his informal car parts shop

than he could through formal employment. The same unobserved personal characteristic which makes him likely to choose informal work may also be raising his wage in the informal sector above that of a person of the same observed characteristics drawn randomly from the population. If everyone in the informal sector shares that “certain something,” then the wages we observe for informal workers are higher than those expected for the population as a whole, and the OLS parameter estimates are biased.<sup>16</sup>

Well-established techniques are available for incorporating this simultaneity into empirical work when the selection rule is clear. Unfortunately, as we have noted, there is no clear consensus about how workers sort themselves--or are sorted-- between sectors.

We propose a very simple model which is consistent with our hypothesis that those who choose to work can also freely choose the sector in which they will work. Consider an individual’s multiplicatively separable utility function  $U = Y(M)Z(N)$ , where  $M$  is the quantity consumed of a marketed composite good and  $N$  is the non-marketed, non-wage benefit accruing from work in a particular sector. The indirect utility function giving utility as a function of the price level  $p$ , expenditure  $E$ , and  $N$  is  $V = X(p,E)Z(N)$ . Assume that  $X$  is linearly homogeneous in  $E$  and that  $E$  is equal to the wage rate  $w$  (in terms of  $p$ ) times hours worked  $h$ . Both the wage rate and the level of non-wage benefits depend on the sector chosen; we assume that hours do not. Utility derived from work in sector  $s$  will equal  $V_s = w_s h X(p) Z(N_s)$ . Taking logs and subtracting, the difference in utility between sector 1 and sector 0 is:

$$\ln V_1 - \ln V_0 = \ln w_1 - \ln w_0 - [\ln Z(N_0) - \ln Z(N_1)]$$

A worker will choose sector 1 if  $\ln V_1 - \ln V_0 > 0$ . Let  $B$  represent the net relative non-wage benefit of working in sector 0,  $B \equiv [\ln Z(N_0) - \ln Z(N_1)]$ . Then a worker will choose the formal sector ( $s=1$ ) if the formal log wage premium, which could be either positive or negative, exceeds the net non-wage benefit associated with informal work,  $B$ , which could also be positive or negative:

---

<sup>16</sup> For a similar problem with self-employed workers, see Bernhardt (1994).

$$(1) \quad \begin{cases} \ln W_f - \ln W_i > B + \varepsilon_0 \Rightarrow s = 1 \\ \ln W_f - \ln W_i \leq B + \varepsilon_0 \Rightarrow s = 0 \end{cases}$$

$$\varepsilon_0 \sim N(0,1)$$

The notion of net non-wage benefits of informal work may require some explanation. Formal work often brings non-wage benefits, but non-wage benefits may be associated with informal work, as well. One of the women of Nezahualc6yotl said that she valued the sense of independence which comes with informal employment; this is a non-marketed, non-wage benefit. Informal work is often done close to one's home under flexible scheduling, and, as another of the women said, it may be carried on in conjunction with child care and home vigilance. Some might see this as a way to avoid the childcare costs associated with formal employment; others might see it as a quality of informal employment which is desirable in itself. The term B is meant to capture the balance of non-wage benefits between sectors.

We expect B to vary with the characteristics of a worker and the person's position in the household. We suggest, in an admittedly *ad hoc* manner, that importance of the independence and flexibility of informal employment to an individual will depend on whether or not one is married (legally or through *uni6n consensual*), on whether or not one is the household head, on the number of infants in the household who have not yet completed their first year, and on the number of inactive adults (alternative caregivers?) present in the household. All of these effects are likely to depend on whether one is male or female.

We also hypothesize that the earnings of other members of the household affect sectoral choice. There is no wage uncertainty in our model, but our data show that the variance of the informal wage exceeds the variance of the formal wage, and it seems plausible that informal incomes may in fact be more subject to shocks than formal incomes. If so, the earnings of other members of the household might serve as a sort of insurance fund, permitting a worker to

take the risk of accepting informal work, and we include this variable as a possible determinant of B.<sup>17</sup>

All the variables determining the wage in the earnings functions of Section 2 and the variables thought to influence B were substituted into Equation 1, and the reduced form probit was estimated. The results are given in Table 9, using the criterion of coverage by the social security system.<sup>18</sup> Likelihood ratio tests show the variables we have chosen to have significant explanatory power.<sup>19</sup>

The probability of formal employment increases with schooling and first increases, then decreases, with age or experience. The results for the other variables differ somewhat from country to country; the Peruvian coefficients are often statistically insignificant. Consider men first. In general, for men, the probability of formal work increases with marriage and with being a household head. The number of infants has no effect. The number of inactive adults in the household raises the probability of being formal, as expected. The probability of being formal also rises, unexpectedly, with the other labor income of the household. For women, marriage either has no significant effect or decreases the probability of formal work, as do the number of infants in the household and the earnings of other members of the household. The number of inactive people in the home again raises the likelihood of formal work.

