PN-ACL-225

Working Paper Series

The Center for Institutional Reform and the Informal Sector (IRIS) has two main purposes: expanding knowledge about institutions in economic development through research, and assisting reform efforts in the third world and in countries undergoing transitions to a market economy. The premise of the IRIS Center is that in unsuccessful economies the existing rules establish poor incentives, often forcing economic activity into the informal economy, and that appropriate reforms improve economic performance. IRIS is especially concerned with the legal and policy framework needed for democratic societies with competitive markets.

IRIS
2105 Morrill Hall
College Park, MD 20742
(301) 405-3110

CENTER FOR INSTITUTIONAL REFORM AND THE INFORMAL SECTOR

University of Maryland at College Park

Center Office: IRIS Center, 2105 Morrill Hall, College Park, MD 20742
Telephone (301) 405-3110 • Fax (301) 405-3020

EMPLOYER SIZE, WAGES AND THE NATURE OF EMPLOYMENT IN PERU

September, 1993

Julie Anderson Schaffner
Working Paper No. 79

This work was funded through the Center for Institutional Reform and the Informal Sector (IRIS).

Author: Julie Anderson Schaffner, Stanford University

Employer Size, Wages and the Nature of Employment in Peru

Julie Anderson Schaffner*
Stanford University

September 1993

Abstract: The paper first describes wage, benefits and job tenure distributions across establishment size categories among urban Peruvian male private sector wage earners over 14 years of age. Second, it assesses the extent to which the pattern of superior wages and benefits paid in larger establishments merely reflects compensating differentials or selection of more productive workers (along unobserved productivity dimensions) into larger establishments. This allows assessment of the (residual) extent to which wage differences reflect "premiums," in the sense that identical individuals find employment superior in larger establishments. Evidence suggests that substantial premiums exist. Finally, the paper assesses the importance of two traditional institutional explanations for large employer premiums, higher incidence of unionization or higher compliance with minimum wage legislation, and the importance of various new institutional explanations, including several efficiency wage arguments.

* Department of Economics, Stanford University, Stanford, CA 94305-6072. The author gratefully acknowledges helpful comments from Greg Anderson, John Pencavel and Frank Wolak, as well as Jaana Remes' excellent research assistance, Paul Glewwe's help in obtaining the data, and funding through the Center for Institutional Reform and the Informal Sector (IRIS) at the University of Maryland.

I. Introduction

Wage differences by establishment size in Peru are higher than in the U.S. and comparable to the highest levels found in developing countries. If such differences reflect that identical workers in different establishments receive compensation that does not leave them equally well off, then the "Law of One Price" is violated. This would imply inefficiency in the allocation of labor and the existence of income differences that constitute rewards for luck or connections rather than compensation for human capital acquisition. Identifying whether and why such differences arise may ultimately aid policy formulation. Careful study of such differences in Peru, where the legal and industrial structure are very different from the U.S., may also aid evaluation of explanations for size-wage effects in the developed countries.

This paper has three objectives: First, to describe wage, benefits and job tenure distributions across establishment size categories among urban Peruvian male private sector wage earners over 14 years of age. Second, to assess the extent to which superior wages and benefits in larger establishments reflect "premiums," in the sense that identical individuals find wages and employment conditions superior in larger establishments. This involves eliminating two explanations for size-wage patterns that are consistent with perfectly functioning labor markets: that higher wages merely compensate for inferior working conditions and that higher wages are merited by workers who are on average of higher quality along dimensions not controlled for. Finally, since premiums are found to exist, the paper seeks to assess the importance of several possible explanations for their existence. Two "traditional institutional" explanations are that greater (threat of) unionization or higher compliance with minimum wage legislation in larger

establishments drives the size-wage differences. Several "new institutional" explanations point to the greater importance of asymmetric information, coordination and morale problems in larger establishments, leading to greater reliance on efficiency wage strategies.

Most of the study involves analysis of the Peruvian Living Standards Survey, collected jointly by the World Bank and Peruvian government institutions in 1985-86. The survey reached a representative sample of 5000 Peruvian households, out of which 1133 males over age 14 lived in urban areas, were wage employees in private sector establishments with at least two workers, reported positive wages, and provided information on several other characteristics. Given the modest size of the sample, special care has been taken to assess the robustness of results; and as many cuts at each question are taken as are possible. Variable descriptions, including discussion of imputation of some missing values, are found in the appendix. Additional information on data cleaning is available from the author.

The focus is on differences in wages and employment conditions across four establishment size categories, derived from workers' reports of the number of workers at their place of employment. The categories are: micro (2 to 5 workers), small (6 to 20 workers), medium (21 to 200 workers) and large (over 200 workers). Three of the four size categories refer to establishments that would be considered small by U.S. standards; but the great majority of Peruvian workers are employed in such establishments. Firm size, which has an independent effect on wages in U.S. studies, is not measured independently. Firm and establishment size are correlated; so the establishment size effects examined here capture some firm size effects. Firm size probably varies substantially less than establishment size, since small and medium firms in

Peru are thought to have strong incentive to divide workers across various micro workplaces in order to prevent detection and the enforcement of cumbersome and costly regulations.

The size category divisions not only correspond to common size groupings in Peruvian firm size studies (see, e.g. Villaran, et al., 1988), but also allow some interface with the large literature on "formal" and "informal" sectors in Peru and other developing countries. Definitions of, and reasons for interest in, the informal sector vary, but operational definitions often include workers in establishments of fewer than 5 or 20 workers, together with the self-employed and unpaid family workers (see, e.g., Carbonetto, et al., 1988). Inspired by Harris and Todaro (1970) and modified to account for the perception of widespread urban unemployment in developing countries (see Fields, 1975), the most popular models of developing country labor markets distinguish the urban formal sector, in which employers pay premium wages to the privileged few, and the informal sector, in which otherwise identical workers find tenuous employment and earn a low, market-clearing wage. The failure of competition to bid wages to equality across sectors is thought to explain slow employment expansion in the formal sector, inefficient rural-urban migration, and depressed wages and persistent poverty in the informal sector. de Soto (1989) popularized a very different view of the informal sector, defined largely as production outside the law and seen as the locus of entrepreneurial skill and economic potential in an economy otherwise made moribund by excessive regulation.

The interface with the informal sector literature is useful because raging debates about the informal sector in Peru have generated a variety of descriptive studies shedding some light on differences between small and large

establishments. The current study may also be of interest to those studying the informal sector, first, because it describes additional differences between what some would call formal and informal sectors. More importantly, it may illustrate a more fruitful approach to the study of urban labor markets than the formal/informal sector debate has tended to generate. The inability to identify a dividing line based on firm or employee characteristics that neatly separates out a formal sector with uniformly higher wages has generated scepticism about the usefulness of the formal/informal sector distinction (Peattie, 1987; Kannappan, 1988). Such attempts have been guided by overly simple models - with only two employer types and no acknowledgement of the tremendous heterogeneity of labor. Examination of differences in wage distributions for similar workers across more than two employer types, allowing distributions to overlap but nonetheless differ significantly, may prove more fruitful.

II. Basic Wage and Benefits Patterns

As Table 1.a demonstrates, wages rise steadily with establishment size in Peru. Mean cash wages in large establishments exceed those in micro establishments by 135 percent, while the difference for a wage measure including the cash value of in-kind payments (total wage, see appendix) is even higher, at 211 percent. Far from being driven by a few outliers, mean wage differences reflect employer size differences throughout the wage distribution, though absolute differences are higher at the high end of the wage distribution. Larger establishments employ more highly schooled, more experienced and possibly more stable (as indicated by household headship) workers; but as Table 2 demonstrates, controlling for such characteristics in standard log wage regressions leaves substantial differences in wage

distributions unexplained. Coefficients on establishment size indicators imply that employees in small, medium and large establishments earn on average 27, 49 and 73 percent higher total wages than employees in micro establishments. Comparison of columns 1 and 2 indicates that including in-kind payments augments the largest establishment differential, though it moderates slightly the small and medium coefficients. Quantile regressions indicate a shifting of the entire wage distribution as establishment size rises. Twenty-fifth and seventy-fifth quantile regressions (not shown) are quite similar to the median regression. The tenth and nintieth quantile regressions demonstrate, however, that establishment size has a much bigger (percentage) effect on the lower tail of the distribution than on the upper tail.

Such establishment size wage differences appear higher than those found in the U.S. The May 1988 Current Population Survey allows U.S. workers' establishment sizes to be grouped into four categories: 1-9 workers, 10-49 workers, 50-99 workers and at least 100 workers. The Peruvian size categories may be regrouped for close comparison into ranges: 2-10 workers, 11-50 workers, 51-100 workers, and over 100 workers. Fifty percent of the CPS sample, and 71 percent of the Peruvian sample, falls into the smallest two size categories. When dummies for the three largest categories are included in log wage regressions made as similar as possible to those in Table 2, the coefficients in the CPS sample are .143, .215 and .297, while the coefficients in the Peruvian sample are .291, .340 and .425 (all strongly significant in both samples).[1]

Though comparability is even more difficult to achieve when examining results from other developing countries, Peruvian size-wage differences appear

similar to those found in other developing countries. Little, et al. (1987), using data on 5000 male workers in Bombay, India, find that after controlling for workers' schooling, age and indicators of training, language and blue collar status, workers in establishments of over 500 employees earn on average 42 percent more than employees of establishments with 10-99 workers (p. 256). A similar study in Malaysia found that workers in enterprises of 1-9 workers earned 32 percent less than those in larger enterprises (p.275). When workers' productive traits are not controlled for, evidence from a variety of countries reflect average wages around twice as great in large establishments relative to small, with size cutoffs ranging from 100 to 500.

Higher wages are accompanied by greater incidence of paid holidays, sick leave and pensions in larger establishments in Peru, whether or not basic worker characteristics are controlled for, as seen in Table 1.b and Table 3. For a "representative individual" with mean schooling and experience and who is a household head, the predicted probability of receiving a pension, for example, is 30, 57 and 65 percent higher in small, medium and large establishments relative to micro establishments, according to Table 3 probit results.[2]

Stability of employment also appears to rise with establishment size. Mean job tenure rises little in the move from micro to small establishments, and somewhat more with subsequent size increases (see Table 1.c). As with wage comparisons, it is useful to make tenure comparisons while controlling for differences in worker composition across establishment sizes. Duration analysis yields more easily interpretable results than ordinary least squares tenure regressions, and is employed in Table 4.[3]

Rather than specifying the functional relationship between the mean of

tenure's distribution and the independent variables, as in ordinary least squares, in duration analysis the researcher specifies the relationship of the hazard function of the job tenure distribution to the independent variables. The hazard function is a function of job tenure t, as well as the independent variables, and relates the probability that a job ends at any t, conditional on it not having ended prior to t. Functional form assumptions about the hazard function have more intuitive interpretation than assumptions about mean tenure, because they allow the explanatory variables to affect tenure through the probabilities governing a sequence of simple binary outcomes: stay on the job another period versus quit or be fired. The effects of the explanatory variables may readily be translated into, say, percentage changes in the probability of quitting or being fired in a given period.

