Institute of Social Studies

Graduate School of Development Studies

OVEREDUCATION IN PERU: DETERMINANTS AND EFFECTS ON LABOR INCOMES: 1995-1999

A Research Paper presented by:

CARLOS R. CALIENES VILLANUEVA
(PERU)

In Partial Fulfillment of the Requirements for Obtaining the Degree of:

Master of Arts in Development Studies
Specialization:

ECONOMICS OF DEVELOPMENT

Members of the Examining Committee:
Dr. Jose Cuesta
Dr. Arjun Bedi

The Hague, December 2003

Enquiries:

Postal Address:
Institute of Social Studies
P.O. Box 29776
2502 LT The Hague
The Netherlands

Telephone: +31-70-4260 460
Telefax: +31-70-4260 799
email: zubli@iss.nl

Location:
Kortenaerkade 12
2518 AX The Hague
The Netherlands

Acknowledgements

A Chinese saying: "when you drink water, think of its source." While feeling the happiness for the birth of the paper, I can't forget the people who have nurtured its gestation.

Firstly, my deepest gratitude goes to Dr. Bamidele Olowu who is my first supervisor. I have continuously and greatly benefited from his ceaseless encouragement and highly valuable guidance. I still remember his word: "I believe that you can do it well". From the very beginning of the research paper, I have benefited from the illuminatin discussions with him and the relevant references he recommended and rendered. Moreover, I am especially grateful that he sacrificed a lot of valuable time to help me polish my final version for making the paper more readable.

I also would like to express my heartfelt thanks to Dr. George Tsogas, who is my second supervisor, for his substantive advice and comments on my paper that helped me sharpen my research focus.

At ISS, I learned a lot from my colleagues at class as well as outside the class. I would like to extend my special thanks to HRE staff group, namely; Dr. Irene Van Staveren, Dr. Freek Schiphorst, Dr. Abbas Abdelkarim and of course Tanya Kingdon, the program administrator.

We were just like a happy family here and we learned from each other, also, shared rich experience each other. I would like to take this opportunity to appreciate among my colleagues--my study group members: Katherine, Bello, Cheka and Lysda who took their time to read my research paper design and offer very invaluable comments on my preliminary research ideas. Besides, during the past one year, I really learned a lot from the efficient study group, which helped me finish the program successfully. I will cherish this deep friendship, which had built up within us. Dear friends, thank you very much for

Index

1. Introduction

The term "overeducation" is commonly used to characterize a situation in which individuals have schooling level in excess of the educational level required by their current job.

In the last decades, education started to be broadly recognized as an instrument for economic growth, social mobility and a mechanism to increase the possibilities to be employed in higher quality jobs. Investment in education and the educational level of the population increased significantly in many countries, included Peru.

However, since the 70's the rates of unemployment in skilled workers increased and it became more frequent to find high educated workers performing jobs with low educational requirements.

In consequence, some theorists started to question the economic value of education: is really the labor market able to absorb the huge increase of educated workers?, will the workers have the opportunity to use all the knowledge and skills they have acquired?, is it really economically profitable the investment in education?.

The present research aims to respond the previous questions for the case of a developing country (Peru) in a context of economic liberalization (1995-99) with two sub periods: high economic growth (95-97) and economic recession (98-99).

Works about overeducation has been focused mainly on developed countries where the presence of skilled workers in low skill jobs seems to be a common feature of the labor markets. However, the incidence of overeducation is also high in some developing countries. In Peru, the first symptoms of overeducation appeared after the 70's when the rates of unemployment and underemployment of skilled workers started to increase. Before that, and since the 60's, investment in education (public and private) had started to increased sharply.

The importance of overeducation as an economical problem is valued in different ways depending on the theoretical framework considered. For Human Capital Theory overeducation is not a real problem since, according to this theory, in competitive environments, all the human capital of the individual is fully utilized by the firms and there is no room for underutilization of resources. In the other extreme, Job Competition Theory does not give any value to personal characteristics of the workers and claims that productivity is actually linked to job characteristics. In this case, the excess of education in relation with the job performed does not have any marginal return.
In the middle of these extremes there are other theories (Job Matching and Assignment Theories), that takes into account aspects of the formers and get intermediate conclusions.

This research has two main objectives. The first one is to analyze the determinants of overeducation in the Peruvian labor market. The second one is to evaluate the different theories that explain the economic value of overeducation and find which one is more relevant for the Peruvian case. The data utilized is the one of the National Household Survey of Peru (Encuesta Nacional de Hogares-ENAHO) for the years 1995-1999 and for the case of the capital city, Lima Metropolitana that group almost one third of the Peruvian labor market.

The research is organized in the following way. Chapter 2 defines overeducation and presents the different methodologies that measure it. Here it is also presented what has been done so far in Peru about the measurement of overeducation. Chapter 3 presents some theoretical frameworks that treat overeducation, its determinants and its effect over labor earnings and productivity. Chapter 4 gives an overview to the Peruvian economy, its labor market and the Peruvian tertiary educational system during the last decade. Chapter 5 is the empirical one and two models are carried on. The first one evaluates the determinants of overeducation through a *probit* model. The second one evaluates the theoretical approaches presented before and its consequence over labor income. In this case earning models with the so called ORU specification are carried on.

2. Overeducation: Definition, measurement and advances in Peru

2.1 Definition

Overeducation refers to the situation in which individuals have schooling level in excess of the educational level required by their job (Moullet, 1999). In other words, a worker is considered overeducated if he/she has more education than the required to perform his/her job, irrespective of the salary paid. According to Green (1999) the concept of overeducation is more complex since it is not easy to estimate the education required to do a particular job. Furthermore, a deeper definition of overeducation and its consequences highly depends on the theoretical approach taken to understand the functioning of the labor market and particularly the mechanism to allocate workers to jobs.

2.2 Measurement of Overeducation

The estimation of the education required to perform a particular job is a problematic topic that has been assessed in different ways by social scientists[1]. The measurement of overeducation is closely related with the former and shows the same controversy in its treatment.

There are at least three approaches that are commonly used to measure overeducation, each of them, however, shows limitations. They are the Self Assessment, the Job Analysis and the Statistical approaches.

2.2.1 Self Assessment Approach

In this method, survey respondents (workers or employers) are asked to make their own assessment of the minimum educational level or qualifications required to perform their job adequately. This method was used by Duncan and Hoffman (1981), Sicherman (1991) and Alba Ramirez (1993) and its main advantage is that the information collected is not only up-to-date but also highly accurate in the sense that doesn't deal with aggregations but with the specific job carried on by the worker. However information of this kind is not usually available and requires an *ad-hoc* collection process.

[1] Halaby (1994) makes a complete overview of the different perspectives that social sciences have about this topic.

Problems of the method

1. There could be a classification error since it is difficult to know how the worker made his/her judgment. As Green (1999) mentions, the answer could be more focused in the level required to *enter* to the job and not to *perform* the job, or the workers can report their own skill level or use any other criterion.

2. The method left room to subjective answers: one person's assessment may not match with the evaluation of other person doing a similar job.

3. The method could lead to biased results because respondents who are genuinely overeducated may have a bad attitude about their jobs and be less likely to respond questions on this issue.

2.2.2 Job Analysis Approach

This measure uses job analysis data and the information available in the occupational code or dictionaries of occupational titles. These sources contain information about the educational requirements of a wide range of occupations and have been made by specialists that follow an objective procedure. The method was used by Rumberger (1987), Thurow (1975), and Hartog (1980). Its main advantage is that it analyses jobs starting from the technology and the kind of activities performed following an explicit objective.

Problems of the method

1. Usually under the same occupational title there are different jobs so, as Green (1999) mention, it is highly possible that workers sharing the same occupational title do not undertake exactly the same work. The problem is even higher if we consider that the assumption that all jobs with the same title have the same educational requirement does not necessarily hold.

2. The method assumes that all years of schooling has the same value i.e. the quality of education is irrelevant and all workers with the same years of education are substitutes. It is not necessarily true, especially in educational systems with high quality differences. That could be the case of private and public centers or progressive quality deterioration along the years.

3. Jobs change over time and occupational data are often out of date. The problem is deeper in contexts of rapid technological transformation that affect the work environment.

2.2.3 Statistical Approach

This approach estimate overeducation considering the mean years of education and the standard deviation for a range of occupations and define an overeducated worker as the one that has more than one standard deviation above the mean education for his/her occupation (Green, 1999). This method was used by Verdugo (1989), Clogg (1984) and De Gripp (1998). The main advantage of this approach is its easy measurement that is especially useful when additional data is not available.

Problems of the method

1. The choice of one standard deviation as the cut off point is arbitrary.
2. If it is considered that the level of education for a particular job can follow a normal distribution, the method tends to generate symmetric estimates of the incidence of over and under education because only measures the tails of the distribution (Green, 1999). In a normal distribution this proportion would be near 15%.
3. Since the method is based on the observed distribution of education for a given occupation, the result is sensitive to cohort effects. This sensitivity is higher under contexts of rapid change in the educational level required for a given occupation. (Chevalier, 2000).
4. Like in the Job Analysis Approach, this method assumes homogeneity in quality education and that all occupational titles have identical skill requirements.

Finally it is worth to mention that Hartog (1997) finds that the Self Assessment Approach and the Job Analysis Approach yields highly close results measuring overeducation, however most of the authors consider that the Job Analysis is superior since its criteria is more objective and trained people are in charge of grading the jobs. On the other hand, authors agree that the Statistical Approach is the weakest one.

For the Peruvian case, since *ad-hoc* data about educational requirements is not available, the Job Analysis methodology will be used.