The possible importance of incorporating sectoral selection into the wage regressions can be seen as follows. Equation 1 implies that those in the formal sector are workers for whom:

$$(2) \quad \varepsilon_0 < \ln W_f - \ln W_i - B$$

The earnings function for formal workers, as estimated in Section 2, had the form:

$$(3) \quad \ln W_f = X_f' \beta_f + \varepsilon_f$$

Estimating Equation 3 by OLS over those observed in the formal sector is appropriate only if the expected value of  $\varepsilon_f$  is zero.

---

<sup>17</sup> Non-labor income is not available in the Mexican data.

<sup>18</sup> As before, the results from the firm size criterion are shown in the Appendix.

<sup>19</sup> The test statistic, which is equal to twice the difference between the maximized value of the log likelihood function for the unconstrained case and the value for the case in which all coefficients but the constant are constrained to equal zero, follows a chi-square distribution. The value of this statistic is also given in Table 8.

Table 9. PROBITS FOR BENEFITS CRITERION

	MALES				FEMALES		
	El Salv.	Peru	Mexico		El Salv.	Peru	Mexico
Years Schooling	.0595 (7.19)	.1654 (12.89)	.0382 (6.20)		.1270 (10.84)	.2008 (9.34)	.0313 (2.99)
H.S. Diploma	.3464 (4.34)	-.5810 (5.90)	.0517 (1.00)		.2256 (2.03)	.2977 (1.60)	.2581 (3.39)
Experience	.0186 (3.46)	.0674 (8.73)	.0167 (4.41)	[Age]	.0710 (4.38)	.1115 (4.83)	.0497 (4.69)
Exper Squared	-.00036 (4.45)	-.00111 (8.38)	-.00049 (7.61)	[Age^2]	-.00100 (4.90)	-.00120 (3.98)	-.00077 (5.63)
Married	.2638 (4.27)	-.0079 (.10)	.2127 (2.23)		.0336 (.45)	.1655 (1.32)	-.2337 (4.00)
Head of HH	.2671 (3.73)	.1004 (1.20)	.1154 (2.23)		.1761 (1.88)	-.0765 (.42)	-.1627 (2.13)
# Babies	-.0022 (.04)	-.13E-04 (.00)	-.0457 (1.13)		-.2566 (2.54)	-.1401 (1.01)	-.0179 (.23)
Inactive People	.0521 (3.02)	.0518 (2.62)	.0420 (4.17)		.0483 (2.03)	.0798 (2.51)	.0543 (3.24)
Other Lab Income	.646E-04 (2.90)	.970E-05 (.79)	280E-04 (2.11)		.425E-05 (.15)	.115E-04 (.80)	-.508E-04 (2.77)
Urban 1	.3042 (6.76)	.2025 (3.68)	.0069 (.21)		.4285 (6.94)	.2395 (2.54)	-.0395 (.72)
Urban 2			.3287 (10.29)				.1851 (3.48)
Agricultural Ind.	-1.84 (12.56)	-1.54 (9.18)	-1.74 (8.51)		-1.17 (4.30)	-1.43 (2.63)	-1.18 (2.35)
Manufacturing	-.9296 (7.60)	-.7991 (6.83)	-.2727 (2.99)		-.9380 (4.63)	-.5015 (2.43)	-.1438 (1.14)
Construction	-.9292 (7.12)	-1.31 (8.94)	-1.19 (12.04)		-.8279 (1.97)		-.1200 (.48)
Commerce	-1.18 (9.45)	-1.40 (11.94)	-1.21 (13.28)		-1.53 (7.60)	-1.11 (5.58)	-.9843 (8.01)
Transport & Comm	-1.14 (8.63)	-1.10 (8.33)	-1.16 (11.61)		.7739 (2.39)	.2431 (.70)	.2265 (1.15)
Services	-.3344 (2.76)	-.5631 (4.99)	-.8808 (9.76)		-.3818 (1.91)	.2388 (1.24)	-.2651 (2.19)
Constant	-.7279 (4.82)	-1.58 (8.47)	.1630 (1.46)		-2.05 (5.85)	-5.13 (10.24)	-.2257 (.95)
Sample Size	4794	2779	9958		3197	1446	4038
$\chi^2$	1507	922	1604		1631	800	760