Table 4 presents results of estimating a simple duration model of job tenure in which the hazard function for a randomly sampled job spell is assumed to take the Weibull form:[4]

$$h(t;X) = \exp(X^T\beta)\alpha t^{a-1}$$

where t is elapsed tenure, X is a vector of individual and establishment characteristics, and the scalar α and the vector β are parameters to be estimated. Because the data to be examined represent a sample of tenure to date for spells in progress at the interview date rather than a random sample of completed job spells, the Weibull assumption must be combined with an assumption about the inflow into employment (for each X) over previous periods to derive the likelihood function for the tenure to date observations. The derivation lying behind Table 4's results employs the dubious but popular assumption, discussed more in the next section, that inflow into employment over previous periods was constant for all X.[5]

Column 1 of Table 4 controls for schooling and experience (and its square) at time of hire, column 2 controls additionally for household structure variables at time of hire (which may influence workers' mobility), and column 3 controls for these and industry (which may be associated with cyclical sensitivity and fire rates). All three columns suggest that hazard rates out of employment in small establishments are not significantly different from those out of micro establishments, while hazard rates out of medium and large establishments are significantly lower.[6] The coefficients in column 3 imply that exit rates are 19 and 34 percent lower in medium and large establishments than in micro establishments. (Column 4 is discussed below.) Thus, at least on the face of it, while expected job tenure is similar in micro and small establishments (after controlling for worker and industry composition), workers in medium and large establishments appear to have more stable employment as well as higher wages and superior benefits.

III. Do Premiums to Employment in Larger Establishments Exist?

Assessing whether "premiums" exist involves ruling out two possible explanations for observed correlations between establishment size and wages that are possible even when labor markets leave identical individuals equally well off. First, higher wages could be required to compensate for greater unpleasantness or costs of holding jobs in larger establishments, as a result of greater routinization of tasks, more impersonal relations in the workplace, or longer commutes (compensating differential arguments). For larger establishments to remain in the market, the higher wages must in turn be compensated by lower costs of capital or some other privilege. In principle, higher wages may also compensate for inferior nonwage benefits or greater job instability, but such explanations are largely ruled out by the results of

Part II. Second, the higher wages may be merited by workers who are of higher quality along lines as yet uncontrolled for (the selection argument). Such systematic differences in worker quality may arise out of differing technological requirements across establishment sizes.[7]

Attempts to assess directly the importance of several compensating differential arguments are possible using the Peruvian data. First, working conditions probaby differ across industries, and industry composition differs by establishment size, admitting the possibility of compensating differences as an explanation for size effects. Table 5 suggests that average wages differ across industries (possibly because of differences in working conditions), that even within industries wages rise with establishment size, and that two high-wage industries, mining and electricity, contribute few observations to the dataset and are largely concentrated among larger establishments. In a study of within-industry establishment size effects those industries are best discarded. The first three columns of Table 6 indicate that discarding observations on workers in mining and electricity reduces the large establishment coefficient a little, from about .55 to .49, and that including controls for the remaining industries in the smaller sample if anything increases the size coefficients. Thus this first potential source of compensating differentials does little to reduce the apparent larger establishment wage premiums.

A second possible compensating differential argument arises because locational mix differs by establishment size. As shown in Table 7.a, larger establishments are more likely to be in Lima, the congested capital city. Any cost of living differences between Lima and other urban areas not captured in the deflator used for constructing the wage variable, as well as greater costs

of transportation to work in Lima, might then lead to higher observed wages in larger establishments that do not reflect true premiums. The fourth column of Table 6 demonstrates, however, that when an indicator of residence in Lima is included in the log wage regression, size effects are reduced by very modest amounts, even though residence in Lima appears associated with wages about 27 percent higher on average.

A third compensating differential argument arises because large establishments probably have to recruit workers over larger geographical areas, and must compensate them for higher commuting costs.[8] Though the Peruvian data do not contain geographical commute distance, they do contain information on mode of transport to work and time spent travelling to work. Tables 7.b and 7.c illustrate that the most common modes of transport to work are public transportation and walking, that the relative importance of public and other motorized transportation rises with establishment size, and that median time spent travelling to work rises with establishment size.

Controlling for commuting distances measured in minutes in log wage regressions is complicated for several reasons. First, mode of transport is endogenous. Workers receiving higher wages (for reasons unrelated to the commute), living at any given distance from work, are more able to purchase motorized transport and reduce their commuting time. Second, even conditional on mode of transport, commuting time is endogenous. Workers earning more may choose either to move closer to work to reduce commute time and increase leisure, or to live farther away in more desirable suburbs. Tables 7.d and 7.e show some tendency for mean commute time conditional on mode of transport to fall eventually as establishment size rises, providing some reason to be concerned about endogeneity.

To control for commuting differences as carefully as possible, attention is focussed on workers who take motorized transport or walk. Column 5 of Table 6 shows that this change in sample size alone has very little effect on the results so far. The econometric model employed may be described as follows:

$$\theta^* = Z\gamma + \varepsilon$$

$$M = Z\delta_M + u_M$$

$$W = Z\delta_W + u_W$$

If $\theta^* > 0$, then $\theta=1$, M is observed, and $Y = X\beta + \phi_M M + \eta$

If $\theta^* \leq 0$, then $\theta=0$, W is observed, and $Y = X\beta + \phi_W W + \eta$,

where θ^* is the latent propensity to take motorized transport, θ is the observed indicator of whether an individual takes motor transport, M and W are distances from work in time spent ln motor transport and walking, and Y is the log wage. X contains all the explanatory variables of column 5 of Table 6. Z contains X as well as several variables thought to influence residence location and mode of transport but not wages: household nonlabor income per capita, numbers of other adult males and females in the household, number of the individual's children ages 14 or less, and an indicator of whether the individual ever migrated. ε, u_M, u_W, and η are normally distributed with zero means in the population as a whole. The variance of ε is normalized to one. Nonzero covariances between the errors introduce endogeneity problems and must be accounted for.

Table 8 presents joint maximum likelihood estimates of the θ^*, M and W relations.[9] The log wage relation may be estimated consistently in a second stage, and the results are presented in the last column of Table 6 for ready comparison to previous results.[10] The nature of the second stage regression is

determined by observing that

$$E(Y|Z,\theta=1) = X\beta + \phi_M E(M|Z,\theta=1) + E(\eta|Z,\theta=1)$$

and

$$E(Y|Z,\theta=0) = X\beta + \phi_W E(W|Z,\theta=0) + E(\eta|Z,\theta=0)$$

where

$$E(M|Z,\theta=1) = Z\delta_M + \sigma_M\rho_M \frac{\phi(Z\gamma)}{\Phi(Z\gamma)}$$

$$E(W|Z,\theta=0) = Z\delta_W - \sigma_W\rho_W \frac{\phi(Z\gamma)}{\Phi(-Z\gamma)}$$

$$E(\eta|Z,\theta=1) = \sigma_\eta\rho_\eta \frac{\phi(Z\gamma)}{\Phi(Z\gamma)}$$

and $E(\eta|Z,\theta=0) = -\sigma_\eta\rho_\eta \frac{\phi(Z\gamma)}{\Phi(-Z\gamma)}$.

The four expected value terms may be estimated consistently (at least up to scale) using first stage results. Once $\theta E(M|Z,\theta=1)$ and $(1-\theta)E(W|Z,\theta=0)$, and an appropriately calculated inverse Mills' ratio term (λ) capturing the $E(\eta|.)$ up to scale ($\sigma_\eta\rho_\eta$) are included on the right hand side, the $\theta=1$ and $\theta=0$ relations above may be stacked, and the residuals, though heteroscedastic, are of mean zero and allow consistent estimation of β, ϕ_M and ϕ_W (as well as the scale factor in the $E(\eta|.)$ terms). The standard errors for the estimates must be corrected for the presence of estimated regressors, and are corrected following Murphy and Topel (1985).

The coefficient on λ in the final column of Table 6 indicates, as expected, that the unobservables tending to increase wages also tend to increase the probability of taking motor transport. The coefficient on expected time spent walking has the anticipated sign, indicating higher wages compensating for greater distance, but is insignificant. The surprise is the strongly negative coefficient on expected time spent in motor transport, the coefficient indicating that a 10 minute increase in motor transport time is

associated with a 12 percent decrease in the wage. Though it is possible that people who live farther away for exogenous reasons are deemed less committed to their work and accordingly receive lower wages, this result suggests more research into residence location and commute time might be worthwhile.

Having controlled for commuting time in this way, the size coefficients increase. Thus again a potential compensating differential argument has failed to explain away the size- wage correlation. Given that the size effect is not reduced, and that the results are somewhat equivocal, commute considerations are dropped in what follows.

A fourth possible reason for a compensating differential often controlled for in U.S. studies of employer-size wage patterns is hours. Table 7.f shows that mean usual weekly hours are highest in large establishments, while median hours are constant at 48, the legal work week. The increase in average weekly hours is largely driven by lower incidence of part-time work, which is illustrated in the next section of the table. Being able to offer longer hours may allow employers to offer lower hourly wages to hours-constrained low-income workers; some reason for thinking this is found in the declining incidence of reported search for additional work as establishment size increases, as shown in Table 7.h. Thus the hours differences are unlikely to explain large establishment wage differentials.

The effects of controlling for hours in log wage regressions are shown in Table 9. The first column replicates the fourth column of Table 6 for comparison. The second column demonstrates that restricting the sample to full-time workers if anything increases the large establishment differential. Controlling for hours instead by including hours as a control in a regression on the full sample (column 3) increases slightly the small and medium

differentials as well. Since hours are probably correlated with the error in the log wage regression, the fourth column presents instrumental variables estimates treating hours as endogenous, using per capita household nonlabor income and the number of the individual's children as instruments. In this specification hours have an even stronger negative effect on wages and size coefficients rise; but it should be noted that the first stage fits quite poorly. Since hours considerations fail to explain away size differentials, they, too, are dropped in what follows.