2.3. Estimation of overeducation in Peru: what is known so far?

Most of the studies about overeducation in Peru have been focused in the estimation of its incidence in the labor market and do not explore the consequences over incomes or productivity. Table 2.1 summarizes the results obtained by some of the works[2]. The results obtained are quite different each other showing probably the high dependency on the methodology used.

Table 2.1

Incidence of Overeducation

	Year of analysis	Methodology	Scope	Tertiary university	Tertiary no university
Rodriguez	1992	Job Analysis	Lima Metropolitana	37.2%	73.3%
Arregui	1991	Job Analysis	Lima Metropolitana	49%	42%
Verdera	1993	Job Analysis	Lima Metropolitana	34.2%	47.3%
FORTE – PE	1998	Job Analysis	Urban National	-	34.7%
Haya de la Torre	1999	Job Analysis	Lima Metropolitana	-	12%
Burga	1997	Statistical	Urban National	38.6% – 49,7%	-

Burga and Moreno (2001) also analyzes the determinants of overeducation for workers with tertiary university education using a *logit* model. The authors find that the probabilities of being overeducated increases if the worker comes from a poor household. Besides, among the individual characteristics, married workers are more likely to be overeducated as well as less experienced ones. On the other hand, workers that receive

[2] Rodriguez (1993), Arregui (1993), Verdera (1995), FORTE PE (1998), Haya de la Torre (1998), Burga (2001).

more specific training (vocational courses) have less probability to be overeducated and gender condition is not found significant. Finally, the authors try to approximate quality of education distinguishing private and public education but do not find any significant relationship between the variables with the incidence of overeducation. They explain this result due to the broad dispersion in the quality of the private educational services.

More recently, the Peruvian Ministry of Labor[3] have estimated the compatibility between the level of education acquired by the worker and the one required by the job performed. The methodology used was the Job Analysis one and was based on the Peruvian National Code of Occupations. Due to the reliability of this methodology and its official status, it is the one used in this study.

The procedure is to identify the different levels of education acquired by workers and compare them with the occupations that are classified in an ad-hoc code[4]. It was considered four formal educational levels:

1. Incomplete secondary and below
2. Complete secondary and incomplete tertiary non university
3. Complete tertiary non university and incomplete tertiary university
4. Complete tertiary university

The classification of occupations takes into account 5 criteria:

1. The tasks carried on
2. The level of responsibility
3. The degree of authority and the levels of subordination
4. The status of the worker in the firm
5. The degree of training required

Finally, the occupations are classified in five groups:

[3] Ministerio de Trabajo y Promoción del Empleo (MTPS): Boletín de Economía Laboral # 17 - "Estimación de la compatibilidad entre la educación adquirida y la ocupación ejercida por los trabajadores" (2000).

1. Low skilled occupations
2. Intermediate skilled occupations
3. Technical occupations
4. Professional occupations
5. Occupations of management

There are three possible exclusive results in the matching between the educational level and the occupational groups identified: overeducated, adequately educated and undereducated. The worker is, overeducated if carry on a job with a lower educational requirement, adequately educated if education acquired and required are the same and undereducated if performs a job that requires a higher level of education. The possible results of the degree of correspondence are presented in Table 2.2.

Table 2.2
Degree of equivalence between education level and occupation

Educational Level			
	Undereducated	**Adequately educated**	**Overeducated**
1. Until incomplete secondary	3, 4 and 5 (*)	1 and 2	
2. Complete secondary and incomplete tertiary non university	3, 4 and 5	2	1
3. Complete tertiary non university and incomplete tertiary university	4 and 5	3	1 and 2
4. Complete tertiary university		4 and 5	1, 2 and 3

(*) the numbers correspond to the occupational group

Taking into account the relations showed in Table 2.2, a variable of educational correspondence is constructed. Besides, from the information obtained it is also possible to calculate years of required education, overeducation or undereducation. These variables are used in the empirical section.

[4] It is the National Code of Occupations developed by the MTPS and the Peruvian Institute of Statistics (INEI)

3. Theoretical Approaches to Overeducation

The economical relevance of overeducation in the labor market has been evaluated under different perspectives depending on the theoretical approach considered.
This research works with theories that study two main fields in overeducation literature: the effects of overeducation on labor incomes and the determinants of overeducation.

About the effects of overeducation over labor incomes, there are five theoretical approaches that predict different returns to years of overschooling: Human Capital Theory, Job Signaling Theory, Job Matching Theory, Job Competition Theory and the Motivational Approach. Some of these theories will be contrasted in Chapter 5 in order to find which one fits better with the Peruvian case.

About the determinants of overeducation, the Heterogeneous Labor Supply, the Occupational Mobility and the Regional/Global market approaches emphasize different aspects of the workers and the labor market on the probability of a worker to be overeducated. These predictions are evaluated also in Chapter 5.

It is worth to mention that in the present research it is considered only the pecuniary rewards of education. Education has also non pecuniary returns associated with the enjoyment that students may get from the very fact of being studying. However, there is evidence (Tsang, Rumberger and Levin, 1991) that overeducated workers have less levels of job satisfaction even after controlling by pecuniary and non-pecuniary returns so; overeducation wouldn't be a trivial problem even when only non pecuniary effects are considered.

3.1 Effect of overeducation over labor earnings and productivity: Theoretical approaches

3.1.1 Human Capital Theory

Human Capital Theory (HCT-Becker, 1964) considers that labor markets are full efficient and that every worker is paid according to the value of his/her marginal product. Here,

the concept of overeducation (and its counterpart, undereducation) may be meaningless. Profit maximizing firms fully utilizes their entire worker's human capital and, in consequence, a particular job doesn't have a specific educational requirement attached to it. It also implies that variables measuring job characteristics should not appear in any model of earnings.

For HCT, over investment in education is a short term phenomenon that has its origin in the uncoordinated functioning of demand and supply (Freeman, 1976) and that eventually leads to a change in the production methods of the firms. It implies a higher use of skilled labor and/or a fall in the rate of return to education. According to this point of view, if the supply of skilled workers exceeds its demand, individuals will not be truly underutilized but they just will receive a lower rate return to their education due to an excess supply of that type of labor (Green, 1999).

In this theory, education increases human capital and leads to higher productivity in the labor market. Employers value labor productivity by offering higher wages to workers with higher human capital. In this scenario, and in a competitive environment (where the information available to economic agents is perfect), the entire productive characteristics of the worker are paid in the same way across the labor market. If the firm doesn't pay for the whole set of characteristics, there will be another firm that will do it and the worker will try to find another job. Here, all human capital invested is paid for: education and overeducation are paid with the same return.

Under this approach, the return to education is a consequence of the interaction between demand and supply. When the skilled labor supply is scarce, the return to education increases and more people will desire to invest in education. The education required to carry on satisfactorily a particular job depends on the relative supply and demand of workers. It means that differences in wages and productivity depends only on human capital and are independent from job characteristics. In this sense, two workers with the same level of human capital should have the same labor earnings no matters the job they carry on. The rate of return to each year of overeducation would be the same as the rate of return to each year of *adequate* education i.e. the one required by the job.

To summarize, the crucial point for HCT is that the marginal return on schooling should be no affected by the amount of overeducation.

3.1.2 Job Screening Theory

HCT assumes that agents have perfect information about the behavior and characteristics of the other agents and the functioning of the market. It means that firms know the productivity of each worker and pay according to it. However, in general, the assumption of perfect information doesn't hold. Usually, firms do not have enough information about the productivity of their workers and therefore can not establish their productivity accurately.

In this context, the Job Screening Theory (JST – Spence, 1973) proposes that education acts as an indicator enabling employers to identify the most productive workers. This role of education is also recognized by the workers and encourages them to invest in education.

Under this approach, education performs a signaling role and part of this function is to help to distinguish between workers with different innate ability. It is possible when it is taken into account that the cost of education is composed by learning and financial costs. It is reasonable to think that higher ability workers have lower learning costs and that-assuming equal financial cost for all individuals-, investment in education will be profitable only for higher ability workers.

However if learning costs of education fall down (for instance due to the decline of academic requirements to get a professional title), education could become profitable also for lower ability workers and the average level of education of new entrants to the market will increase. The difference is that education no longer functions as a signal and, in consequence, employers have to upgrade their *nominal* educational requirements. This phenomenon is called *qualification inflation* because no real change in the educational

level is involved. There is just an increase in the educational requirements of jobs but without any change in job content[5].

An extreme case of this model is to consider education only as a signaling device without any implication on productivity. In this case, *qualification inflation* will increase the resources allocated to education but without any increase in productivity levels. Green (1999) considers that in this case all the workers would be considered overeducated in the sense that education merely identifies innate ability rather than actually providing skills to perform a job.

In this model, education is treated as a positional good and more education implies higher wages, not because an increase in the productivity of the worker but because a credentialist effect. The return to overeducation can be in this case even higher than the marginal productivity of the worker.

3.1.3 Job Competition Theory

According to Job Competition Theory (JCT- Thurow, 1973), wages and productivity are determined primarily by job characteristics and not by the characteristics of the individual.

As we saw above, HC theory suggests that labor market is regulated by wage competition: a negative shock of labor demand lead to a bumping down process that also reduces the return to education. In this context, if economic agents (students in this case) anticipate this effect, educational investment will diminish and, at the end, the labor supply will be adjusted to the demand shock. Empirically, the shift of higher educated workers to lower level jobs has been observed in many countries; however it has not been observed reductions in education investment. It can suggest that labor market is not regulated by wage competition but by job competition.