Note: t-values in parentheses

Given a variance of  $\varepsilon_0$  equal to one and covariance between  $\varepsilon_f$  and  $\varepsilon_0$  equal to  $\sigma_{0f}$ , it is well known that:

$$(4) \quad E(\varepsilon_f | \varepsilon_0 < \ln W_f - \ln W_i - B) = -\sigma_{0f} \frac{\phi(\ln W_f - \ln W_i - B)}{\Phi(\ln W_f - \ln W_i - B)} \equiv -\sigma_{0f} \lambda_f$$

where  $\phi$  is the density and  $\Phi$  the cumulative distribution of a standard normal random variable and

$$\lambda_f \equiv \frac{\phi(\ln W_f - \ln W_i - B)}{\Phi(\ln W_f - \ln W_i - B)}$$

(Maddala, 1983, p.224). Therefore:

$$(5) \quad E(\ln W_f | \varepsilon_0 < \ln W_f - \ln W_i - B) = X_f' \beta_f - \sigma_{0f} \lambda_f$$

Including  $\lambda_f$  as an additional regressor in Equation 3 permits unbiased estimation by ordinary least squares, even if the unobserved personal characteristics associated with the error terms at the two levels of the equation are correlated with one another. For those observed to work in the informal sector, the equation corresponding to (5) is:

$$(6) \quad E(\ln W_i | \varepsilon_0 \geq \ln W_f - \ln W_i - B) = X_i' \beta_i + \sigma_{0i} \lambda_i$$

where:

$$(7) \quad E(\varepsilon_i | \varepsilon_0 \geq \ln W_f - \ln W_i - B) = \sigma_{0i} \frac{\phi(\ln W_f - \ln W_i - B)}{1 - \Phi(\ln W_f - \ln W_i - B)} \equiv \sigma_{0i} \lambda_i$$

and

$$\lambda_i \equiv \frac{\phi(\ln W_f - \ln W_i - B)}{1 - \Phi(\ln W_f - \ln W_i - B)}$$

We used our reduced form probit model of sectoral assignment to calculate a selectivity factor,  $\lambda_f$ , for each person in the formal sector, and a factor  $\lambda_i$  for each person in the informal sector. The selectivity factors were added to the wage regressions, which were then run using ordinary least squares with TSP's White-style correction for heteroskedasticity (Hall, 1993, p.46). The estimated coefficients on the selectivity factors in the wage regressions are given in Table 10, again using the social security criterion of informality.<sup>20</sup>

---

<sup>20</sup> Those for the size criterion are shown in the Appendix.

Consider first the informal wage equation. In Peru, the coefficients are insignificantly different from zero, suggesting that  $\sigma_{0i}$  is not significantly different from zero and that the sample selection does not bias the straight OLS informal sector coefficient estimates of Section 2. For Salvadorans and for Mexican men, the coefficient on  $-\lambda_i$  is negative, implying that  $\sigma_{0i}$  is positive and that those most likely to be selected into the informal sector are also those most likely to have high wages there. For Mexican informal women the opposite effect was found. In the formal wage equation, insignificant effects were found for Salvadoran and Peruvian women. For Mexican women, the coefficient on  $\lambda_f$  was significantly positive, implying that  $\sigma_{0f}$  is negative, and that those most likely to choose formal employment are those with characteristics generating a high formal wage, as expected. For all men, however,  $\sigma_{0f}$  is positive, which is an unexpected result. In these cases, high  $\epsilon_0$ , which imply a low probability of entering the formal sector, tend to be associated with high  $\epsilon_f$ , which imply higher formal wages than would be expected in a random draw. One would expect the opposite sign if people with unobserved idiosyncrasies making them especially well-rewarded in the formal sector were able to select themselves into that sector.

Table 10. COEFFICIENTS ON SELECTIVITY VARIABLE IN WAGE REGRESSIONS

	<u>MALES</u>			<u>FEMALES</u>		
	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>
Formal Wage Equation: Density/Cum. Density	-.271** (3.30)	-.960** (4.27)	-.406** (3.78)	-.207 (.98)	-.071 (.28)	1.17** (8.87)
Informal Wage Equation: (-Density)/(1-Cum. Density)	-1.12** (9.38)	-.262 (.90)	-1.36** (9.22)	-.517** (3.13)	-.206 (.94)	1.04** (4.65)

\*\* Indicates significant at the .01 level.