The final compensating differential argument that may be studied using the Peruvian data pertains to familial atmosphere. It is sometimes argued that smaller establishments, which are more often family concerns, provide a more personal environment for which workers are willing to take pay cuts. As shown in the last segment of Table 7, the share of workers with a relative as boss drops from 27 percent among employees in micro establishments to 1 percent in large establishments. Comparison of the first and fifth columns of Table 9 demonstrates, however, that inclusion of an indicator of relative as boss increases the size coefficient slightly. Having a boss as relative furthermore has a positive (though insignificant) effect on wages, suggesting that it may be a better proxy for connections than for familial atmosphere. Restricting the sample to only workers with nonrelatives as bosses (not shown) also has little effect. While no evidence is found that the smallest establishments pay less because they more often involve working with family, it remains possible that micro establishments behave more like family concerns, whether or not employers and employees are related.

The net result of all the direct attempts to control for compensating differences is that the small and medium coefficients have been shown robust,

while the large establishment differential has declined somewhat but continues to be very large. The list of differences has been far from comprehensive. In their study of U.S. size-wage correlations, Brown and Medoff (1989) also control for shift work, overtime, danger, subjective measures of comfort and relationships. Even with this longer list of controls, however, they find that size-wage effects are little affected.

An indirect method of assessing compensating difference arguments involves examining and modifying the job tenure duration analysis presented in the previous section. If workers in larger establishments receive premium remuneration, then they should be more reluctant than others to quit. If, on the other hand, their higher wages merely compensate for inferior working conditions, they may be equally likely to quit. The hazard rate results discussed in Part II thus are consistent with the conclusion that higher wages in small establishments (relative to micro establishments) merely compensate for inferior working conditions, while at least part of the higher wages in medium and large establishments constitute premiums.

Unfortunately, there are at least two problems with this interpretation. First, the hazard rates conflate quit and fire rates. The lower hazards in larger establishments may reflect greater stability of larger, older establishments, and thus lower fire rates. It may also reflect greater enforcement of employment stability legislation. Similarly, the lack of significant differences in hazards across micro and small establishments may obscure higher fire rates in "more capitalist" but neophyte small enterprises, together with lower quit rates induced by premium wages.[11] Second, the assumption of constant inflow into employment for all establishment sizes required to derive the likelihood function for Table 4 results is probably

misleading. It is widely believed that small and informal enterprises have been rising in numbers relative to larger, formal establishments since the mid-1970s in Peru (Romani and Abugattas, p.ii). The small and informal sectors also seem to have maintained their sizes better than the formal sector in the recession that bottomed in 1983, and picked up somewhat more rapidly from then to the interview year. Greater recent inflow into employment in new small establishments would tend to make micro and small establishment spells look shorter even if conditional exit rates were the same across establishment sizes.

A refinement of the inferences may be made by controlling for (the log of) the wage. If higher wages in larger establishments constitute premiums, the wage should have a negative effect on the hazard, and its inclusion should reduce the absolute effects of medium and large establishment sizes on the hazard. If higher wages in general merely reflect the presence of inferior working conditions for which they perfectly compensate, then the wage should have little effect on the hazard and its inclusion should leave size coefficients unaffected. Column 4 of Table 4 shows that the log of the wage reduces job exit rates significantly; its coefficient is an elasticity indicating that a 10 percent increase in the wage would reduce the hazard by nearly 3 percent. The inclusion of ln(wage) also reduces the medium and large size effects, the medium coefficient becoming insignificant. These results are somewhat less vulnerable to the criticisms of the previous paragraph and strengthen the claim that higher wages in medium and large establishments contain some premiums inducing lower quit rates. They are less helpful in aiding interpretation of the result that micro and small hazards are not significantly different.

Table 10 provides a small measure of corroborating evidence for the existence of premiums in larger establishments. The rising age at hire as establishment size rises is consistent with workers moving into superior jobs as opportunities arise for the privileged few. Since the rising age at hire could simply reflect higher education requirements and later entry into the labor market, age at hire differences are examined using regressions controlling for education as well as establishment size. Quite consistent with the duration analysis results, age at hire appears not to be significantly different between micro and small establishments, but is significantly higher in medium and large establishments. A crude indicator of whether the current job is the first job (a worker's report that what they did prior to their main job in the last twelve months was anything other than a job) also declines with size, further corroborating the evidence that individuals move into larger establishments as opportunities arise. On the other hand, workers' reports of search for replacement jobs shows no strong pattern by establishment size. On net, the indirect evidence for the existence of a premium, at least in medium and large establishments, is reasonably compelling.

Before concluding that premiums exist, however, selection arguments must be examined, since some of them are consistent with the duration analysis results as well. In particular, if larger establishments hire more high productivity workers who receive higher wages and who also have inherently lower quit rates, then the wage (which essentially identifies worker quality) would be negatively correlated with the hazard and its inclusion could reduce size effects as seen above. In fact, the productivity dimension of greater importance to larger establishments may be workers' stability itself (for one

such argument, see Mazumdar, 1983).

Selection explanations for wage differences by establishment size are based on the notion that technological and organizational differences associated with larger establishments may lead employers to fill a higher share of positions with higher ability workers, where "ability" is some dimension of workers' productivity not captured by observed education and experience measures. To the extent that the other ability dimensions are correlated with observed measures, the effects have already been picked up. It is perhaps difficult to believe that the uncorrelated components of unobserved ability could explain the large remaining size coefficients, but some attempt should be made to examine this empirically. Not having direct measures of ability in the data at hand, several indirect and incomplete methods of examining this argument must be resorted to.

First, consider differences in ability requirements that are reflected in differences in occupational mix. Controlling for occupation should eliminate such selection bias. Table 11 presents log wage regression establishment size coefficients for several occupational subsamples. The first two columns are broad divisions based on self-reported blue and white collar status.[12] The last four columns pertain to the four specific occupations with highest numbers of observations in the dataset. "Construction workers" is a reasonably homogeneous skill category, not including workers claiming to be supervisors, electricians, painters, stone cutters and the like. "Salesmen," "transport equipment operators," and "bookkeepers" are similarly homogeneous groups. Size effects are jointly significant at least at the five percent level in all occupation groups except transport equipment operators. Size effects are significantly different from

each other for all but construction workers and transport equipment operators. The samples are small and the results are somewhat mixed, but they leave the impression that establishment size effects are substantial even within occupations.

Second, consider diminsions of ability that are shaped by genetics and family upbringing and thus are more closely correlated among brothers than among random pairs of individuals. If all other variables are measured without error and there are no endogeneity problems after controlling for family effects (admittedly big ifs), then regressions run on data differenced between brothers should produce size effect estimates much less biased by selection. The first column of Table 12 presents results from such an exercise. 59 pairs of brothers were identified in the sample of workers excluding those in mining and electricity.[13] The sample is small and the results consequently weak, but they suggest that size effects, at least between micro and all larger establishments, are substantial even when family effects are controlled for. The size variables are jointly just about significant at the 10 percent level. They are not, however, significantly different from one another; and when constrained to be equal they exhibit a coefficient of .309 that is significant at the 5 percent level.

Finally, consider more general dimensions of unmeasured ability, the effects of which may in principle be accounted for if one is willing to make strong assumptions about the distributions governing the selection process. Lee (1983) proposes the most palatable assumptions that lead to tractable methods for dealing with endogenous selection across more than two (here four) categories. Under the assumptions that the selection of workers into establishment size categories may be represented by a multinomial logit model,

and that the error in the log wage regression and an appropriate transformation of the errors underlying the multinomial logit model are joint normally distributed, functional forms may be derived for the expected value of the log wage regression error terms conditional on observed size and workers' characteristics. Much like inverse Mills' ratio terms in models controlling for selection across two categories, these terms (Lamda0-Lamda3 in the second column of Table 12) capture the notion that if, say, higher schooling and higher unobserved ability both make people more likely to work in large establishments, then (holding other observed variables constant) people with lower schooling who are observed to work in large establishments are expected to have greater ability. Because the conditional expectations depend on observed charactersitics in this way, they vary even within size categories and in principle allow separate identification of the effects of establishment size and differences in unobserved ability. Results from such an exercise (column 2 of Table 12) provide little evidence of endogenous selection bias in size coefficient estimates. It should be admitted, however, that they provide little evidence of anything, because the size and selection correction terms are estimated very imprecisely. The estimates are furthermore quite sensitive to the choice of variables included in the first stage multinomial logit model, which at best predicts only a small fraction of the size categories.

A final method for attempting to control for endogenous selection is proposed by Idson and Feaster (1990). It is similar in spirit to the Lee methodology but based on the assumption that the selection across size categories may be captured by an ordered probit model. This assumption is somwhat inferior conceptually to the multinomial logit assumption, because it

constrains the "ability" composite of concern to employers of any size to be the same, larger employers simply requiring more of the composite. It seems likely that the traits that matter to employers in micro, small, medium and large establishments differ, making the more flexible multinomial logit model more attractive. The benefit of the ordered probit is that it has far fewer parameters, providing some hope of greater precision. As seen in the last column of Table 12, this hope is not borne out. Though the point estimates of the size effects are much closer to zero, their standard errors (corrected for inclusion of estimated regressors following Murphy and Topel, 1985) are so large as to admit their usual higher values as highly probable. The selection correction term itself is not significantly different from zero.

Taking all of the selection control attempts together, it seems reasonable to conclude that there is little evidence in favor of selection arguments. If endogenous selection does take place, it must be of a sort that can explain the duration analysis results. That is, it involves selection of more stable individuals, or of higher ability individuals who also happen to be more stable, into larger establishments. There is also some reason to believe that it is more likely important in explaining wage differences between small, medium and large establishments, than between these and micro establishments.

Direct attempts to account for compensating differentials also failed to explain much of the observed size-wage correlation. Thus a reasonable conclusion to draw from this section is that substantial wage premiums exist, at least in medium and large establishments.

IV. Are Premiums Caused by Stronger Unionization Threats or Greater Compliance with Minimum Wage Legislation?

If true premiums exist, understanding why they arise is of obvious importance. The first two possible explanations for premiums have long been given as reasons for the existence of a Harris-Todaro formal sector with high and rigid wages in developing countries, and have the most direct implications for labor market policy. Larger firms may pay more on average either because they are more likely to be unionized or because they are more likely to be punished for evasion of minimum wage laws.[14] Presumably these disadvantages are compensated by cheaper access to capital, greater monopoly power in output markets, or government-bestowed privileges, allowing larger establishments to remain in business. If these explanations are correct and complete, then in the absence of minimum wage laws and organized workers' market power, employers of all sizes would pay identical wages to identical workers.[15]

As seen in Table 13.a, unionization does increase with establishment size. An individual is recorded as being in a unionized job if he reports that there is a "labor union or other workers' organization" at his workplace. The numbers should be treated with care, because by law unions, which are organized by firm and often within single collar status groups, must have 20 members (Gonzales, 1980, p.225). Within-enterprise worker groups with between five and 20 members can elect a representative to negotiate with the employer, but are not officially given union status. Thus the few workers in micro and small establishments coded as unionized are reporting a workplace organization without official union status, or are in establishments that form parts of larger, multi-establishment unionized firms, or have misunderstood the question.