Under Job Competition, workers are ranked according to their ability in a labor queue where the first places are given to the most able or educated candidates and in the tail-end are placed the less schooled ones. Firms assume that the formers require less amount of training in order to perform well a job. On the other hand, firms rank job positions

[5] In this case, assessment of overeducation based on employers' perceptions would underestimate the real incidence of overeducation.

following a similar criterion: the best positions (the ones that are more complex and have higher wages) are put in first place. In this framework, the most productive jobs go to the workers with lowest training cost (those who have higher education), and the least productive positions are offered to the applicants with less schooling. The relationship between education and wages is still positive not because of the effect over productivity but because of the positional effect that education has. In this sense, in general, two persons with the same amount of human capital (education, experience, etc.) will have different wages. They would have the same wage only in the particular case of performing the same kind of job and every type of job is characterized by a constant wage level that is explained by wage mark-ups due to monopolistic competition (Snower, 1983).

Van der Meer (2000) suggests that in this model the presence of overeducation in the labor market is highly probable. In periods when labor supply is abundant or demand decreases, high skilled workers, which are in the top of the queue, will be located in jobs that need less schooling than the actually acquired by the individual: everyone will shift some places downward in the queue. According to this theory, this excess of schooling will not be rewarded. Since productivity is linked only with job characteristics, overeducation will not increase it. The marginal rate of return of each year of overeducation will be zero.

3.1.4 Job Matching / Assignment Theory

The Job Matching Theory (JMT – Sattinger (1980, 1993), Hartog (1981, 2000)) proposes that productivity and earnings depends on the human capital of the worker, the characteristics of the job and the relationship between both of them. In other words, it is the degree of fit or kind of combination between individual and job characteristic that determines the labor productivity: If a worker is employed in a job that doesn't fit with his/her characteristics, his/her acquired skills are underutilized and his/her productivity - and wages- are less than the optimal. The allocation of workers will be optimal when every worker is matched to a job in which he performs better than anyone else.

The model assumes the existence of heterogeneous workers and heterogeneous jobs. In this context, wages not only reward productivity but also facilitate the allocation process between workers and jobs.

Wolbers (2002) mentions that the incidence of overeducation depends on the difference between the proportion of available jobs and the existence of workers with the same education. The matching model includes both, the human capital and job competition theory and postulates that educational mismatch can be a permanent feature of the labor market.

In an economy without an assignment problem where there is only one observable skill, a worker will get the same wage no matter which job he/she has. As Sattinger (1993) mention, "hiring would be reduced to the trivial problem of taking the first worker that came along. Unemployment would only arise if the number of workers exceeded the number of jobs". However, clearly, it is not a realistic situation. Economies face an assignment problem and spend significant amounts of resources in the process of matching workers with jobs in the best possible way.

Sattinger (1993) systematizes a number of models to explain how the profit or utility maximizing decisions of workers or firms generate a particular assignment. Two of the models presented by Sattinger that illustrate formally the assignment problem are the Labor Market Comparative Advantage Model and the Scale of Operation Effect Model.

3.1.4.1 Labor Market Comparative Advantage Model (LMCA)

LMCA takes as assumptions a fixed proportions technology and employers that need to have a set of tasks performed to get a determined level of output. Besides, workers do not have preferences for some jobs. If a_{ij} is the number of times that worker i can perform job j's task per period and if:

$$\frac{a_{11}}{a_{21}} > \frac{a_{12}}{a_{22}} \qquad (1)$$

It can be said that worker 1 has a comparative advantage at job 1 and worker 2 has it in job 2.

The assignment in the labor market under LMCA is given in the following way. The equilibrium wage for individual i is w_i and employers try to get the job performed minimizing the costs but taking the wage rate as given. The cost of employing worker i is

$\frac{w_i}{a_{ij}}$ and worker 1 will be preferred by employer j whenever:

$$\frac{w_2}{a_{2j}} > \frac{w_1}{a_{1j}} \qquad (2)$$

However, according to (1), it is not possible for employer 1 to hire worker 2 and for employer 2 to hire worker 1. It is so because this assignment requires the condition $\frac{a_{11}}{a_{21}} < \frac{w_1}{w_2} < \frac{a_{12}}{a_{22}}$ that, given (1), doesn't hold. The possible assignments, depending on the wages rates, are: (1) both employers prefer worker 1, (2) both employers prefer worker 2 and (3) employer one prefer worker 1 and employer 2 prefer worker 2. Given the technology assumed only assignment (3) is possible.

3.1.4.2 Scale of Operations Effect Model

Comparative advantage is not the only production principle underlying the assignment of workers to job. With cooperating factors of production, it is necessary to consider the opportunity cost of the factor (machinery, for instance) and subtract it from the value of output to yield the earnings. Under this technology, the optimal arrangement indicates that the workers with greater abilities should have more resources (capital, labor or responsibilities) because in this way the effect over the output is greater. In turn, since output is more sensitive to the abilities of workers, wage differentials increases favoring the most skilled ones. In addition this approach can explain why some workers are unemployable (Akerlof, 1981). Hierarchical models also arrive to similar conclusions. Calvo and Wellisz (1979) develop a model where the low performance of supervisors leads to low performance of workers under his/her supervision so; firms will place more able workers at the higher levels in the hierarchy. Under this technology and "with the scale operation effect, the wage ratio for the 2 workers will not lie between the ratios of outputs as in the LMCA case because of the presence of opportunity costs from the use of

a machine" (Sattinger, 1993). In this context of cooperating factors and with only one worker per machine per unit of time, wages are determined in the following way:

p_j = price of a unit of output produced per period by worker i at machine j

a_{ij} = output produced by worker i at machine j

w_i = wage for worker i

The owner of machine j takes the wage as given and chooses the worker that maximizes the residual $p_i a_{ij} - w_i$. If the owner of machine 1 chooses worker 1 and the owner of machine 2 chooses worker 2:

$= p_1 a_{11} - w_1 \geq p_1 a_{21} - w_2$

$= w_1 - w_2 \leq p_1 (a_{11} - a_{21})$

And,

$= p_2 a_{12} - w_1 \leq p_2 a_{22} - w_2$

$= w_1 - w_2 \geq p_2 (a_{12} - a_{22})$

So,

$p_2 (a_{12} - a_{22}) \leq w_1 - w_2 \leq p_1 (a_{11} - a_{21})$

As can be observed, the difference in wages must lie between the difference in the value of output produced by the two workers on machine 1 and the corresponding difference on machine 2. The assignment of worker 1 to machine 1 and worker 2 to machine 2 is possible only if $p_2 (a_{12} - a_{22}) \leq p_1 (a_{11} - a_{21})$.

The two models show how maximizing behavior in employers generate assignments of workers to jobs. Besides these simple models, assignment theory works more sophisticated ones that consider worker preferences (Tinbergen, 1951 and its further developments) and examine how the allocation of workers to jobs generates wage differentials and determine the distribution of income among workers (Sattinger, 1979). In summary, the two models presented show the relevance of the assignment process in the determination of labor earnings: the marginal return to education for a correctly allocated worker will differ from the one of the misallocated worker, the over and under educated one.

3.1.5 Motivational Approach

According to the Motivational Approach (MA- Rumberger (1981), Tsang (1985)), overeducation is a source of frustration in workers and affects negatively their productivity. Overeducated workers would have higher level of job dissatisfaction, more health problems and higher rates of absenteeism than the workers correctly allocated. In addition, overeducated use to change jobs more often and training investments of firms are less profitable since the recovery time is shorter[6].

These effects reduce the overall productivity of firms that hire overeducated workers so it can be predicted that firms will try to avoid hiring this kind of workers. Since the effect on productivity is negative, the effect on wages and the marginal rate of return to each year of overeducation will be also negative. This approach is totally opposite to the previous ones where overeducation has a positive impact (HCT), or at least no impact (JCT) on productivity.

Different authors have worked empirically this approach and have gotten different conclusions. Tsang's (1987) finds a negative effect of overeducation over productivity for the case of 22 US Bell companies. To get that result, Tsang links two separate causal effects: overeducated workers are less satisfied with their jobs and low satisfaction is linked with lower productivity.

Buchel (2000), on the other hand, also studies, the relationship overeducation-productivity through job satisfaction for the German case but find a positive relation between the variables: overeducated are more productive, enjoy better health, have higher rates of participation in training programs and show longer firm tenures. In summary, Buchel finds that overeducated workers are more productive than their corrected allocated colleagues. So, it can not be proved productivity losses from overeducation.

[6] Most of the works about this approach are developed for the US case: Kornhauser (1964), Vroom (1964), Kasl (1974), House (1974), Coburn (1975), Quinn (1975).

The difference in both approaches could be the different methodology followed. Tsang compare the satisfaction of an overeducated worker with the satisfaction of a correctly allocated worker that has his/her same level of education. Here the straightforward conclusion is that the overeducated worker is less satisfied. On the other hand, Buchel compare the satisfaction of the overeducated worker with the satisfaction of any worker doing the same job and find that it is not the case that the overeducated worker is less satisfied than others performing a similar job.

Finally, some considerations must be taken into account to measure the economic returns to overeducation. According to some authors, (Doeringer, 1971 and Sthincombe, 1979) the labor market is a segmented one where the allocation of labor is organized differently in each segment. Competitive segments of the market would function as HCT predicts and no competitive segments would work under different processes of allocation. Van der Meer (2000) evaluates this approach considering private and public sector for The Netherlands. According to this approach, private sector would function more as a competitive market and employers will have to pay their workers according to their marginal productivity that is related to worker's schooling (including overeducation). On the other hand, public organizations are not involved in a competitive environment and do not follow profit maximization goals. For public sector, Job Screening Model or Job Competition model would be more appropriate. The results found by Van der Meer did not confirm the predictions of this approach.
Besides, it can also be considered formal and informal sectors: Taking into account that the informal sector is not regulated it would function like a competitive one. In consequence it would be expected that this sector functions as HCT predicts.