Generally, our results suggest that people likely to earn more than observationally equivalent peers in informal employment do, in fact, choose the informal sector. Men with unobserved qualities likely to be highly rewarded in the formal sector, however, seem *not* to

have access to that sector. Does this support the image of the informal sector as easily-entered while formal opportunities are limited? Could it be that formal firms are choosing from queues of observationally equivalent male workers the ones at the bottom of the wage distribution, as might be the case if wage rates do not solely reflect productivity? Once immobilities are added to the model, one would not necessarily expect good hunters to hunt and good fishers to fish. It is also possible, however, that our model of selection into the two sectors is misspecified; little microeconomic work has been done on sectoral selection in developing economies. Reviewing similar corrections in the literature on union wage gaps, Lewis concluded that the straight OLS estimates are as reliable as those “corrected” for selection (1986, p.1144).

At any rate, the fundamental question for us is whether the best model we can estimate of sectoral assignment can explain the sectoral wage differences which we observed in Section 2. In Peru and El Salvador, informal workers are, if anything, drawn from the upper end of the possible informal wage distribution, and formal workers from the lower end of theirs. The formal wage premium which we found is, then, if anything, understated. The premium which we found for informal work among Mexican women is also likely to be understated; formal Mexican women are drawn from the upper end of the formal wage distribution, and informal women, if biased at all, from the lower end of the informal distribution. In the case of Mexican men, however, the selection factor could help to explain the observed gap.

#### **Section 4: Summary and Lines of Research**

Our analysis reveals patterns in the personal characteristics of informal workers, who represent one-third to one-half of all paid workers in these three countries, as well as patterns in the characteristics of their jobs. The young and the old are more likely to be informal than are prime-aged workers. Formal workers generally have more schooling than do informal workers. Retail trade is by far the largest single activity of informal sector workers in all three countries, but construction, transportation and personal services are also largely informal

activities. Jobs permitting very short or very long work weeks are likely to be in the informal sector.

Running OLS wage regressions separately for men and women in each sector, we found the returns to schooling and experience to be of the expected signs and magnitude in all cases. Substantial returns to schooling in the informal sector do not fit the characterization of a secondary sector as described by Dickens and Lang (1985). A Chow test rejected the hypothesis that the two sectors are characterized by identical wage-setting patterns.

The wage regressions were used to calculate unexplained wage gaps between the sectors. In El Salvador and Peru, we found that significant wage premia are associated with work in the formal sector. Unexpectedly, we found in Mexico that the premium is associated with *informal* work.

Using a reduced form probit selection model including sectoral wage differentials and the balance of non-wage costs and benefits of different sorts of work, we found that, for men, the probability of formal employment increases with heading a household, with marriage, and with the number of inactive adults in the household. For women, the model shows the probability of formal employment to decrease with marriage and with the number of infants in the household and to increase with the number of inactive adults in the household.

The selectivity factor proved in most cases to reinforce the wage gaps which we found in straight OLS estimation. If there is any bias at all, the informal workers of Peru and El Salvador are drawn from the upper end of the informal wage distribution, but the formal workers are not drawn from the upper end of the formal wage distribution. Therefore, our measure of the formal wage premium, if anything, understates the true premium. The selection results also reinforce the finding of an *informal* premium for Mexican women. We do find some evidence, however, that the informal premium for Mexican men may in part be due to this sort of sample selection according to comparative advantage.

We have been as inclusive as possible in our analysis, including not only very young and very old workers and part-time workers but also the self-employed. Since self-employment is

at the heart of the informal experience, we do not wish to exclude it from our analysis. However, both earnings and hours worked may be measured with special imprecision by the self-employed. We have also been careful to include alternative definitions of the informal sector. Results by social security coverage have been presented here; qualitatively similar results under the firm size definition are presented in the Appendix.

Additional data would be useful. Wages might be affected by job tenure and, in some countries, by ethnicity. The choice of sector might depend on household nonlabor income and on whether the worker receives social security coverage through the work of a parent or spouse. Data from additional years might also shed some light on these results. Woodruff (1995) examines changes in firm size-wage effects in Mexico between 1987 and 1990.

This research, which has taken care to examine data from three different countries in a consistent way, casts doubt on the received wisdom that the informal sector, always and everywhere, is a poorly-paid but easily-entered refuge for those who have no other option for employment.