Including an indicator of unionization and observing what happens to the establishment size coefficients allows assessment of a simple story in which unionized firms pay more, and since they constitute a higher fraction of larger establishments, the larger establishments appear to pay more. The connection between unionization, size and wages may be more subtle, however, if larger establishments face greater threats of unionization and pay more to stave it off. If this is the case, the union-nonunion differential should be positive, but smaller in larger establishments. Allowing the union coefficient to differ by establishment size also allows for the possibility that the "average meaning" of the unionization dummy varies by establishment size, as suggested above.

Again as seen in Table 13.a, the incidence of wages below the legal minimum declines as establishment size rises. While this evidence is consistent with the claim of higher minimum wage compliance driving higher wages in larger establishments, it alone sheds relatively little light on the validity of the claim for several reasons. First, it at best indicates the effect of compliance with legislation on the lower tail of the wage distribution. Some would argue that for sociological reasons an increase in wages at the bottom of the wage distribution must be accompanied by increases throughout the distribution to maintain morale. Second, unless the lower incidence of illegally low wages may be associated with some indicator of compliance, they may just arise because larger establishments pay higher wages for other reasons. Thus it would be useful to examine log wage regressions including a proxy for compliance with labor legislation.

The best proxy for compliance with labor legislation available in the dataset is an indicator of whether the individual is entitled to the benefits

of the Instituto Peruano de Seguridad Social (IPSS). As seen in Table 13.a, IPSS incidence also rises with size. IPSS is in principle mandatory and provides health care as well as disability insurance and retirement benefits. It is funded in part by a tax of five percent on employers' wage bills. Compliance with IPSS regulations and minimum wage legislation should be closely related, since signing up for IPSS benefits makes the establishment more vulnerable to monitoring with regard to its other labor practices. "Formality" is often treated as a package deal, employers complying either with all regulations or with none. Some support for using IPSS as a proxy for minimum wage compliance is found in Tables 13.b and 13.c, which show that the incidence of illegally low wages is lower among those reporting IPSS within each size category. It is also reassuring that IPSS and paid holidays - also mandated by labor legislation - are closely correlated. IPSS may be a better proxy for compliance with minimum wage legislation among larger establishments, where the risk of monitoring and punishment is higher overall, and where compliance with one regulation may thus be more closely associated with another. Thus the IPSS effect, too, should be allowed to vary by establishment size. If IPSS captures only compliance with minimum wage legislation, it should enter positively in log wage regressions. To the extent, however, that IPSS provides benefits of some value to workers, for which they are willing to accept lower wages, the coefficients will tend to understate the minimum wage effect.

It is useful to allow size, union and IPSS effects to differ across collar statuses for several reasons. First, as became evident in Table 11, size effects are greater among white collar workers than among blue collar workers. Second, unions are largely organized within collar status groups.

Third, for many decades labor legislation treated blue and white collar workers differently, though by the time of the survey most laws had been unified. The regressions of Table 14 allow for such differences, but attempt to remain parsimonious by restricting coefficients on a set of variables, for which the null hypothesis of equality across collar statuses could not be rejected, to equality across collar groups.

Including only simple indicators of unionization and IPSS entitlement in the second column of Table 14 has little effect. Allowing both effects to vary by establishment size, as in the third column, reveals that while unionization continues to have little importance, IPSS effects are relatively strong among blue collar workers in larger establishments. Size coefficients for blue collar workers remain jointly significantly different from zero but are no longer different from each other, while white collar size coefficients increase slightly, remaining significant and significantly different. These size coefficient effects remain the same when all union terms and white collar IPSS terms (jointly insignificant in column 3) are eliminated (column 4).

Unionization appears to be of little importance in explaining size differentials. The lack of significant union effects is reasonable in light of the nature of the union movement in Peru, and arises even though unions became important influences on labor developments in the years prior to the survey. The national confederations became quite powerful, all the 1985 presidential candidates appealing for their supoprt. But the confederations' focus was on national labor and macroeconomic policies, and on organizing national strikes, rather than on collective bargaining at the plant level, which has a greater tendency to create union-nonunion differentials (see Bollinger, 1987).

The IPSS results admit the interpretation that minimum wage compliance plays an important role in generating wage differences among blue collar workers in small, medium and large establishments, though they do not explain differences between blue collar workers in micro and larger establishments, or among white collar workers. Some additional reason for associating the IPSS effects with minimum wage compliance is found in the last columns of Table 14. One might expect that if wage scales are being pushed up by binding minimum wages, then the increases should be greatest at the lower end of the wage distribution. And indeed, the medium and large IPSS coefficients show some tendency to be greater at the lower end of the wage distribution (column 5's 25th quantile regression) than at the upper end (column 6's 75th quantile regression).

The interpretation of the blue collar IPSS coefficients is subject to some uncertainty. If IPSS and other labor legislation compliance is largely exogenous (determined by location, size, etc., but not correlated with errors in the log wage regression), then the results imply that the elimination of the law would eliminate differentials in wages for blue collar workers in small, medium and large establishments. It remains possible, however, that compliance is endogenous. That is, establishments desiring to pay higher wages (for reasons discussed in the next section) face lower costs of compliance and thus are more likely to comply. Under such circumstances eliminating the law may fail to eliminate differentials, though it may reduce them.

Perhaps the most striking result of this section is the robustness of the large size differentials among white collar workers, which cannot be explained away by traditional institutional explanations. This, together with

the conclusion that premiums exist, suggest that the newer institutional explanations of the next section merit attention.

V. Might One or More "New Institutional" Explanation Lie Behind the Premium?

Most other explanations for payment of premium wages by larger establishments point to more fundamental labor market malfunctions, which cause employers of different sizes to pursue different wage and employment strategies even in the absence of government intervention. A first set of explanations, often cast as efficiency wage models, provides reasons why larger employers are more likely to choose to suffer higher wages in order to reap higher productivity. Paying wages above market-clearing levels may induce workers to work carefully and hard (not to shirk) out of fear of losing future premiums (see, e.g., Bulow and Summers, 1986); and the benefits of such a strategy may be more likely to exceed the costs in larger establishments if their technology and organization are such that direct monitoring and supervision are more difficult or the costs of shirking are greater. Higher wages may also increase productivity by reducing downtime and training costs, through reducing turnover and absenteeism (Stiglitz, 1974) or through increasing applications (Lang, 1991, Montgomery, 1991). Presumably such considerations are more important for larger establishments if more capital remains idle per vacancy or if their more complicated technology implies higher training costs. High wages may also improve the average quality of the applicant pool (Weiss, 1980), or improve worker productivity by increasing norms of production (Akerlof, 1982) or by improving esprit de corps within the organization (Clague, 1977). Efficiency wages for improvement of the applicant pool explain higher wages in larger establishments if larger employers have greater need for high quality workers but cannot distinguish

workers by quality (unlike in the simple screening argument considered above). Higher wages to raise norms or esprit de corps may be more valuable in larger establishmsents, if such efforts are substitutes for establishing personal ties with workers, easier to achieve in smaller establishments.

All of these arguments, with the likely exception of the norms and esprit de corps arguments, would also provide employers with reason to promise workers upward-sloping wage-tenure profiles. Such positive dependence of wages on tenure allows lifetime premiums sufficient to prevent shirking or quitting, and to attract high quality applicants, to be achieved at lowest lifetime cost to the employers. Thus if these arguments are to explain size-wage correlations, one might expect to see stronger wage-tenure profiles in larger establishments. Table 15 provides some reason to believe wages do rise more rapidly with tenure in medium and large establishments than in micro and small establishments. When the tenure effect is undifferentiated by size in a log wage regression on the entire sample (column 1), the apparent effect is small by U.S. standards. When, however, the tenure effect is differentiated by establishment size, the effects are negligible in micro and small establishments and closer to the magnitudes found in U.S. data in medium and large establishments. When tenure effects are differentiated by collar status as well as size in regressions like those in Table 14, the hypothesis that the tenure-size interaction terms are equal across collar statuses cannot be rejected. The third and fourth columns of Table 15 present results from such a regression when tenure-size interactions are constrained to be equal across collar groups. They tell a similar story to that of column 2. In general, despite significant tenure effects, size coeffcients change very little, indicating that premiums remain.

Since the wage-tenure associations are drawn entirely from cross section analysis, one must consider the possibility that they reflect higher wages throughout their tenure for people who tend to remain on the job longer, rather than wages that rise with tenure for individuals. If large establishments pay premium wages to only some of their workers, or if they hire more high quality, stable individuals who are paid more from the outset, such cross section patterns might arise. A simple attempt to control for the endogeneity of the tenure variables is presented in the last two columns of Table 15, where experience (which is experience at hire plus tenure) and its square, and the tenure variables are instrumented by experience at hire, its square, the number of adults in the household at time of hire and the individual's number of children at time of hire, each fully interacted with the size categories.[16] Surprisingly the tenure coefficients increase slightly, but with the exception of the ill-identified large establishment tenure coefficient they retain the pattern of steeper wage-tenure profiles in larger establishments. Table 15 is thus at least weakly consistent with some efficiency wage stories.

Of the efficiency wage stories, the most popular but least compelling in the Peruvian context is the shirking-based efficiency wage argument. Higher wages produce lower shirking only if combined with a credible threat that shirkers can be caught with some probability and that they will be dismissed if caught. Employment stability legislation in Peru, however, makes dismissal very costly and difficult, and thus reduces greatly the credibility of such dismissal threats, especially in large establishments. In the years prior to the survey employers were allowed to fire workers freely only during their first three months on the job. After three months, workers had to be given

three months' notice or three months' wages as indemnity for dismissal; and after three years on the job it became virtually impossible to fire workers without "grave cause," which was difficult to establish. Alan Garcia came into office in mid-1985 promising to reduce the probationary period before workers were vested with "absolute stability" from three years to three months, and such legislation was implemented by mid-1986. Thus firing workers became even more difficult. Two other observations weaken the shirking-based efficiency wage argument. First, the most significant size effects are observed between micro and small establishments. If Peruvian firm studies are any guide, the mean number of workers among micro establishments is three or four, while among small establishments it is around eight (Villaran, et al., 1988, p.11). It seems unlikely that monitoring would become so much more difficult at that level. Second, Table 16 reveals some differences in size effects across sectors that are difficult to reconcile with the shirking story. For example, size effects are greater in construction than manufacturing, though one might expect that workers' responsibilities and monitoring practices differ more with establishment size in manufacturing.