3.2 Theoretical Approaches about determinants of overeducation

3.2.1 Heterogeneous Labor Supply – Grade Drift

Green (1999) and Chevalier (2000) links overeducation with the phenomenon of *Grade drift* and postulate the existence of education mismatch due to changes in the quality of the education.

In this case some workers receive less qualification than others in the formal education system. Due to these quality differentials, it is no longer possible to consider workers with the same level of education as homogeneous.

According to some authors[7], the quality of educational services has declined in the last decades and not to take into account this fact can lead to overestimate *undereducation*. What appears as undereducation can truly be a problem of grade drift: older workers that acquired high quality education require less *amount* of education than younger workers with low quality education. In this context, employers can upgrade the educational requirements of jobs and put as the standard measure the educational level of the less quality workers (for instance, m years). In consequence, higher quality workers with n years of education (where $m>n$) will appear as undereducated. Actually, older high quality workers have less years of education but (at least) the same level of "skills" compared with new less quality workers.

Chevalier (2000), on the other hand, concludes that not to consider the decline in quality education overestimates *overeducation* and not *undereducation* as Green points out. The difference between both explanations is the standard of educational level considered to measure any particular job. As mentioned before, for Green the standard is years of schooling of the *new* workers (or less quality ones) so, old workers appear as "undereducated" (have less years but with higher quality). Chevalier, on the other hand, considers that the standard measure is the schooling of old workers (or high quality ones), so, new workers appear to be overeducated.

Under this approach, overeducation incidence could be more an apparent than a real problem since both hide problems in quality education. A more accurate empirical procedure to measure real educational mismatch would have to focus on skills rather than on formal education.

In the case of the Peruvian educational system, the heterogeneity among workers is also linked with the kind of educational service received. Usually, public education has less quality than private education. So, workers that come from public schools are more likely

[7] Chevalier (2000), for the European case, and Burga (2001), for the Peruvian case, point out this pattern.

to appear as overeducated if the standard measure of required education for a particular job is given by the level of the private education. Alternatively, workers that come from private schools would be more likely to be undereducated if the standard measure of required education for a particular job is given by the level of the public education. For the case of analysis, the former is the relevant one.

3.2.2 Theory of Occupational Mobility

According to this theory, undereducation is a temporary phenomenon that is explained by the labor cycle of the workers: at the beginning of their labor life, workers could perform a job in which they are overeducated but that will give them some abilities that allow them to perform a job adjusted to their education. In this sense, the incidence of overeducation would be higher in younger and less experienced workers who also would have higher probabilities to shift to other jobs. This theory is complementary to the HCT because postulates that the presence of overeducation is consequence of the lack of certain kind of human capital (experience). The persistence of overeducation along the labor life of a worker is explained in the theory by problems of information. Alba Ramirez (1993) for he case of Spain find some empirical evidence that show that the average duration in an employment for an overeducated worker is 7.6 years less than the correspondent average of an worker correctly allocated (i.e. with a similar level of acquired and job required education).

3.2.3 Access to regional and global labor markets

Frank (1978) incorporated spatial determinants in the probability of being overeducated. According to this author, the access to restricted regional labor markets in opposition to global ones increases the probability of a worker to be overeducated. This phenomenon would be more frequent in dual earner households where women with family responsibilities use to search for a job in the regional market in order to be able to fulfill reproductive activities. On the other hand, the male member of the family would be able to look for a job in the global market reducing the probability of overeducation. In general, any resource that facilitates the access to the global labor market (like labor

information, social networks and adequate transport means) reduces the likelihood to be overeducated.

Tables 3.1 and 3.2 summarize the principal implications of each of the theories presented about the returns to overeducation and its determinants.

Table 3.1

Effect of overeducation over labor earnings and productivity

Theory	Returns to overeducation (excess of education in relation with the education required by the job)	Returns to overeducation in relation with returns to education required by the job	Transmission channel education → wages	Determinants of wages	Determinants of productivity
HCT	(+)	=	Productivity	Human Capital of workers	Human Capital of workers
JST	(+)	<, =, >	Credentials	Human capital as a positional good	Innate abilities of workers
JCT	Zero	<	Positional effect	Job characteristics	Job characteristics
JAT	(+)	<	Productivity	Human capital and job characteristics	Human capital and job characteristics
MA	(-)	<	Productivity	Human Capital	Human Capital of workers

The econometric model developed in chapter 4 will contrast four of the five theories presented above. Job Signaling Theory will not be evaluated since it requires a different specification model to test its predictions.

Table 3.2

Overeducation determinants

Theory	**Theoretical Highlight**	**Determinants of overeducation**
Heterogeneous Labor Supply / Grade drift	Heterogeneity of the labor supply and the quality of the educational services	Higher incidence of overeducation in workers that receive less quality education
Occupational Mobility	Workers follow a labor cycle where shortages of one kind of HC (experience) are compensated with other ones (more education).	Higher incidence of overeducation in young individuals and workers with less experience.
Access to regional and global labor markets	Regional labor markets offers a smaller range of job opportunities	Higher incidence of overeducation in workers with less access to global markets (women and married workers and workers with less developed social networks).

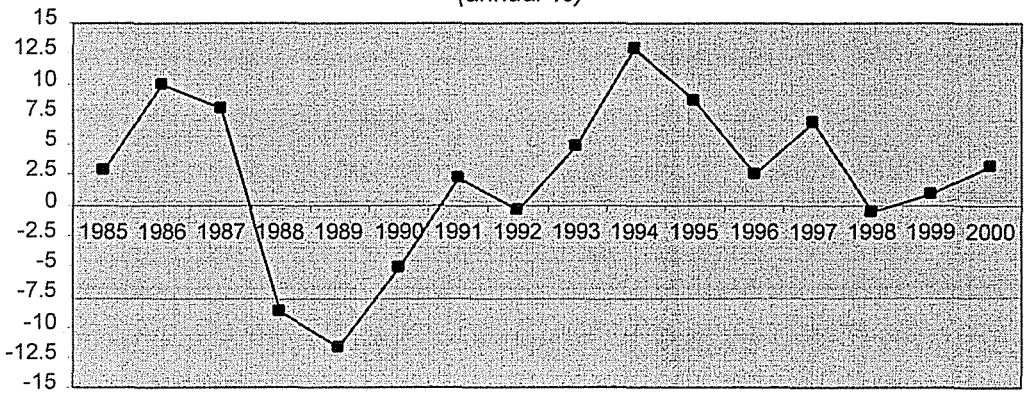
The econometric model developed in Chapter 5 will consider, among others variables, the relationships presented in Table 3.2. In the case of Grade drift and heterogeneity of the labor supply, it will be considered public and private education since it is expected that private schooling has higher quality. In the case of Occupational Mobility Theory, job tenure and age will be taken into account. Finally, to evaluate the predictions of the Regional / Global labor market approaches, it could be necessary to have spatial variables

that are not available in the data set so; these effects will be estimated indirectly taking as proxy gender, civil status conditions and access to social networks.

4. Overview of the Peruvian economy, the labor market and the tertiary educational system in the 90's

The decade of the 90's in Peru was characterized by a structural adjustment program at the beginning of the decade followed by structural reforms that increased the role of the market and reduced the size of the government in the economy. Before 1990; and since 1985, Peru had two clear different periods. During 1985-87 (Graph 4.1) the country had a brief period of high growth (10% in 1986) based on an active role of the public expenditure financed by a loose monetary policy. The consequences of the *populist* growth were felt since 1988 when the GDP started to decline and the inflation increased dramatically. For 1988, the first year of the crises, GDP fell down in 8.7% and inflation became unmanageable passing from 667% in 1988 to 3400% in 1989 and 7500% in 1990 (Graph 4.2).

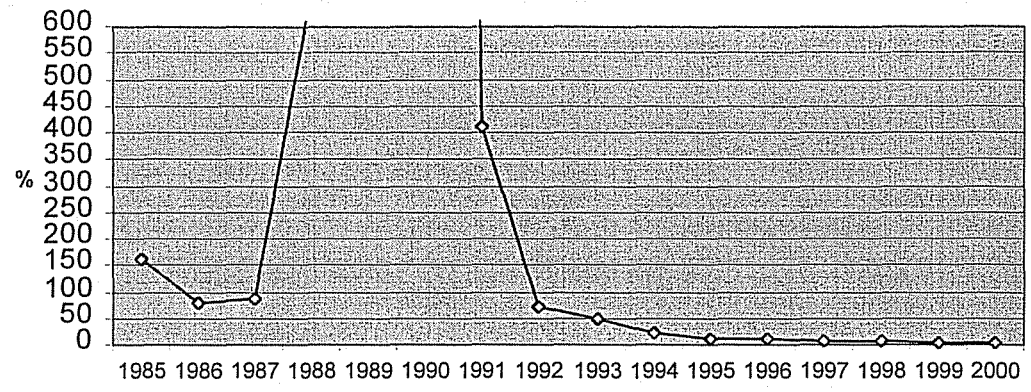
Graph 4.1
Growth of GDP
(annual %)



Source: Peruvian National Institute of Statistics - INEI

The structural reform program consisted on the redesign of the state size (reduction of the bureaucracy), elimination of public monopolies in the commerce of food, fuel, the privatization of public companies, and the liberalization of public prices, financial markets, the international trade, and the labor market.

Graph 4.2
Rate of Inflation
(Variations in the Consumer Price Index)



Source: Peruvian National Institute of Statistics - INEI

In the labor market the principal measures in the reform were the elimination of firing rigidities, the freezing of the minimal wages, and the prohibition of wage indexation.

The reforms of the 90's stabilized the economy reducing the inflation rate to international levels. As Graph 4.2 shows, inflation followed a decreasing pattern since 1991 until reach 3.7% in 1999. However, the adjustment had hard consequences on employment and production generating an initial recession. In 1990, the GDP declined in 5.1% and started to recover only after 1993. The period 1993-97 was the longest growth period in Peru in the last 30 years with an average annual growth of 8%. The period of growth finished in 1998 with the beginning of a brief recession period. In the last two years of the period analyzed the economy had a slight recovery.