## REFERENCES

- Bernhardt, Irwin, "Comparative Advantage in Self-Employment and Paid Work," Canadian Journal of Economics, Vol. 27, No.2, May, 1994, pp. 273-289.
- Brown, Charles, and Medoff, James, "The Employer Size--Wage Effect," Journal of Political Economy, Vol.97, No.5, October, 1989, pp.1027-1059.
- Castells, Manuel, and Portes, Alejandro, "World Underneath: The Origins, Dynamics, and Effects of the Informal Economy," in Alejandro Portes, Manuel Castells, and Lauren Benton, eds., The Informal Economy: Studies in Advanced and Less Developed Countries, Baltimore: Johns Hopkins University Press, 1989, pp.11-40.
- de Soto, Hernando, The Other Path: The Invisible Revolution in the Third World, translated by June Abbott, New York: Harper & Row, 1989.
- Dickens, William T., and Lang, Kevin, "A Test of Dual Labor Market Theory," American Economic Review, Vol. 75, No. 4, September, 1985, pp. 792-805.
- Fields, Gary S., "Rural-Urban Migration, Urban Unemployment and Underemployment, and Job Search Activity in LDC's," Journal of Development Economics, Vol.2, No.2, June, 1975, pp.165-187.
- Fields, Gary S., "Labour Market Modelling and the Urban Informal Sector: Theory and Evidence," in David Turnham, Bernard Salome, and Antoine Schwarz, eds., The Informal Sector Revisited. Paris: OECD Development Centre, 1990, pp.49-69.
- Funkhouser, Edward, "The Urban Informal Sector in Central America: Household Survey Evidence," Working Paper in Economics #23-94, Department of Economics, University of California, Santa Barbara, May, 1994.
- Gindling, T. H., "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., No., 1991, pp.585-605.
- Hall, Bronwyn H., Times Series Processor Version 4.2 User's Guide (Palo Alto: TSP International, February 1993).
- Harris, John, and Todaro, Michael, "Migration, Unemployment, and Development: A Two-Sector Model," American Economic Review, Vol.60, No.1, March, 1970, pp.126-142.
- Hart, Keith, "Informal Income Opportunities and Urban Employment in Ghana," Journal of Modern African Studies, Vol.11, No.1, 1973, pp.61-89.
- Heckman, James J., and Hotz, V. Joseph, "An Investigation of the Labor Market Earnings of Panamanian Males: Evaluating Sources of Inequality," The Journal of Human Resources, Vol. 21, No. 4, Fall, 1986, pp.509-542.
- International Labor Organization (ILO), Employment, Incomes, and Equality: A Strategy for Increasing Productive Employment in Kenya, Geneva: ILO, 1972.
- King, Elizabeth, "Does Education Pay in the Labor Market? The Labor Force Participation, Occupation, and Earnings of Peruvian Women," LSMS Working Paper #67, Washington, DC: The World Bank, 1990.

Lubell, Harold, The Informal Sector in the 1980's and 1990's, Paris: OECD Development Centre, 1991.

Maddala, G.S., Limited-Dependent and Qualitative Variables in Econometrics, New York: Cambridge University Press, 1983.

Magnac, Th., "Segmented or Competitive Labor Markets?" Econometrica, Vol. 59, No. 1, January, 1991, pp. 165-187.

Marcouiller, Douglas, and Young, Leslie, "The Black Hole of Graft: The Predatory State and the Informal Economy," American Economic Review, Vol. 85, No. 3, June, 1995, forthcoming.

Mazumdar, Dipak, "Segmented Labor Markets in LDCs," American Economic Review: Papers and Proceedings, Vol.73, No.2, May, 1983, pp.254-259.

Mezzer, Jaime, Informal Sector as in PREALC, Santiago: PREALC, 1990.

Moock, Peter, Musgrove, Philip, and Stelcner, Morton, "Education and Earnings in Peru's Informal Nonfarm Family Enterprises," LSMS Working Paper #64, Washington, DC: The World Bank, 1990.

Oaxaca, Ronald L., and Ransom, Michael R., "On Discrimination and the Decomposition of Wage Differentials," Journal of Econometrics, Vol. 5, No. 21, 1994, pp.5-21.

Pérez Sáinz, Juan Pablo, Informalidad urbana en América Latina: Enfoques, problemáticas e interrogantes, Caracas: Editorial Nueva Sociedad, 1991.

Rakowski, Cathy A., editor, Contrapunto: The Informal Sector Debate in Latin America, Albany: SUNY Press, 1994.

Roberts, Bryan, "The Changing Nature of Informal Employment: The Case of Mexico," Texas Papers on Mexico, No. 90-02, 1990.