Screening-based efficiency wage stories are called into question by the inability to find evidence of endogenous sample selection in Section III. The payment of high wages to improve the applicant pool when employers cannot distinguish worker types should not itself produce patterns of endogenous selection; but one might expect the two processes to go hand in hand. Employers would be interested in avoiding premiums by doing as much screening as possible. Presumably after workers had been on the job a while employers would also be able to differentiate wages by skill, and such differentiation should also be discernible in tests for endogenous sample selection.

Morale and esprit de corps efficiency wage stories are consistent with widespread size effects across occupations and industries. They lead to no obvious explanation of differences in size effects across industries or occupations, however, and perhaps have greater difficulty explaining wage differences between medium and large establishments. One might expect that the ability to maintain personal ties with workers, the substitute for efficiency wages, is exhausted at fairly small sizes.

The most compelling efficiency wage stories concern the maintenance of a steady workforce through reduction of turnover and absenteesim, and perhaps through the increase of applications. Several decades ago western firms' frequent complaint about operating in developing countries was the instability or fickleness of the workforce. Those complaints seem to have died down, perhaps because firms discovered wage and employment strategies to deal with the problem. Secondary reports reflect steady increases in capital per worker as establishment size rises, at least in manufacturing (Villaran, et al., 1988), allowing these stories to explain size-wage correlations throughout the size distribution. This story may also be more readily reconcilable with greater size effects in construction than in manufacturing, since it is quite possible that the importance of timeliness and coordination rises more with establishment size in construction.

A final set of explanations for the existence of wage premiums in larger establishments arise because larger firms are in a privileged position with respect to small firms. Insider-outsider models (Lindbeck and Snower, 1986) suggest that when total rents are higher, workers are able to obtain higher rents through bargaining (not necessarily through unions). Simple political economy might also suggest that large firms might be willing to pay higher

wages in order to remain in the government's good graces, so that privileges continue. These arguments seem most relevant, however, in explaining higher wages in only a few truly large establishments in Peru, and again have some difficulty in explaining wage differences throughout the size distribution.

VI. Conclusion

Larger employers provide substantially better wage, benefits and job security packages in Peru, even after controlling for differences in standard measures of their workers' productive characteristics. Part III found little evidence that higher wages merely compensate for inferior working conditions, or that higher wages are merited by workers of higher quality along unobserved dimensions. It led to the conclusion that substantial premiums exist, at least in medium and large establishments. The analysis of Part IV found that greater unionization or greater threat of unionization in larger establishments does not explain the observed large employer premiums. Greater compliance with minimum wage legislation appears to be associated with higher wages in larger establishments among blue collar workers, though the relationship may be a subtle one. Employers interested in paying higher wages for other reasons may find compliance less costly and thus choose to comply with labor legislation. Part V briefly discussed several more recent explanations for higher wages in larger establishments, some or all of which must be relevant in explaining white collar size differentials, and which may interact with compliance considerations in explaining blue collar size differentials. Few direct tests are possible using the data at hand, but several common sense arguments suggest that the most plausible of these explanations concern the greater interest of larger establishments in guaranteeing continuity in operations. Higher capital per worker and other

differences in technology and organization may render them more interested in reducing turnover and absenteeism, and perhaps increasing applications, by offering wages above market-clearing levels.

A final observation is that a number of analyses suggested that the explanations for wage differences between micro and all other establishments might be quite different from the explanations relevant to wage differences between small, medium and large establishments. It remains possible that micro establishments indeed constitute an informal sector, which operates quite differently from the larger, possibly more capitalist enterprises. Duration analysis suggests that either working conditions differ substantially between micro and small establishments (and presumably medium and large establishments as well) or that premiums unexplained by working condition differences exist but that higher fire rates in small establishments counteract lower quit rates induced by the premiums. (This would imply an informal sector with work-sharing or other mechanisms to increase tenure, rather than one with more tenuous employment relationships.) Endogenous selection correction results suggest that if selection effects are important, they are important in explaining differences between small, medium and large establishments, and not between micro and small establishments. Compliance with minimum wage legislation furthermore seems to explain differences between small, medium and large establishments, but not between micro establishments and the rest. Dividing establishments arbitrarily into only two groups, formal and informal, on the basis of size might have obscured some of this complexity. But if a single line is to be drawn between formal and informal sectors, it appears to belong between micro and small establishments, rather than at any higher size level.

References

Abraham, K. and H. Farber (1987), "Job Duration, Seniority and Earnings," American Economic Review 77(3),pp.278-297.

Akerlof, G. A. (1982), "Labor Contracts as Partial Gift Exchange," Quarterly Journal of Economics 97, pp.543-569.

Bollinger, W. (1987), "Peru," in G. Greenfield and S. Maram (eds.), Latin American Labor Organizations, New York: Greenwood Press, pp.607-666.

Brown,C. and J. Medoff (1989), "The Employer Size-Wage Effect," Journal of Political Economy 97(5), pp.1027-1059.

Bulow,J. and L. Summers (1986), "A Theory of Dual Labor Markets with Application to Industrial Policy, Discrimination, and Keynesian Unemployment," Journal of Labor Economics 4(3), pp. 376-413.

Calvo,G.(1978), "Urban Unemployment and Wage Determination in LDCs: Trade Unions in the Harris-Todaro Model," International Economic Review 19(1), pp.65-81.

Carbonetto, D., J. Hoyle and M. Tueros (1988), Lima: Sector Informal, Lima: Letra S.R.L.

Clague, C. (1977), "Information Costs, Corporate Hierarchies, and Earnings Inequality," American Economic Review Papers and Proceedings 67(1), pp.81-85.

de Soto, H. (1989), The Other Path: the Invisible Revolution in the Third World, New York: Harper and Row.

Fields, G. (1975), "Rural-Urban Migration, Urban Unemployment and Underemployment, and Job-Search Activity in LDCs," Journal of Development Economics 2, pp.165-187.

Glewwe, P. (1988), The Distribution of Welfare in Peru 1985-86. Living Standards Measurement Survey Working Paper No. 42, Washington, D.C.: The World Bank.

Gonzales, G. (1980), Neuvo Derecho Peruano del Trabajo, Lima: Libreria Studium.

Harris, J. and M. Todaro (1970), "Migration, Unemployment and Development: A Two-Sector Analysis," American Economic Review, pp.126-142.

Idson,T. and D. Feaster (1990), "A Selectivity Model of Employer-Size Wage Differentials," Journal of Labor Economics 8(1), pt.1, pp.99-122.

Kannappan, S. (1988), "Urban Labor Markets and Development," World Bank Research Observer 3, pp.189-206.

Lancaster, T (1990), The Econometric Analysis of Transition Data, Cambridge: Cambridge University Press.

Lang, K. (1991), "Persistent Wage Dispersion and Involuntary Unemployment," Quarterly Journal of Economics, pp.181-202.

Little, I.M.D., D. Mazumdar and J.M. Page (1987), Small Manufacturing Enterprises: A Comparative Study of India and Other Economies, New York: Oxford Univeristy Press.

Lee, L.-F. (1983), "Generalized Econometric Models with Selectivity," Econometrica 51(2), pp.507-511.

Lindbeck, A. and D. Snower (1986), "Wage Rigidity, Union Activity and Unemployment," in W. Beckerman (ed.), Wage Rigidity and Unemployment, Baltimore: Johns Hopkins, pp.97-125.

Mazumdar, D. (1983), "Segmented Labor Markets in LDCs," American Economic Review 73(2), pp.254-259.

Montgomery, J.D. (1991), "Equilibrium Wage Dispersion and Interindustry Wage Differentials," Quarterly Journal of Economics, pp.163-179.

Murphy, K. and R. Topel (1985), "Estimation and Inference in Two-Step Econometric Models," Journal of Business and Economic Statistics, Vol 3., No. 4, pp.370-379.

Peattie, L. (1987), "An Idea in Good Currency and How it Grew: The Informal Sector," World Development 15, pp.851-860.

Romani, H. and L. Abugattas (between 1986 and 1990), Ensayos sobre el Desarollo Industrial en el Perú, Lima: Asociación Acción y Pensamiento Democrático.

Stiglitz, J.E. (1974), "Alternative Theories of Wage Determination and Unemployment in LDC's: The Labor Turnover Model," Quarterly Journal of Economics, pp.194-227.

Villaran, F., S. Alvarez and C. Mendoza (1988), Perspectivas del Desarrollo de la Pequeña y Micro Industria en un Contexto de Crisis Económica, Lima: Centro de Ingeniería para el Desarrollo Laboral - CIPDEL.

Vuong,Q. (1989), "Likelihood Raio Tests for Model Selection and Non-Nested Hypotheses," Econometrica 57(2), March, p.307-333.

Weiss, A. (1980), "Job Queues and Layoffs in Labor Markets with Flexible Wages," Journal of Political Economy 88, pp.526-38.

Appendix

Variable Descriptions

Variable Name	Description
Cash Wage	Real hourly cash wage for main job in last seven days in June 1985 Intis. Deflator calculated by Glewwe (1988) pertains to global expenditures in last three months, and varies by region and date of interview.
Total Wage	Real hourly wage including bonuses and cash value of in-kind payments for main job in last seven days in June 1985 Intis. Deflator same as for Cash Wage. In-kind payments include those in meals, groceries, housing, clothing and transport. For 26 observations with positive but missing values for at least one in-kind payment, missing pieces were filled in using predicted values from regression on schooling, experience and its square and industry indicators, corrected for endogenous selection into the positive in-kind payment sub-sample. Results are little affected by setting these values to zero or omitting these observations.
(Potential) Experience	Age minus highest grade attained minus years of school repeated minus five.
Tenure	Months or years (as indicated) worked in main job in last seven days. Tenures of x months plus 2 or more weeks rounded to x+1. Tenures of x months plus 1 week rounded to x, except when x=0, in which case tenure is rounded to 1 month. Tenure results little affected by omission of the 1 week tenure observations.
Years Primary	Highest grade attained in primary schooling, ranging from zero to six.
Years Secondary	Highest grade attained in secondary schooling, ranging from zero to five.
Years Post-Secondary	Years of post-secondary schooling, ranging from zero to five.
Hours	Usual weekly hours worked in main job of last seven days.
Boss-is-Relative	Indicator equal to one if respondent answers yes to: Is your employer, manager or director related to you?
Age at Hire	Age minus tenure in years.