Table 4.1 summarizes some macroeconomic indicators of Peru during the 90's

Table 4.1
Macroeconomic indicators
Peru 1990-2000

	90	91	92	93	94	95	96	97	98	99	2000
GDP*	-5.14	2.17	-0.43	4.76	12.82	8.58	2.49	6.75	-0.53	0.95	3.13
Trade Balance **	399	-189	-341	-607	-997	-2165	-1988	-1721	-2466	-616	
Net International Reserves**	682	1933	2425	2910	6025	6693	9862	7982	7134	7674	7747
External debt**	22856	25444	26612	27447	30191	33378	33805	28635	29477	27966	27460
Inflation	7481.66	409.5	73.53	48.58	23.74	11.13	11.54	8.562	7.248	3.47	3.757
Devaluation	10316.1	93.4	63	32.5	-1.4	9.4	10.7	5.1	15.4	11.2	1.0

* % variation

** Millions of US$

4.1 Some characteristics of the Peruvian labor market in the 90's

In general terms, the main characteristic of the Peruvian labor market is the high incidence of underemployment that is closely linked with low income levels and low productivity. In effect, as it can be observed in Table 4.2, underemployment rates fluctuated in the second part of the 90's between 42% and 44% On the other hand; unemployment doesn't seem to be a massive problem since its rates were never above 8%.

Table 4.2
Labor Market Indicators
Peru 1990-1995

	95	96	97	98	99
Labor Force	100.0	100.0	100.0	100.0	100.0
Workforce	92.4	93.0	92.3	92.4	92.3
underemployment	42.4	42.6	41.8	44.1	43.4
adequately employed	50.0	50.4	50.5	48.3	48.9
Unemployment	7.6	7.0	7.7	7.6	7.7

4.1.1 Characteristics of the Labor Supply

In the case of the labor supply, the rate of growth of the working age population was declining since 1985. During 85-91, the average rate of growth per year for the urban sector was 4.8% and during 91-97 the rate reduced to 2.1%.

As a consequence, the age composition of the labor supply changed (Table 4.2). The participation of young workers declined in relation to older workers for both, male and female groups.

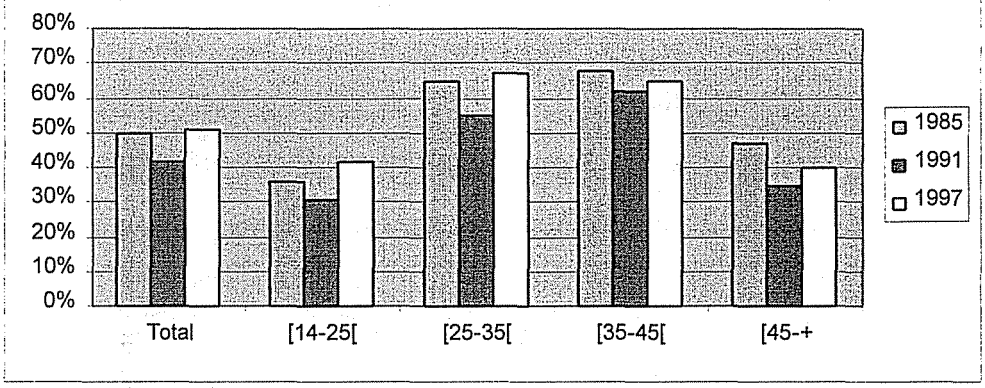
Table 4.2
Age Composition of the labor supply
1985-1997
(Urban sector)

	1985	**1997**
Male		
Under 25	37%	32%
Above 45	25%	30%
Female		
Under 25	36%	31%
Above 45	24%	29%

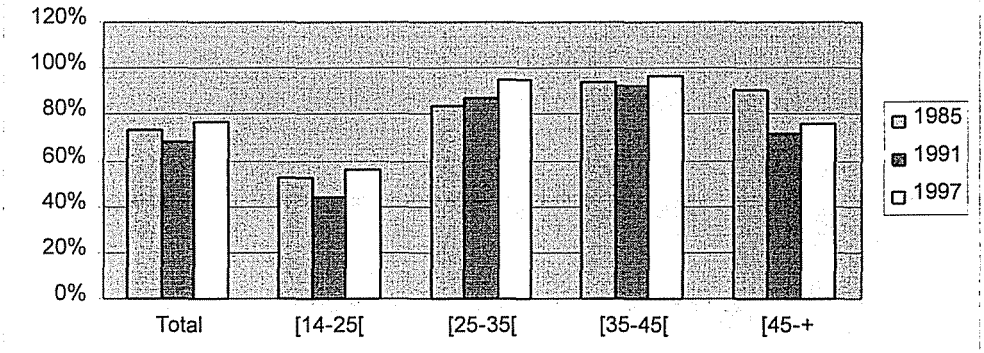
Source: Encuesta Nacional de Hogares sobre Niveles de Vida.
Taken from Saavedra and Muruyama (1999)

Despite the decline of the demographic pressure over the labor supply, it did not fall down but rather increased. The reason was the sustained increase of the labor participation rate. As can be observed in Graph 4.3, the period 1985-1991 the labor participation reduced hardly due to the economic crises of the end of the 80's and the adjustment program of the 90. However, the economic growth after 1992 increased the willingness of people to work and the participation rate augmented for the entire population and especially for young workers and females. As a consequence, the labor supply increased during the 90's.

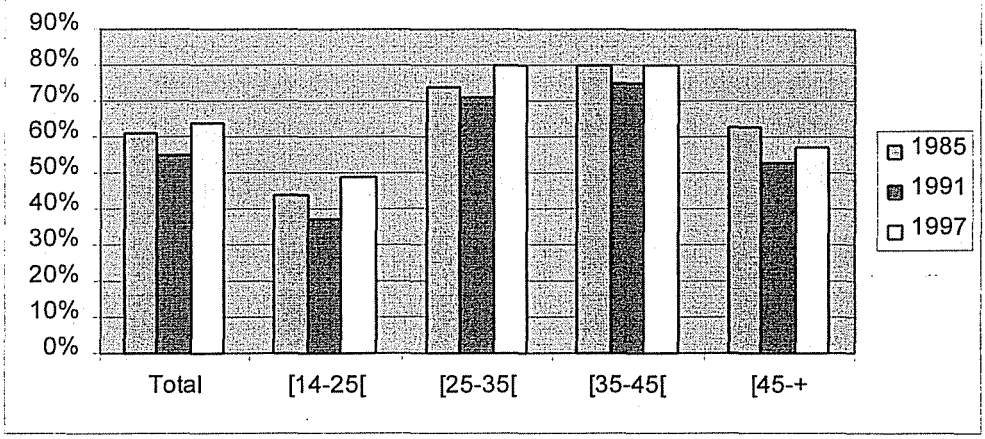
Graph 4.3
Labor Participation - Females by age group



Labor Participation - Males by age group



Labor Participation- Total by age group



Source: Encuesta Nacional de Hogares sobre Niveles de Vida.
Taken from Saavedra and Muruyama (1999)

Education of the labor supply

The level of education of the Peruvian population has increased sharply since the 50's. Between 1950 and 1997, enrolment in primary education increased 6.6 times while population increased in 3 times. Average years of education for the population of 15 years or more increased also from 1.9 to 8.6 and illiteracy decline from 58% to 11% in the same period (Wu, 2000). Education of the Peruvian labor supply followed the same pattern. As Table 4.3 shows, the proportion of workers with primary or no education decreased and the proportion of workers with tertiary education increased both, for males and females. In 1985, 30% of the male labor supply had primary or no education and in 1997 this proportion reduced to 23%. On the other hand, male workers with tertiary education increased from 21% to 29% in the same period. Female population shows the same patterns.

Table 4.3
Educational Level of the Peruvian Labor Supply by sex
(Urban sector)

	1985	1997
Primary/no education		
Male	30%	23%
Female	44%	29%
Tertiary Education		
Male	21%	29%
Female	15%	28%

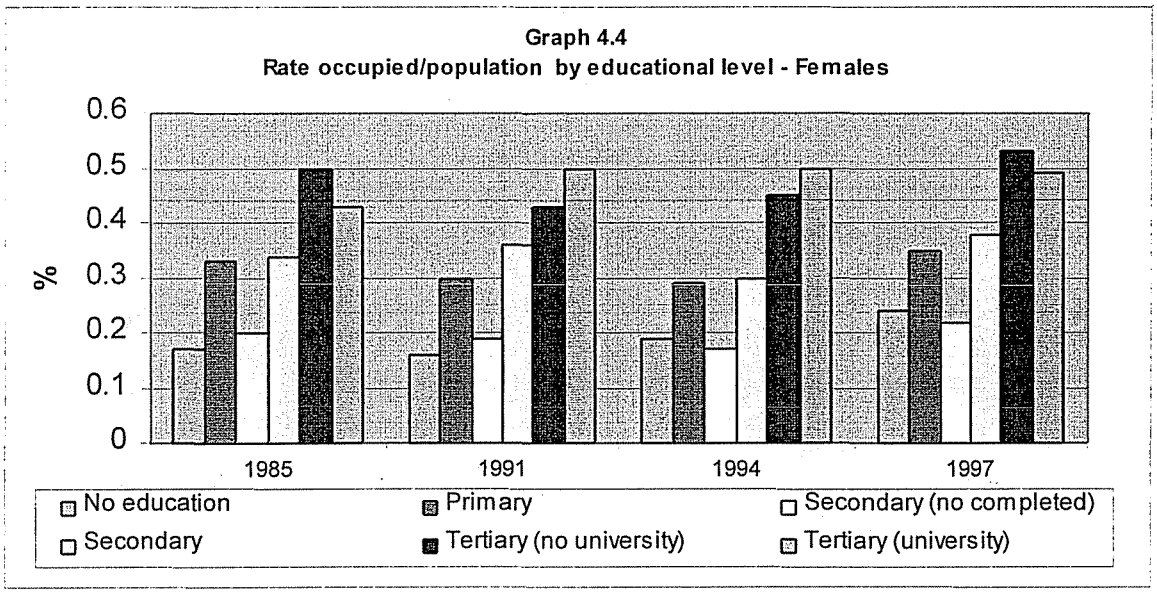
Source: Encuesta Nacional de Hogares sobre Niveles de Vida.
Taken from Saavedra and Muruyama (1999)