Selby, Henry, Murphy, Arthur D., and Lorenzen, Stephen, The Mexican Urban Household: Organizing for Self-Defense, Austin: University of Texas Press, 1990 .

Telles, Edward E., "Urban Labor Market Segmentation and Income in Brazil," Economic Development and Cultural Change, Vol.41, No.2, January, 1993, pp.231-250.

Terrell, Katherine, "Analysis of the Wage Structure in Guatemala City," The Journal of Developing Areas, Vol. 23, April, 1989, pp.405-424.

Tokman, Victor E., ed., Beyond Regulation: The Informal Economy in Latin America, Boulder: Lynne Rienner, 1992.

Tokman, Victor E., "Policies for a Heterogeneous Informal Sector in Latin America," World Development, Vol.17, No.7, July, 1989, pp.1067-1076.

Woodruff, Christopher. "Inflation Stabilization and *The Case of the Vanishing Size-Wage Effect*." Working Paper, Graduate School of International Relations and Pacific Studies, University of California, San Diego, February, 1995.

## APPENDIX: RESULTS UNDER FIRM SIZE DEFINITION OF INFORMAL SECTOR

In the body of this paper, we developed two different operational criteria for identifying the informal sector, one based on coverage by social security systems, the other based on firm size. In general only the “social security” results were presented in the body of the paper; the results for the firm size criterion are shown here.

Tables A.1 and A.2 give the estimated coefficients of the sectoral earnings functions. These results correspond to Tables 3 and 4 in the paper. The shape of the regressions is similar to those displayed previously. The Chow tests (Table A.3) once again reject the hypothesis that the wage setting patterns are identical across sectors.

Table A.1 FORMAL / INFORMAL SECTOR WAGE EQUATIONS  
SIZE CRITERION--FEMALE WORKERS

	FORMAL			INFORMAL		
	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>
Years Schooling	.0649 (10.64)	.0703 (4.74)	.0798 (16.78)	.0365 (4.73)	.0909 (5.20)	.0320 (2.90)
H.S. Diploma	.1192 (2.01)	.1153 (.72)	.0306 (.90)	.2326 (2.43)	.0663 (.52)	.1234 (1.37)
Age	.0447 (5.99)	.0878 (4.37)	.0464 (9.18)	.0406 (6.12)	.1025 (8.07)	.0534 (6.59)
Age Squared	-.00036 (3.73)	-.00087 (3.22)	-.00044 (6.17)	-.00044 (6.00)	-.00103 (6.70)	-.00058 (6.12)
Urban 1	.0945 (2.93)	.1121 (1.52)	.0002 (.01)	.1066 (2.41)	.0291 (.44)	.1534 (2.50)
Urban 2			.0823 (3.55)			.1163 (1.88)
Agricultural Ind.	-.3725 (3.87)	-.2403 (.76)	-.0716 (.23)	-.0198 (.06)	-1.40 (4.98)	.3793 (.95)
Manufacturing	-.2247 (2.97)	-.3265 (2.35)	-.2104 (4.97)	-.0213 (.06)	-.3191 (1.37)	.0448 (.18)
Construction	-.1601 (.92)			-.5196 (.88)		
Commerce	-.3589 (4.76)	-.2634 (1.85)	-.2373 (5.42)	-.0569 (.17)	-.0880 (.40)	.0065 (.02)
Transport & Comm.	.2772 (2.18)	.0006 (.00)	-.0205 (.31)	.1935 (.28)	.6396 (1.75)	.5492 (1.51)
Services	-.0222 (.32)	.0360 (.29)	-.1821 (4.55)	-.0340 (.10)	-.3046 (1.36)	.0775 (.32)
Constant	-.2840 (1.85)	-1.13 (2.82)	-.5312 (5.27)	-.3969 (1.11)	-1.70 (4.89)	-.4614 (1.50)
Sample Size	1225	449	3056	1972	997	982
R2	.48	.30	.31	.06	.21	.08

Note: t-values in parentheses.