Numbers of Other Males and Females at Time of Hire	Numbers of males other than the individual, and females in the individual's household, who were: at least 15 years of age at the individual's time of hire for relatives of the individual; and at least the Peruvian medium age at marriage for men (25.7) or for women (23.2) at the individual's time of hire for spouses of the individual or his relative.
Number of the Individual's Children at Time of Hire	Number of the individual's children, either living in the household or living elsewhere and below 30 years of age at the interview date, whose age minus the individual's tenure in years is between 0 and 14.
Household Nonlabor Income Per Capita	Total real annual nonlabor income of the household, including retirement pension, unemployment benefits, food pension, alimony, children's allowance, medical or life insurance payments, income from non-profit organizations, interest on savings, dividends or profit shares, rental income, winnings from gambling, interitances, and other miscellaneous income, in June 1986 Intis, divided by the number of household members. For deflation purposes, missing month of income report was replaced by 6 months prior to date of interview (four households affected). The deflator is the same as for Cash Wage, except where the month referred to is before the earliest availability of the deflator, in which case the deflator for the appropriate region was extended back using the consumer price index from the IMF's International Financial Statistics. Positive but missing values of interest income were imputed using savings stock at time of interview and three-month CD interest rate from month of interview (42 observations affected). Positive but missing values for other types of nonlabor income were replaced by predicted values of regressions on household head's schooling, experience and its square, numbers of children and adults in the household, whether the household head was born in the countryside, years of schooling of the household head's father, and an indicator of residence in Lima (13 observations affected).

Notes

[1] The CPS regression differs from the Peruvian sample in two ways. First, it does not exclude employees who work alone in their workplaces (a very small number in Peru). Second, the CPS wage is gross of taxes and social security payments, while the Peruvian wage is more nearly take-home pay.

[2] In all relations in Tables 2 and 3, except the nintieth quantile regression, the hypothesis that the coefficients on small, medium and large indicators are equal may be rejected soundly. In most relations the hypothesis that the coefficients on primary, secondary and post-secondary schooling are equal is just barely rejected.

[3] See Lancaster (1990) for an exposition of duration analysis.

[4] One restriction implied by this form is that the hazard is monotonic in elapsed duration, increasing, constant or decreasing as α is greater than, equal to, or less than one. Job tenure hazards are often found to rise first and then to decline with elapsed tenure. A hazard with this characteristic is the log-logistic, of the form

$$h(t;X) = \frac{\exp(X^T\beta)\alpha t^{\alpha-1}}{1+t^{\alpha}\exp(X^T\beta)},$$

when $\alpha>1$. This assumption was also employed, but a likelihood ratio test for non-nested model selection suggested by Vuong (1989) rejects the null hypothesis that the models are equally far from the true model (by the Kullback-Liebler Information Criterion) in favor of the alternative that the Weibull model is closer to the true model.

[5] For the Weibull model, the likelihood contribution for tenure to date (s) under constant inflow is:

$$g(s) = \frac{\exp[-\exp(X^T\beta)s^{\alpha}]}{\exp(-X^T\beta/\alpha)\Gamma(1+1/\alpha)},$$

where $\Gamma(.)$ is the gamma function. For the log-logistic model (see footnote 5) the likelihood contribution is:

$$g(s) = \frac{\alpha\exp(X^T\beta/\alpha)}{(1+s^{\alpha}\exp(X^T\beta))\Gamma(1/\alpha)\Gamma(1-1/\alpha)}.$$

[6] A likelihood ratio test rejects the hypothesis that the coefficients on small, medium and large are equal.

[7] It is possible that worker and job heterogeneity are such that some worker-job matches are more productive than others, inducing wage variation. If in larger enterprises it is easier to shuffle workers and jobs to find superior matches, then again employer size-wage correlations could arise that do not reflect true premiums. Unfortunately, no direct test of this is possible with available data. It seems unlikely that shuffling into somewhat better matches could account for employer size-wage differences of the great magnitudes described below.

[8] This is a pure compensating differential argument only if employers may differentiate wages across otherwise identical workers who reside at different distances. If employers were constrained for sociological or legal reasons to pay all workers the same wage, then the need to compensate the marginal worker for transport costs would lead to wage premiums for inframarginal workers. Peruvian employers appear to make some effort to differentiate wages. The shares of workers in micro, small, medium and large establishments reporting receipt of some in-kind payment related to transport are .07, .10, .19 and .33.

[9] The likelihood contribution for this standard endogenous sample selection model is:

$$\frac{1}{\theta\sigma_M+(1-\theta)\sigma_W}\,\phi\left(\theta\left(\frac{M-Z\delta_M}{\sigma_M}\right)+(1-\theta)\left(\frac{W-Z\delta_W}{\sigma_W}\right)\right)$$

$$\cdot\ \Phi\left((-1)^{(1-\theta)}\frac{Z\gamma+\theta\rho_M\left(\frac{M-Z\delta_M}{\sigma_M}\right)+(1-\theta)\rho_W\left(\frac{W-Z\delta_W}{\sigma_W}\right)}{(1-\theta\rho_M^2\ -\ (1-\theta)\rho_W^2)^{.5}}\right),$$

where $\phi(.)$ and $\Phi(.)$ are standard normal density and distribution functions, σ_M and σ_W are the standard deviations of the errors in the M and W relations, and ρ_M and ρ_W are the correlation coefficients between the errors in those equations and the error in the θ^* relation.

[10] The entire model may be estimated consistently in at least three ways. A three-stage procedure, in which the first stage of the text is estimated in two stages using standard endogenous sample selection techniques, produces very similar results. Full maximum likelihood for the entire model is also possible in principle, but in practice failed to converge.

[11] At least one characterization of micro and small enterprises in Peru makes the distinction that micro establishments function for subsistence maintenance, while small enterprises are more like young but growing businesses (Villaran, et al, 1988, pp.8-9).

[12] Historically, labor legislation differed for "obreros" (blue collar) and "empleados" (white collar), a legal distinction in Peru. In recent years most legislation has been unified.

[13] There were 52 groups of exactly two brothers, 5 groups of three from each of which two were chosen randomly, and 2 groups of four which were broken randomly into two pairs each.

[14] Harris and Todaro themselves raised the minimum wage story. For a Harris-Todaro model with high wages induced by union bargaining, see Calvo (1978).

[15] Normative conclusions about the merits of such policies must be based on a broader set of considerations, of course.

[16] The inclusion of an expected completed job duration estimate based on duration analysis results, following Abraham and Farber (1987), was also attempted, but produced very weak results.

Table 1

Descriptive Statistics by Establishment Size
(Standard Errors in Parentheses)

	All	Micro	Small	Medium	Large
a. Entire Sample					
No. Observations	1133	346	304	318	165
Percent	100.0	30.5	26.8	28.1	14.6
Mean Cash Wage	4.63	2.92	4.25	5.71	6.85
	(6.17)	(3.24)	(5.22)	(6.95)	(9.14)
Mean Total Wage	5.98	3.56	4.92	6.99	11.06
	(13.29)	(3.86)	(5.71)	(8.72)	(30.64)
90th Quantile Total Wage	11.30	6.29	9.46	14.50	15.45
75th Quantile Total Wage	6.17	4.23	5.66	7.27	9.85
Median Total Wage	3.62	2.52	3.32	4.37	6.13
25th Quantile Total Wage	2.27	1.71	2.30	2.96	3.61
10th Quantile Total Wage	1.59	1.01	1.69	2.03	2.18
Mean Highest Grade Attained	9.40	8.74	9.06	9.82	10.62
	(3.53)	(3.23)	(3.53)	(3.70)	(3.39)
Mean Potential Experience	18.68	17.01	17.84	19.96	21.26
	(14.15)	(15.53)	(12.92)	(13.93)	(13.17)
Percent Household Head	53	44	51	56	67
b. Subsample with Nonmissing Benefits Indicators					
No. Observations	1121	343	298	316	164
Percent with Paid Holiday	51	20	48	73	80
Percent with Sick Leave	48	17	42	71	80
Percent with Pension	45	13	38	68	81
c. Subsample with Nonmissing Tenure					
No Observations	1132	346	304	317	165
Mean Tenure in Years	6.45	5.33	5.60	6.98	9.31
	(8.01)	(7.39)	(7.14)	(8.57)	(8.87)
Median Tenure in Years	3.00	2.00	2.71	3.00	6.00

Table 2

OLS and Quantile Log Wage Regression Coefficients
(Standard Errors in Parentheses)

Dependent Variable	Ln (Cash Wage)	Ln (Total Wage)	Ln (Total Wage)	Ln (Total Wage)	Ln (Total Wage)
Method	OLS	OLS	10th Quantile	Median	90th Quantile
No. Observations	1133	1133	1133	1133	1133
Constant	-.454**	-.131	-1.136**	-.189	.828**
	(.147)	(.154)	(.260)	(.169)	(.320)
Years Primary	.051**	.048*	.063	.068**	-.000
	(.025)	(.026)	(.041)	(.028)	(.053)
Years Secondary	.114**	.102**	.103**	.083**	.148**
	(.011)	(.012)	(.021)	(.013)	(.025)
Years Post-Secondary	.148**	.150**	.111**	.149**	.224**
	(.014)	(.015)	(.028)	(.016)	(.030)
Experience	.042**	.036**	.0462**	.040**	.042**
	(.005)	(.006)	(.010)	(.006)	(.011)
Experience^2/100	-.047**	-.039**	-.061**	-.048**	-.044**
	(.009)	(.010)	(.015)	(.011)	(.020)
Household Head	.173**	.168**	.151	.166**	.157
	(.053)	(.055)	(.099)	(.061)	(.102)
Small	.280**	.239**	.377**	.175**	.060
	(.052)	(.054)	(.096)	(.060)	(.108)
Medium	.416**	.400**	.498**	.341**	.189*
	(.052)	(.054)	(.096)	(.060)	(.109)
Large	.451**	.546**	.725**	.532**	.284**
	(.064)	(.067)	(.124)	(.074)	(.131)
R-Squared	.400	.366			
Adj. R-Squared	.395	.361			
Pseudo R-Squared[1]			.184	.217	.247

* Significant at 10 percent level.

** Significant at 5 percent level.

[1]Calculated as one minus the ratio of the sum of weighted absolute deviations about the estimated quantile to the sum of weighted absolute deviations about the raw quantile, where the weights are given by 2q if the deviation is positive and 2(1-q) otherwise, q being a fraction representing the quantile to be estimated.

Table 3

Benefits Profit Coefficients
(Standard Errors in Parentheses)

Dependent Variable	Paid Holiday	Sick Leave	Pension
No. Observations	1127	1130	1128
Constant	-2.118**	-2.457**	-3.115**
	(.322)	(.336)	(.354)
Years Primary	-.007	.027	.067
	(.053)	(.054)	(.056)
Years Secondary	.129**	.144**	.131**
	(.025)	(.025)	(.027)
Years Post-Secondary	.039	.032	.051
	(.030)	(.031)	(.031)
Experience	.080**	.074**	.082**
	(.112)	(.012)	(.013)
Experience^2/100	-.105**	-.094**	-.095**
	(.021)	(.023)	(.022)
Household Head	.158	.226**	.266**
	(.111)	(.111)	(.114)
Small	.703**	.688**	.824**
	(.111)	(.114)	(.123)
Medium	1.318*	1.363**	1.563**
	(.113)	(.114)	(.123)
Large	1.438**	1.590**	1.884**
	(.145)	(.148)	(.157)
Log-Likelihood	-589.65	-578.30	-531.46

* Significant at 10 percent level.