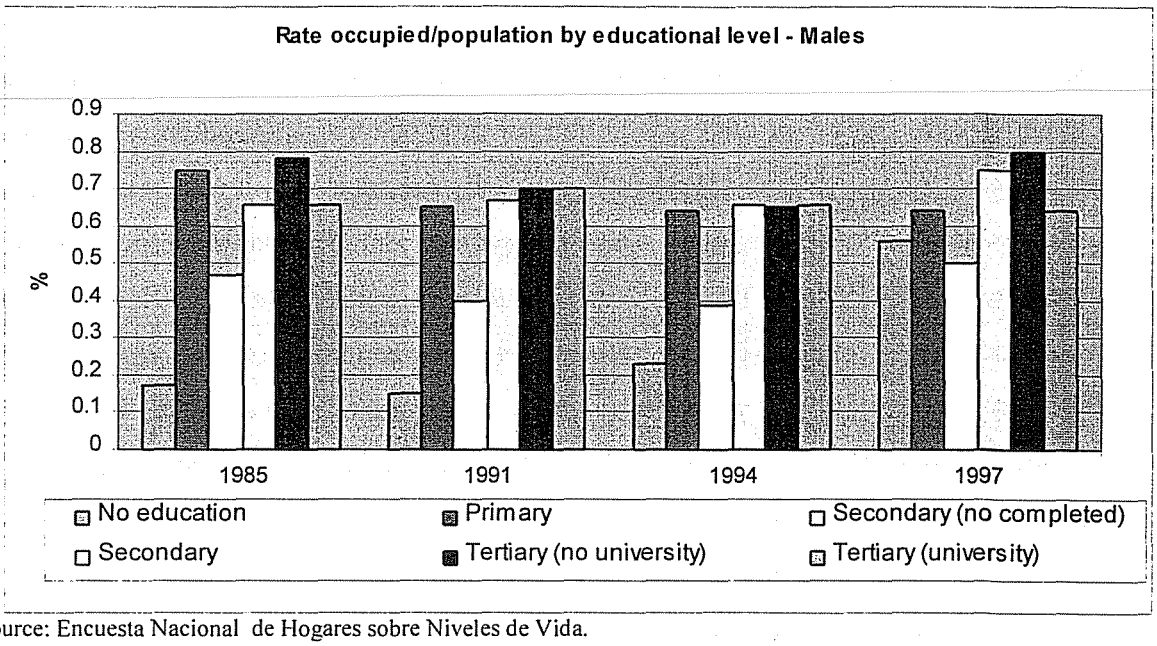
The structural reforms and the changes in the labor supply produced a new scenario for the Peruvian labor market in the 90's. First at all, the rate occupation/population increased for young workers (males and females) while the same ratio for older workers kept more stable and for male old workers even declined.

On the other hand, the rate of employment for female population increased sharply, especially for the younger group (less than 35). In general, the total volume of employment increased until 1997 but the opportunities where biased towards the young workers against the older ones.

The same analysis is done for the population divided by educational level and gender (Graph 4.4). In the case of male population the ratio occupation/population follows a pro cyclical pattern in the period 1985-1997, especially for the lowest levels of education. For workers with higher educational level, the pattern is more stable despite the strong economic fluctuations. It is worth to mention, however, that in 1997 the rate of employment for workers with university level is lower than the correspondent rate for secondary and tertiary no-university workers. Maruyama and Saavedra (1999) explain this phenomenon with the liberalization of the labor market that reduced the firing cost and allowed the subscription of temporary labor contracts. High firing costs and labor stability had created before an incentive for employers to hire workers with higher productivity and reduce the demand for less productive ones. The liberalization of the market increased the demand for less productive workers, especially the younger ones.

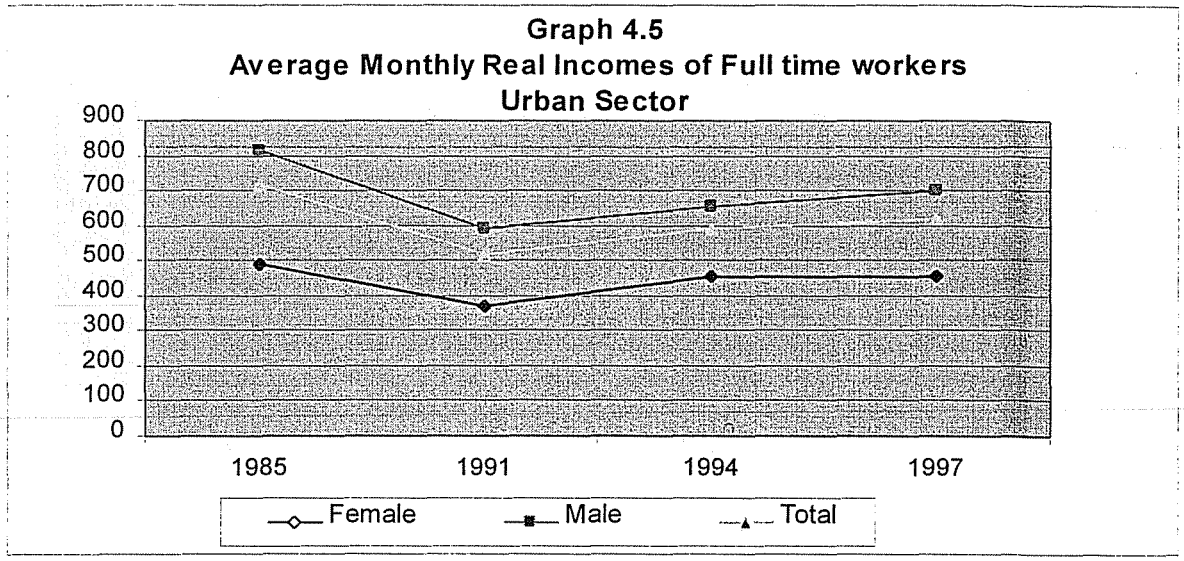
On the other side, for female population the pro cyclical pattern is less pronounced but it is more clear the positive relationship between employment and education. Like in the case of males, the rate of employment is stable for workers with university level but increases for the workers with tertiary no university level.





Source: Encuesta Nacional de Hogares sobre Niveles de Vida.
Taken from Saavedra and Muruyama (1999)
* Occupied are defined as the persons that work at least 10 hours per week and receive a kind of payment for that activity

Besides the reduction of the labor participation at the end of the 80's, real incomes also functioned as mechanism of adjustment in the labor market. In effect, as it is observed in Graph 4.5, real incomes reduced drastically during the period of crises (1987-1990). This is the reason why unemployment did not increase in this period.



Source: Encuesta Nacional de Hogares sobre Niveles de Vida
Taken from Saavedra and Muruyama (1999)
Income in *Nuevos soles* of June 1994

4.1.2 Characteristics of the Labor Demand

The structural reforms of the 90 led to a decline of the real exchange rate and to the increase of the relative price in the non tradable goods. In consequence, some non tradable sectors were the most dynamic in terms of employment (Diaz, Saavedra and Torero *et al.* 2000): finance sector, services to firms, transport and commerce had rates of employment growth above the average level. Other sectors, like non personal services, consultancy services and construction grew near the average level while manufacture and public administration were below it. As a consequence of this dynamic, the participation in employment of manufacture, personal services and public administration declined while the one of commerce, transportation and financial services increased (Table 4.3).

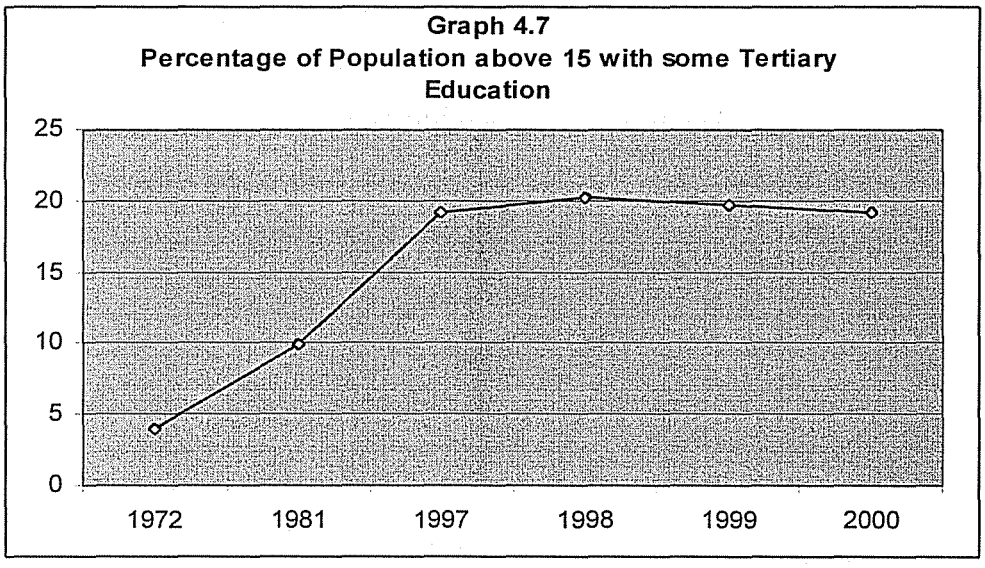
Table 4.3
Sectoral distribution of employment – Lima Metropolitana
1986-1997