Table A.2 FORMAL / INFORMAL SECTOR WAGE EQUATIONS  
SIZE CRITERION--MALE WORKERS

	FORMAL			INFORMAL		
	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>
Years Schooling	.0830 (19.82)	.1279 (15.69)	.0908 (26.98)	.0466 (6.49)	.0984 (8.63)	.0391 (7.45)
H.S. Diploma	.0970 (2.37)	-.0863 (1.11)	.1246 (4.46)	.2552 (2.97)	-.0685 (.85)	.2134 (4.66)
Experience	.0385 (15.95)	.0461 (8.63)	.0386 (22.73)	.0364 (10.54)	.0575 (12.82)	.0430 (18.16)
Experience Sq	-.00037 (9.16)	-.00048 (4.14)	-.00050 (14.67)	-.00049 (9.99)	-.00080 (10.37)	-.00062 (15.30)
Urban 1	.1248 (5.22)	.1914 (4.34)	.1604 (8.59)	.1539 (3.55)	.1161 (2.59)	.2220 (8.11)
Urban 2			.1148 (6.83)			.1419 (5.01)
Agricultural Ind.	-.5410 (9.14)	-.1059 (.83)	-.3066 (2.27)	-.6280 (2.84)	-1.05 (7.97)	-.9211 (4.83)
Manufacturing	-.1617 (3.19)	-.1199 (1.71)	-.1256 (3.56)	-.0120 (.05)	-.4196 (3.73)	-.3069 (1.83)
Construction	-.1256 (2.14)	.0623 (.61)	-.0753 (1.77)	.0763 (.35)	-.2717 (2.16)	-.3126 (1.85)
Commerce	-.2269 (4.10)	-.1703 (2.16)	-.1742 (4.62)	.0700 (.32)	-.3470 (3.26)	-.3084 (1.86)
Transport & Comm.	.0838 (1.36)	-.0662 (.69)	-.0710 (1.56)	.3174 (1.44)	-.0900 (.76)	-.0243 (.14)
Services	-.0646 (1.33)	-.1011 (1.49)	-.1783 (5.06)	.0368 (.17)	-.2576 (2.43)	-.2364 (1.42)
Constant	.1739 (2.69)	-.2840 (2.45)	-.1054 (2.19)	.1151 (.50)	.2253 (1.49)	.4505 (2.57)
Sample Size	2658	1064	6622	2136	1715	3336
R2	.44	.40	.34	.23	.21	.18

Note: t-values in parentheses.

Table A.3 CHOW TESTS FOR DIFFERENCES IN WAGE EQUATION COEFFICIENTS: BY  
FIRM SIZE

	Females	Males
El Salvador F-Value (Deg Freedom)	10.50** (12, 3173)	18.57** (12, 4770)
Peru F-Value (Deg Freedom)	3.67** (11, 1424)	3.17** (12, 2757)
Mexico F-Value (Deg Freedom)	4.32** (12, 4012)	15.37** (13, 9932)

Notes: \*\* indicates the F-value is significant at the .01 level.

Table A.4, which corresponds to Table 7 in the text, gives the total and the unexplained wage premia associated with formal work. Under the firm size definition of informality, the formal premia are reduced in El Salvador and especially in Peru, where the premium for men is now statistically insignificant; conversely, Mexico's informal premium is less under this definition of informality than it was before. Once again, by demonstrating that a formal wage premium cannot always be found, the results cast doubt on the notion that the informal sector is merely a low-wage but easily-entered refuge for those excluded from the formal sector.

Table A.4 TOTAL AND UNEXPLAINED WAGE GAPS: BY FIRM SIZE

	Females	Males
<b>El Salvador</b>		
Mean Ln Formal Wage	1.42	1.37
Mean Ln Informal Wage	0.60	0.84
Diff LnWage	0.82	0.53
Unexpl. (FW)	0.46**	0.30**
(t-value)	(21.6)	(22.6)
Unexpl. (IW)	0.30**	0.27**
(t-value)	(17.7)	(18.0)
<b>Peru</b>		
Mean Ln Formal Wage	1.72	1.79
Mean Ln Informal Wage	1.12	1.54
Diff LnWage	0.61	0.25
Unexpl. (FW)	0.19**	0.02
(t-value)	(4.33)	(0.82)
Unexpl. (IW)	0.25**	0.04
(t-value)	(8.48)	(1.84)
<b>Mexico</b>		
Mean Ln Formal Wage	1.13	1.22
Mean Ln Informal Wage	0.97	1.12
Diff LnWage	0.16	0.10
Unexpl. (FW)	0.03**	0.01
(t-value)	(3.17)	(1.24)
Unexpl. (IW)	-0.04*	-0.08**
(t-value)	(2.19)	(7.41)

Notes: FW uses the mean characteristics of formal sector workers; IW uses the mean characteristics of informal sector workers. The t-values are calculated from pooled regression standard errors.

\* Indicates significant at the .05 level.

\*\* Indicates significant at the .01 level.