** Significant at 5 percent level.

Table 4

Constant Inflow Weibull Monthly Tenure Distribution
Hazard Function Estimates[1]
(Standard Errors in Parentheses)

No Observations	1132	1132	1132	1132
Constant	-1.128** (.196)	-1.375** (.205)	-1.378** (.211)	-1.686** (.213)
Years Primary	.106** (.035)	.097** (.036)	.100** (.037)	.116** (.038)
Years Secondary	.059** (.012)	.047** (.013)	.052** (.014)	.081** (.014)
Years Post-Secondary	.002 (.016)	-.012 (.016)	-.005 (.016)	.042** (.018)
Experience At Hire	-.004 (.005)	-.007 (.006)	-.007 (.007)	.002 (.006)
Experience^2/100	.024** (.013)	.020* (.014)	.021* (.014)	.010 (.013)
Small	-.056 (.057)	-.066 (.060)	-.074 (.063)	.005 (.063)
Medium	-.179** (.059)	-.195** (.063)	-.216** (.007)	-.058 (.071)
Large	-.424** (.079)	-.402** (.083)	-.417** (.090)	-.219** (.093)
Family and Household Vars.[2]	no	yes	yes	yes
Industry Indicators[3]	no	no	yes	yes
Ln (Total Wage)				-.300** (.035)
Alpha	.504** (.032)	.547** (.035)	.562** (.036)	.594** (.036)
Log-Likelihood	-5882	-5845	-5836	-5796

* Significant at 10 percent level.

** Significant at 5 percent level

[1]Variables normalized so that baseline hazard refers to person with average schooling, experience and wages, who is in the manufacturing sector and in a micro establishment.

[2]The variables are: number of other adults in the household at the individual's time of hire, and number of the individual's children at time of hire (see appendix).

[3]Industries controlled for are: agriculture, mining, manufacturing, electricity, construction, commerce, transportation, finance and services.

Table 5

Mean Wages by Industry and Establishment Size
(Number of Observations in Parenthesis)

	All	Micro	Small	Medium	Large
Agriculture	3.59 (62)	1.87 (22)	4.01 (18)	4.97 (20)	5.08 (2)
Mining	11.77 (20)	3.01 (1)	3.68 (1)	6.89 (7)	16.41 (11)
Manufacturing	5.14 (354)	3.01 (71)	3.90 (93)	6.16 (123)	7.24 (67)
Electricity	10.33 (8)	• (0)	• (0)	5.49 (3)	13.2 (5)
Construction	4.22 (107)	2.80 (41)	4.46 (28)	5.58 (34)	5.54 (4)
Commerce	7.11 (253)	4.09 (104)	5.80 (74)	7.82 (53)	24.03 (22)
Transportation	6.07 (98)	4.44 (40)	4.47 (20)	6.64 (20)	10.83 (18)
Finance	9.05 (85)	4.38 (11)	6.06 (23)	12.38 (25)	10.47 (26)
Service	5.51 (146)	3.72 (56)	5.85 (47)	7.74 (33)	6.54 (10)

Table 6

Log Total Wage Regression Coefficients[1]
(Standard Errors in Parentheses)

Sample	All	All Except Min./Elec.	All Except Min./Elec.	All Except Min./Elec.	Walk/Motor Except Min./Elec.	Walk/Motor Except Min./Elec.[2]
No. Observations	1133	1105	1105	1105	1180	1080
Small	.239**	.238**	.258**	.246**	.241**	.334**
	(.054)	(.054)	(.054)	(.054)	(.054)	(.071)
Medium	.400**	.395**	.430**	.399**	.399**	.457**
	(.054)	(.055)	(.056)	(.055)	(.056)	(.065)
Large	.546**	.489**	.514**	.462**	.455**	.518**
	(.067)	(.069)	(.071)	(.071)	(.071)	(.075)
Industry Controls	no	no	yes	yes	yes	yes
Lima				.236**	.259**	.477**
				(.044)	(.045)	(.117)
θ*E(M|Z,θ=1)						-.012**
						(.005)
(1-θ)*E(W|Z,θ=1)						.003
						(.009)
λ						.467**
						(.168)
R-Squared	.366	.356	.365	.381	.380	.404
Adj. R-Squared	.361	.351	.356	.372	.381	.393

* Significant at 10 percent level.

** Significant at 5 percent level.

[1]Regressions also control for years of primary, secondary and post-secondary schooling, experience and its square, and household headship.

[2]Two-stage estimates of model controlling for endogeneity of choices over whether to use motor transport and distance between residence and work in time spent in motor transport or on foot. The last three regressors are estimated using results of a maximum likelihood first stage. For details see text and Table 8. Standard errors corrected for presence of estimated regressors following Murphy and Topel (1986).

Table 7

Various Statistics by Establishment Size[1]

	All	Micro	Small	Medium	Large
a. Number of Workers (Percent in Parentheses)					
Lima	713	185	181	216	131
	(100.0)	(26.0)	(25.4)	(30.3)	(18.4)
Other Urban	420	161	123	102	34
	(100.0)	(38.3)	(29.3)	(24.3)	(8.1)
b. Percentage Distribution by Mode of Transport					
Public	57.3	46.5	60.9	61.3	65.5
Private Motor	10.8	4.0	6.6	14.8	24.8
Private Nonmotor	1.1	2.0	0.3	1.3	0.0
Walk	29.7	46.0	31.9	21.4	7.9
Other	1.1	1.5	0.3	1.3	1.8
All	100.0	100.0	100.0	100.0	100.0
c. Commute Time in Minutes, Entire Sample, 1132 Obs.					
Mean	33.7	29.4	36.0	35.9	34.4
(Std. Dev.)	(32.2)	(28.6)	(36.1)	(33.6)	(28.3)
Median	30	20	25	30	30
d. Commute Time in Minutes, Public or Private Motor Transport, 770 Obs.					
Mean	40.8	41.6	45.4	39.2	36.2
(Std. Dev.)	(32.6)	(32.7)	(39.4)	(27.6)	(28.8)
Median	30	30	30	30	30
e. Commute Time in Minutes, Walking, 337 obs.					
Mean	17.2	17.4	16.9	17.8	14.2
(Std. Dev.)	(17.5)	(16.5)	(15.2)	(23.5)	(8.6)
Median	15	15	15	15	15

f. Usual Weekly Hours					
Mean	48.9	48.1	49.3	48.5	50.3
(Std. Dev.)	(15.8)	(17.8)	(16.6)	(14.1)	(13.7)
Median	48	48	48	48	48
g. Percent Part-Time (Usual Weekly Hours Less Than 35)					
	11.6	17.3	12.5	8.5	3.6
h. Percent Searching for Additional Work					
	12.4	16.2	11.8	10.1	10.3
i. Percent with Relative as Boss, 1132 Obs.					
	12.7	26.7	9.9	6.3	1.2

[1]Unless otherwise noted, out of 1133 observations.

Table 8

Estimates of Joint Maximum Likelihood
Model of Transport Mode, Time in Motor Transport
and Time Walking[1]
(Standard Errors in Parentheses)

Dependent Variable	Motor Transport Indicator	Time in Motor Transport	Time Walking
Small	.297**	10.81**	-4.77**
	(.095)	(3.01)	(2.28)
Medium	.312**	8.87**	-5.92**
	(.109)	(3.61)	(2.53)
Large	.457**	9.15*	-5.75
	(.229)	(5.73)	(5.87)
Sigma[2]		37.30**	23.85**
		(.62)	(.64)
Rho[3]		.994**	-1.00**
		(.004)	(.000)

* Significant at 10 percent level.

** significant at 5 percent level.

[1]All equations originally controlled for years of primary, secondary and post-secondary schooling, experience and its square, household headship, industry, residence in Lima, household nonlabor income per capita, numbers of adult males and females in household, number of the individual's children ages 14 or less, and an indicator of whether the individual ever migrated. In each equation, any of these variables with coefficients insignificant at the 20 percent level in consistent two-stage estimation were dropped from the maximum likelihood specification. Coefficient on industry indicators not significantly different from one another were also constrained to be equal. There are 1080 observations and the log-likelihood is -5321.88.

[2]Standard deviation of error in indicated equation.

[3]Correlation coefficient between error in indicated equation and error in the latent motor transport propensity equation (column 1).

Table 9

Log Total Wage Regression Coefficients[1]
(Standard errors in Parentheses)

Sample	Full[2]	Usual Weekly Hours ≥ 35	Full	Full	Boss-is-Relative Nonmissing
No. Observations	1105	974	1105	1105	1104
Method	OLS	OLS	OLS	IV[3]	OLS
Small	.246**	.223**	.259**	.287**	.259**
	(.054)	(.054)	(.051)	(.051)	(.054)
Medium	.399	.409**	.410**	.436**	.416**
	(.055)	(.055)	(.053)	(.065)	(.056)
Large	.462**	.513**	.494**	.563**	.486**
	(.071)	(.069)	(.068)	(.092)	(.072)
Hours			-.013**	-.041**	
			(.001)	(.018)	
Boss-is-Relative					.094
					(.063)
R-Squared	.381	.419	.435	.185	.381
Adj. R-Squared	.372	.410	.426	.172	.371

* Significant at 10 percent level.

** Significant at 5 percent level.

[1]Regressions also control for years of primary, secondary and post-secondary schooling, experience and its square, household headship, industry and residence in Lima.

[2]The full sample excludes workers in mining and electricity.

[3]Instruments include: household nonlabor income per capita, number of other adult males in household, number of adult females in household, and number of the individual's children ages 14 or less.

Table 10

Various Statistics by Establishment Size

	All	Micro	Small	Medium	Large
No. Observations	1132	346	304	317	165
a. Age at Hire					
Mean Actual	26.9	25.8	26.6	28.1	27.9
Median Actual	24.0	21.8	24.4	26.0	25.0
Mean Predicted[1]		24.1	24.8	26.2	26.4
Prediction Std.Err.		0.59	0.63	0.63	0.83
t-ratio on Size Coefficient			0.81	2.61	2.37
b. Percent Not Reporting Previous Occupation (Proxy for First Job):					
	35.7	45.1	34.2	31.2	27.3
c. Percent Reporting Search for Replacement Job:					
Actual	9.0	9.2	10.2	7.6	9.1
Predicted[1]	-	9.6	10.4	7.4	8.8
Prediction Std. Err.		10.2	10.5	11.4	14.4
t-ratio on Size Coefficient			.324	.97.0	.289

[1]Predictions for individual with median schooling (10 years), based on regression or profit controlling for years of primary, secondary and post-secondary schooling, as well as establishment size.