	1986-89	**1992**	**1997**
Total	100	100	100
Agriculture	1.1	0.8	1.2
Mining	0.7	0.3	0.2
Manufacture	21.7	17.2	16.1
Electricity, gas, water	0.5	0.5	0.3
Construction	5.4	5.7	6.3
Commerce	29.9	33.7	35.8
Transport	6.6	6.9	9.0
Finances and services to firms	5.0	6.1	9.2
Personal services	23.2	23.9	18.3
Public administration	6.1	4.8	3.6

Source: Encuesta Nacional de Hogares 1986-1997
Taken from: Diaz, Saavedra and Torero (2000)

4.2 An overview to the Tertiary educational system in Peru

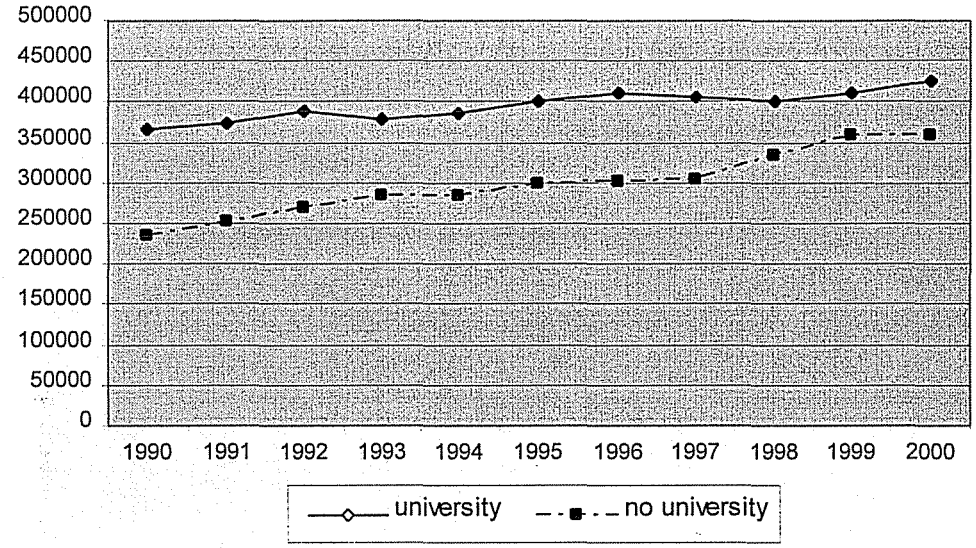
Tertiary education in Peru includes university one (at least 5 years of education) and no university one (2 or 3 years of education). The level of access to tertiary education has increased sharply in Peru in the last 30 years. In 1972 (Graph 4.7) just 4% of the population above 15 had some tertiary education. In 1981 this percentage increased to 10% and in the 90's the proportion was already around 20%.



Source: Peruvian National Institute of Statistics - INEI – ENAHO, Censos 1972, 1981

One feature of the tertiary educational system in Peru during the 90's is the higher increase in non university education. In effect, in 1990 (Graph 4.8) 235 304 students were registered in no university institutions, and in 2000 this number had increased to 359 783 i.e. an increase of 53%. The number of university students also increased but less sharply passing from 366 303 to 426 573 in the same period (16.5%).

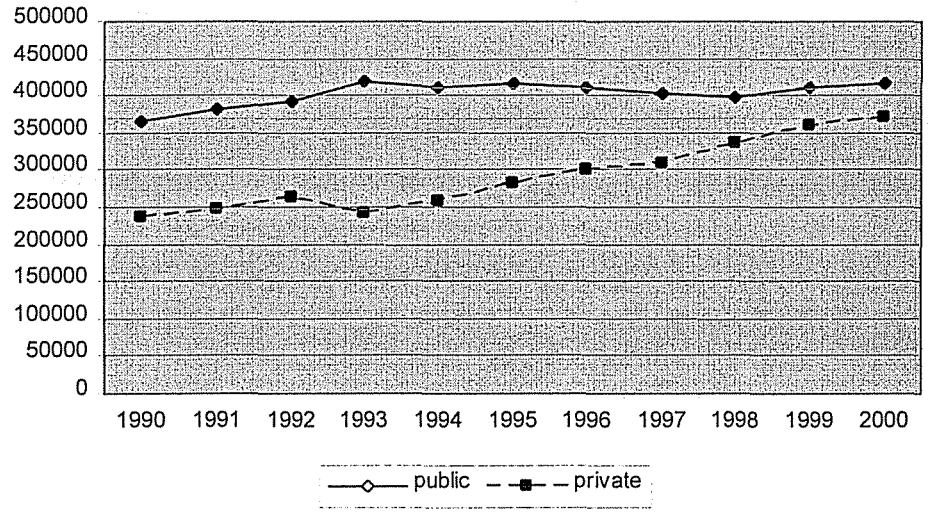
Graph 4.8
Number of Students in Tertiary Education
(university vs no university)



Source: Peruvian National Institute of Statistics - INEI

The other main characteristic of the 90's is the increasing number of students in private institutions. In 1990 (Graph 4.9) 235 405 students belonged to private centers and in 2000 this number increased to 370 439, an increase of 57.3%. The number of students in the public system also increased but in less proportion: from 366 202 to 415 917 (13.6%).

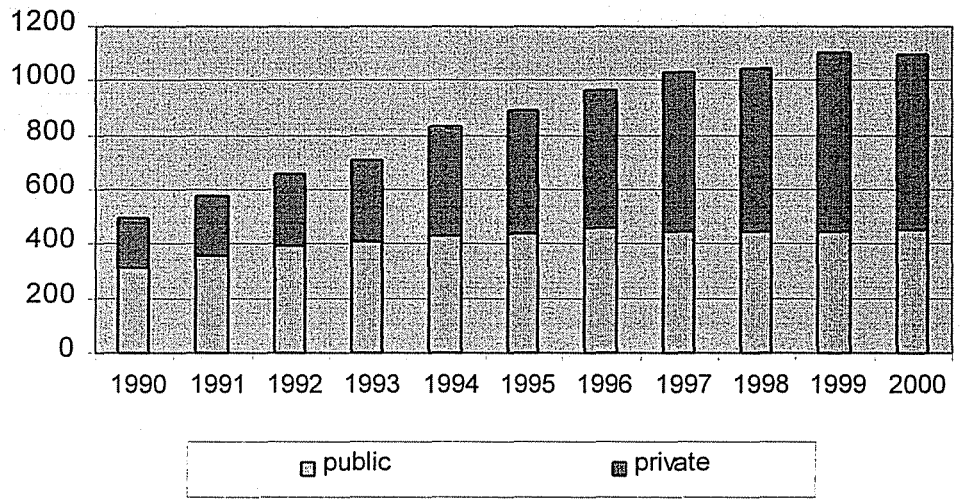
Graph 4.9
Number of Students in Tertiary Education
(public vs private)



Source: Peruvian National Institute of Statistics - INEI

The increase in the number of students involved in tertiary education was accompanied by an increase in the supply of services. In effect, as it is observed in Graph 4.10, the number of institutions that provide tertiary education in Peru increased from 499 in 1990 to 1100 in 2000.

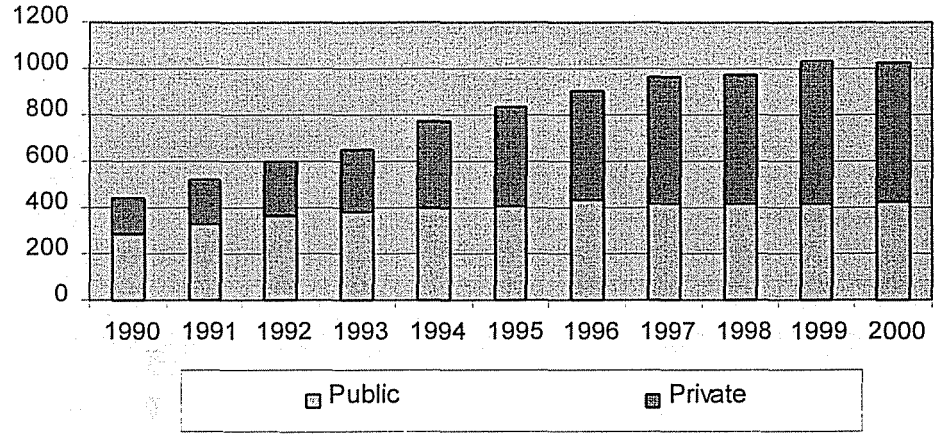
Graph 4.10
Number of Institutions of Tertiary education



Source: Peruvian National Institute of Statistics - INEI

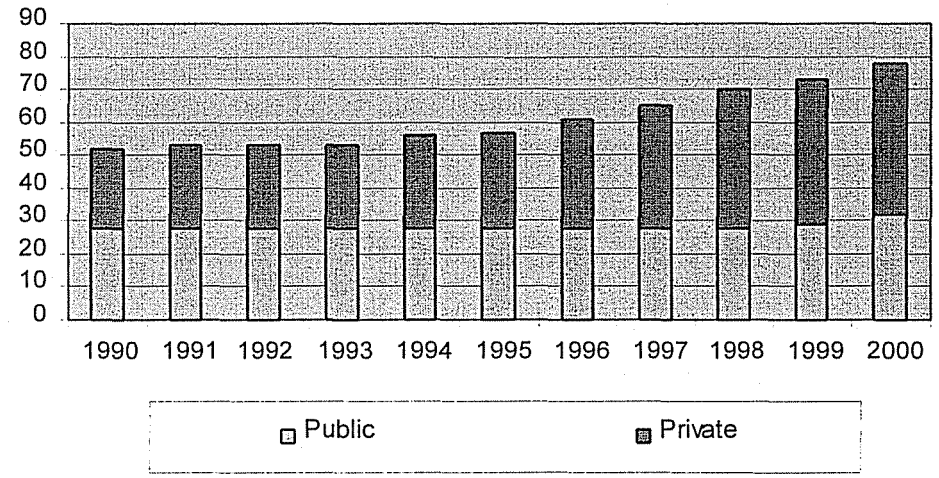
The increase was also much clearer in non university and private institutions. It is illustrated in Graph 4.11 and 4.12: while non university centers increased from 447 to 1022, universities increased from 58 to 78. Besides, private no university institutions increased from 164 to 600 and public no university ones did it only from 283 to 422. In the case of university institutions the increment was also biased towards private centers that grown from 24 to 46 in the period considered while public ones increased only from 28 to 32.