Table A.5 gives the estimated coefficients on the sectoral assignment equation under the firm size criterion of informality. It corresponds to Table 9 in the text. Again, the probability of formal employment is shown to increase with schooling and first to increase, then to decrease, with experience or age. In general, the pattern of the coefficient estimates is similar to that under the alternative criterion of informality, although in some cases a coefficient which was insignificantly different from zero has been pushed to significance one way or the other.

TABLE A.5. PROBITS FOR SIZE CRITERION

	MALES			FEMALES		
	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>
Years Schooling	.0471 (5.83)	.1310 (11.22)	.0620 (9.54)	.1204 (10.77)	.2132 (10.85)	.0586 (4.87)
H.S. Diploma	.3864 (4.71)	-.4905 (5.20)	.1393 (2.55)	.2294 (2.09)	-.3178 (1.93)	.2202 (2.55)
Experience	-.0021 (.44)	.0266 (3.59)	.0050 (1.28)	[Age] -.24E-04 (.002)	.0196 (1.01)	.0263 (2.27)
Exper Squared	-.00013 (1.83)	-.00062 (4.66)	-.00025 (3.85)	[Age^2]-.00021 (1.32)	-.00033 (1.33)	-.00058 (3.97)
Married	.1660 (2.85)	.0920 (1.22)	.1079 (2.25)	-.0380 (.53)	.2180 (1.88)	-.1932 (2.95)
Head of HH	.1084 (1.60)	-.1854 (2.28)	.0437 (.80)	.2002 (2.25)	.0372 (.22)	.0686 (.80)
# Babies	-.1249 (2.13)	-.0653 (.92)	-.1060 (2.54)	-.1769 (1.90)	-.2186 (1.72)	-.0090 (.10)
Inactive People	.0066 (.40)	.0523 (2.74)	.0455 (4.29)	.0270 (1.17)	.0670 (2.31)	.0470 (2.47)
Other Lab Income	.331E-04 (1.43)	.3088-05 (.26)	.152E-04 (1.05)	.591E-04 (2.07)	.120E-04 (.91)	.349E-04 (1.51)
Urban 1	.3037 (6.67)	.3195 (5.98)	-.1286 (3.72)	.4401 (7.32)	.3337 (3.88)	-.1406 (2.26)
Urban 2			.1133 (3.36)			-.0281 (.46)
Agricultural Ind.	-.8856 (5.63)	-.4879 (3.28)	-1.87 (8.93)	.6121 (2.34)	.1382 (.45)	-1.82 (3.42)
Manufacturing	-.9299 (6.11)	-.0316 (.30)	-.4699 (3.71)	-.8443 (3.73)	.1614 (.80)	-.3343 (2.07)
Construction	-1.10 (6.96)	-.4127 (3.11)	-1.08 (8.21)	-.4812 (1.02)		
Commerce	-1.50 (9.75)	-.8240 (7.75)	-1.62 (12.86)	-1.48 (6.59)	-.7827 (4.04)	-1.49 (9.53)
Transport & Com	-1.37 (8.65)	-.7068 (5.71)	-1.52 (11.51)	.4789 (.98)	-.0908 (.27)	-.1969 (.82)
Services	-.3925 (2.57)	-.6070 (6.00)	-1.03 (8.24)	-.2381 (1.05)	-.0795 (.42)	-.4108 (2.60)
Constant	.5610 (3.24)	-1.26 (7.36)	.7731 (5.39)	-.3381 (1.04)	-3.02 (7.10)	.6905 (2.49)
Sample Size	4794	2780	9958	3197	1446	4038
$\chi^2$	1133	619	1925	1777	696	1176

Note: t-values in parentheses.

Table A.6, corresponding to Table 10 in the text, gives the estimated coefficients on the selectivity variable when it was added to the wage equation. The Salvadoran signs are the

same as before and always significant. The one Peruvian coefficient which was significant under the other definition has lost significance. The coefficient for Mexican men in the formal sector now has the expected sign and is significant.

Table A.6. COEFFICIENTS ON SELECTION VARIABLES: SIZE CRITERION

	MALES			FEMALES		
	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>	<u>El Salv.</u>	<u>Peru</u>	<u>Mexico</u>
Formal Wage Equation: Density/Cum. Density	-.531** (4.39)	.372 (1.53)	.349** (2.66)	-.384** (2.79)	-.520 (1.67)	.353* (2.11)
Informal Wage Equation: (-Density)/(1-Cum. Density)	-1.15** (6.25)	.151 (.54)	-.968** (9.22)	-.631** (3.02)	-.348 (1.45)	.261 (1.01)

\* Indicates significant at the .05 level.

\*\* Indicates significant at the .01 level.