Table 11

Descriptive Statistics and Log Total Wage Regression Coefficients
for Several Occupation Groups[1]
(Standard Errors in Parentheses)

Sample	Blue Collar	White Collar	Sales Representatives	Construction Workers	Transport Equipment Operators	Bookkeepers
No. Observations	702	403	87	86	70	62
Percent Micro	34.1	26.3	57.5	44.2	31.4	16.1
Percent Small	28.6	25.3	25.3	30.2	24.3	27.4
Percent Medium	24.9	33.0	12.6	23.3	31.4	41.9
Percent Large	12.4	15.4	4.6	2.3	12.9	14.5
Small	.163** (.062)	.418** (.097)	.116 (.168)	.284* (.158)	.032 (.220)	-.160 (.273)
Medium	.285** (.066)	.587** (.095)	.754** (.249)	.559** (.172)	.008 (.212)	.410 (.258)
Large	.419** (.085)	.610** (.121)	1.067** (.344)	.363 (.412)	.175 (.276)	.684** (.347)
R-Squared	.216	.438	.504	.278	.397	.288
Adj. R-Squared	.200	.418	.439	.171	.229	.075

* Significant at 10 percent level.

** Significant at 5 percent level.

[1]Sample exclude workers in mining and electricity. Regressions also control for years of schooling, experience and its square, household headship, industry and residence in Lima. Schooling coefficients differ significantly across primary, secondary and past-secondary levels only for transport equipment operators; the coefficients are constrained to be equal in all other regressions.

Table 12

Log Total Wage Regression Coefficients[1]
(Standard Errors in Parentheses)

Sample	Brother Pairs	Full	Full
No. Observations	59	1105	1105
Method	**Differenced OLS[2]**	**Multinomial Logit Selection Correction[3]**	**Ordered Probit Selection Correction[4]**
Small	.390**	.191	.054
	(.161)	(.318)	(.182)
Medium	.177	.358	.064
	(.173)	(.355)	(.310)
Large	.385	.309	-.062
	(.274)	(.316)	(.482)
Lamda0		-.251**	
		(.139)	
Lamda1		-.172	
		(.246)	
Lamda2		-.194	
		(.227)	
Lamda3		-.084	
		(.157)	
Lamda			.178
			(.162)
R-Squared	.137	.385	.382
Adj. R-Squared	.038	.373	.372

* Significant at 10 percent level.

** Significant at 5 percent level.

[1]Sample exclude workers on mining and electricity.

[2]Regression also controls for years of schooling, experience and its square. Industry indicators were not jointly significant and were dropped.

[3]Regression also controls for years of primary, secondary and post-secondary schooling, experience and its square, household headship, residence in Lima and industry. The first-stage multinomial logit model controls for years of secondary and post-secondary schooling, residence in Lima, participation in the transport sector, household nonlabor income per capita and number of children. In a more general model years of primary schooling, household headship, other industry indicators and other household structure variables were found jointly insignificant. Standard errors are corrected for the inclusion of estimated regressors following Murphy and Topel (1986). Lamdao-Lamda3 capture up to scale the expected value of the log wage regression error conditional on knowing the individual to be employed in a micro, small, medium or large enterprise.

[4]Regression also controls for years of primary, secondary and post-secondary schooling, experience and its square, household headship, residence in Lima and industry. The first-stage ordered probit controls for years of secondary and post-secondary schooling, experience and its square, residence in Lima, two sets of industries, numbers of adult males and females in the household and the number of children. In a more general model, years of primary schooling, household headship, household nonlabor income per capita and other industry controls were found jointly insignificant. Standard errors are corrected for the inclusion of an estimated regressor following Murphy and Topel (1986). Lamda captures up to scale the expected value of the log wage regression error conditional on knowing the size of enterprise in which the individual is employed.

Table 13

Various Statistics by Establishment Size[1]

	All	Micro	Small	Medium	Large
a. Of those reporting union and IPSS status, 1101 observations:					
Percent Union	23.6	3.2	8.3	33.6	70.3
Percent Above Minimum[2]	71.6	51.0	72.6	85.3	88.5
Percent IPSS	50.1	18.7	43.6	75.2	84.5
b. Of those with IPSS = 1, 552 observations:					
Percent Above Minimum[2]	85.9	70.3	80.3	90.0	92.0
c. Of those with IPSS = 0, 549 observations:					
Percent Above Minimum[2]	57.2	46.6	66.7	71.1	69.6

[1]Samples exclude workers in mining and electricity.

[2]A worker is counted as receiving a wage in compliance with the legal minimum if the nominal hourly cash wage is greater than or equal to the nominal monthly minimum wage divided by (48* the average number of weeks in a month), where 48 is the legal work week.

Table 14

Log Total Wage Regression Coefficients[1]
(Standard Errors in Parentheses)

Method	OLS	OLS	OLS	OLS	25th Quantile	75th Quantile
No. Observations	1101	1101	1101	1101	1101	1101
Blue Collar						
Small	.160** (.064)	.145** (.065)	.194** (.075)	.201** (.074)	.104 (.102)	.249* (.089)
Medium	.277** (.068)	.230** (.078)	.128 (.111)	.145 (.108)	-.018 (.144)	.200 (.131)
Large	.395** (.086)	.343** (.101)	.262 (.176)	.145 (.155)	.113 (.210)	.043 (.178)
Union	-	.024 (.074)	-	-	-	-
IPSS	-	.070 (.065)	-	-	-	-
Union* Micro	-	-	.001 (.239)	-	-	-
Union* Small	-	-	.147 (.165)	-	-	-
Union* Medium	-	-	.077 (.108)	-	-	-
Union* Large	-	-	-.211 (.153)	-	-	-
IPSS* Micro	-	-	.052 (.137)	.052 (.136)	.052 (.179)	.054 (.163)
IPSS* Small	-	-	-.103 (.100)	-.084 (.098)	-.016 (.134)	-.171 (.117)
IPSS* Medium	-	-	.176 (.119)	.200 (.116)	.285 (.155)	.041 (.140)
IPSS* Large	-	-	.375** (.172)	.345** (.166)	.426* (.232)	.358* (.197)

White Collar						
Small	.400**	.408**	.414**	.402**	.436**	.321**
	(.092)	(.094)	(.128)	(.092)	(.128)	(.109)
Medium	.581**	.578**	.633**	.582**	.495**	.697**
	(.089)	(.096)	(.145)	(.089)	(.122)	(.104)
Large	.629**	.585**	.648	.631**	.607**	.550**
	(.114)	(.136)	(.413)	(.114)	(.158)	(.133)
Union	-	.103	-	-	-	-
		(.098)				
IPSS	-	-.050	-	-	-	-
		(.080)				
Union* Micro	-	-	.569	-	-	-
			(.385)			
Union* Small	-	-	.092	-	-	-
			(.269)			
Union* Medium	-	-	.021	-	-	-
			(.134)			
Union* Large	-	-	.175	-	-	-
			(.196)			
IPSS* Micro	-	-	-.033	-	-	-
			(.138)			
IPSS* Small	-	-	-.023	-	-	-
			(.138)			
IPSS* Medium	-	-	-.069	-	-	-
			(.139)			
IPSS* Large	-	-	-.154	-	-	-
			(.395)			
R-Squared	.411	.412	.419	.415		
Adj. R-Squared	.398	.397	.397	.400		
Pseudo R-Squared					.232	.267

* Significant at 10 percent level.
** Significant at 5 percent level.
[1]Sample excludes workers in mining and electricity, and individuals for whom union or IPSS indicat are missing. Regressions also control for years of primary, secondary and past-secondary schooling, household headship and residence in Lima, the coefficients of which are allowed to differ across blue and white collar subsamples, as well as experience and its square, and industry, the coefficients of which are not allowed to differ. The null hypothesis of the equality of the intercept and the indicated slopes across blue and white collar subsamples could not be rejected at any reasonable significance level.

Table 15

Log Total Wage Regression Coefficients[1]
(Standard Errors in Parentheses)

Method	OLS	OLS	OLS		IV[2]	
No. Observations	1104	1104	1104		1104	
			Blue	White	Blue	White
Small	.248**	.268**	.174**	.427**	.109	.367**
	(.054)	(.067)	(.076)	(.099)	(.156)	(.155)
Medium	.395**	.354**	.207**	.558**	.104	.475**
	(.055)	(.068)	(.080)	(.095)	(.147)	(.139)
Large	.449**	.399**	.331**	.620**	.340**	.631**
	(.071)	(.092)	(.103)	(.128)	(.170)	(.191)
Tenure	.006*					
	(.003)					
Tenure* Micro		.003	.003		.004	
		(.005)	(.005)		(.019)	
Tenure* Small		-.001	.000		.015	
		(.006)	(.006)		(.022)	
Tenure* Medium		.010**	.011**		.026*	
		(.005)	(.005)		(.015)	
Tenure* Large		.010	.008		.008	
		(.007)	(.006)		(.015)	
R-Squared	.383	.385	.412		.400	
Adj. R-Squared	.373	.373	.397		.385	

* Significant at 10 percent level.

** Significant at 5 percent level.

[1]Samples exclude workers in mining and electricity, and an individual for whom tenure is missing. The first two regressions also control for years of primary, secondary and post-secondary schooling experience and its square, household headship, industry and residence in Lima. The latter two regressions control for these, some effects allowed to differ across blue and white collar subsamples, as in Table 14.

[2]Instruments for experience, experience^2 and the four tenure variables include experience at time of hire and its square, number of adults in household at time of hire and number of the individual's children at time of hire, each fully interacted with size categories.

Table 16

Log Total Wage Regression Coefficients[1]
(Standard Errors in Parentheses)

	Small	Medium	Large	No. Obs.	R-Squared
Manufacturing	.186* (.100)	.333* (.097)	.407** (.111)	354	.379
Construction	.347** (.161)	.648** (.163)	.618* (.327)	107	.290
Commerce	.147 (.016)	.279** (.120)	.364** (.167)	253	.439
Transport	-.189 (.200)	.037 (.203)	.276 (.225)	98	.295
Finance	.610** (.238)	.843** (.238)	.818** (.237)	85	.531
Service	.371** (.148)	.385** (.176)	.343 (.269)	146	.326
Agriculture	.862** (.243)	1.188** (.268)	1.223** (.243)	62	.423

* Significant at 10 percent level.

** Significant at 5 percent level.

[1]Regressions also control for years of schooling, experience and its square, household headship and residence in Lima.