Graph 4.11
Number of Institutions of Tertiary No university education



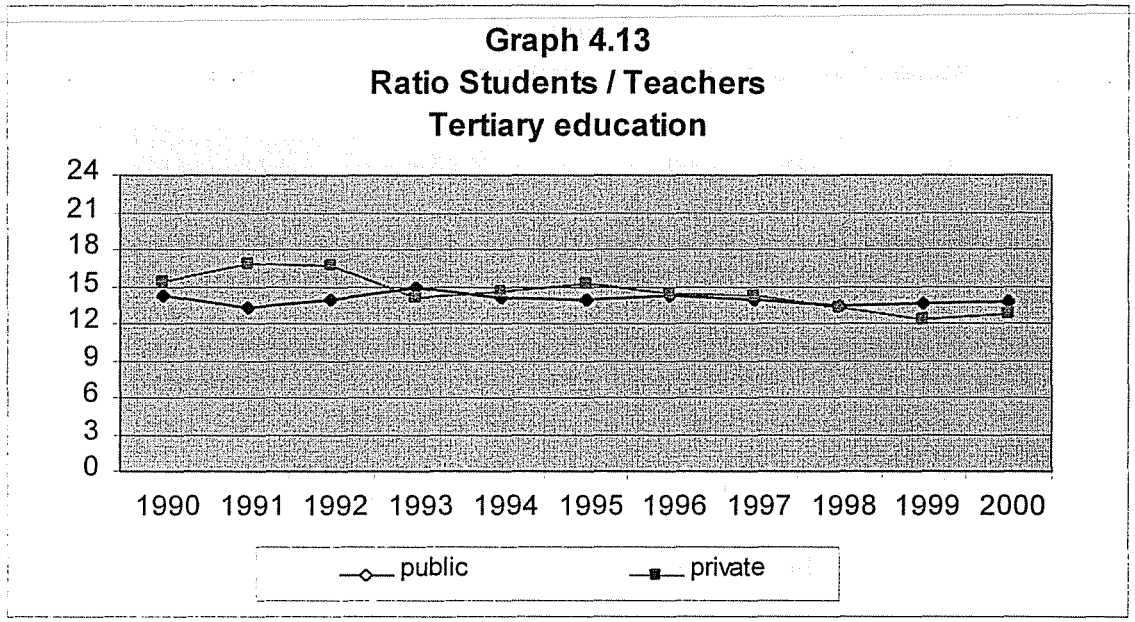
Source: Peruvian National Institute of Statistics - INEI

Graph 4.12
Number of Tertiary university institutions



Source: Peruvian National Institute of Statistics - INEI

In consequence, both, the demand and the supply of tertiary education increased during the 90's. The increment, however, was a little bit higher in the supply. In effect, Graph 4.13 shows that the ratio *students/teachers,* which can be interpreted as a measure of demand/supply and quality education, reduced from 14.2 to 13.7 in the public system and from 15.33 to 12.78 in the private one during the 90's.



Source: Peruvian National Institute of Statistics - INEI

4.3 Returns to education and overeducation

Many works have aimed to analyze returns to education in Peru. Saavedra and Maruyama (1999) used a typical *mincerian* equation to calculate education returns for 1985, 1991, 1994 and 1997[9] and found that the returns to education follow a pro cyclical pattern: from 1985 to 1991, the contractive period, the returns declined from 10.2% to 7.6% and in the 90's during the recovery period, reached 9.9% in 1991 and 10.4 in 1997. Besides, some particular groups have been benefited more than others in the 90's. That is the case of workers with university education whose educational return has increased more than the one of the groups with less education. The authors also found that returns to experience declined during the same period but that it was the more educated workers who were able to capitalize better the returns to experience suggesting their higher capacity to accumulate human capital. The authors evaluate also the return of education by income group and find that the return to tertiary education is higher for the workers that belong to the superior percentiles. According to them, this effect would be a consequence of the access to higher educational quality of the richest segments of the population and to

[9] The survey used correspond to the "Encuesta Nacional de Hogares sobre Niveles de Vida" and takes into account only main urban areas.

informational sources and social networks that allow them to get more easily jobs that rewards education more.

Finally it is found that there are differences depending on the educational system: workers that come from the public system shows less return than the ones that come from the private system. Differences in the socio economical background explain only partially this difference.

Diaz, Saavedra and Torero (2000) calculate returns to education using a different data set[10] and also find the increase in the return for most educated workers in the 90's. They explain it due to the increase of the demand of skilled workers, the increase in the use of new physical capital (whose price decline with the reforms) and the complementarity between both factors.

On the other hand, returns to overeducation are estimated in Burga *et al.* (2000) who use a mincerian equation to estimate the effect of being overeducated on earnings for the Peruvian urban labor market in 1997. The authors find that overeducated workers earn 21% more than adequate educated ones but only when the length of overeducation is more than 3 years.

However, it would be interesting to evaluate more deeply how the educational return of the overeducated worker is compared to the one of the worker with the same level of education but that is correctly allocated.

Finally Arregui (1993) makes an exploratory analysis of the mismatch between educational acquirement and the occupation[11] and find that correctly allocated workers with tertiary education have higher earnings than overeducated and undereducated workers.

[10] The authors use the "Encuesta Nacional de Hogares – ENAHO" and only for the capital city, Lima Metropolitana

[11] Arregui uses the "Encuesta Nacional de Hogares – ENAHO" for Lima Metropolitana and the years 1986 and 1991.

5. Empirical evaluation of Overeducation

This research develops two econometric models. The first one measures the effect of some economic and demographic variables on the probability to be overeducated and consists on a *probit* model. The second one measures the marginal effect of overeducation on wages. In this case a *semi log* earning function is carry on.

5.1 Determinants of overeducation

Table 5.1 shows the incidence of just educated, undereducated and overeducated workers for the years 1995-1999 following the methodology proposed by MTPE (2001) and described in Chapter 2.

Table 5.1
Incidence of Just educated, Undereducated and Overeducated workers by economic sector

	1995		
	Tradable	No tradable	Total
just educated	46.0	59.7	57.3
Undereducated	7.4	13.2	12.1
Overeducated	46.7	27.1	30.6

	1996		
	Tradable	No tradable	Total
just educated	47.3	57.3	54.8
Undereducated	8.5	11.4	10.7
Overeducated	44.2	31.3	34.5

	1997		
	Tradable	No tradable	Total
just educated	42.0	55.9	52.9
Undereducated	10.6	12.4	11.8
Overeducated	47.4	31.9	35.3

	1998		
	Tradable	No tradable	Total
just educated	44.0	56.1	53.7
Undereducated	9.0	15.0	13.7
Overeducated	47.1	28.9	32.6

	1999		
	Tradable	No tradable	Total
just educated	37.6	53.7	50.6
Undereducated	8.4	11.0	10.5
Overeducated	54.1	35.3	38.8

The first important feature to highlight is that the incidence of overeducation is much higher than the one of undereducation. In effect, the proportion of overeducated workers varies during the period between 30% and 38% while the same indicator for undereducated workers varies between 10% and 13%.

According to Job Competition Theory, workers who participate in the market during economic recessions would have more probabilities to be overeducated while during periods of economic growth it is expected a lower incidence of overeducation. As was mentioned in Chapter 4, 1995-97 was in Peru a period of growth and 1998-99 was a recessive one. The information available doesn't allow inferring any systematic relationship between overeducation and economic activity except for 1999, a recessive year that coincides with the highest incidence of overeducation among the period. However, it would be necessary to do a further analysis beyond the scope of this research in order to get more conclusive outcomes.

The second relevant point is the higher incidence of overeducation in the tradable sector compared with the non tradable one. As shown in Chapter 4, tradable sector was the one that lost participation during the 90's compared with the non tradable. The higher incidence of overeducation in tradable sector could be reflecting this performance. In the following section we evaluate if this difference is reflected also in labor earnings.

Taking into account the low proportion of undereducated workers compared with overeducated ones, the probit analysis considers only two categories, overeducated and not overeducated workers (that includes undereducated and just educated) and will try to find the determinants of overeducation. For this purpose, the *probit* model will take the following specification:

$$\Pr ob(Y_i = 1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\beta_1+\beta_2 X_i} e^{-t^2/2} dt$$

Where t ~ N (0,1) and the dependent variable Y_i takes the following values:

1 if worker *i* is overeducated

0 if worker *i* is just educated or undereducated

And X is a vector of socioeconomic and demographic characteristics that includes the following variables[12]:

1. **Education**

1.1 Amount of specific HC

Workers with more specific HC have less probability to be overeducated: It is expected that vocational education replaces formal education and workers that are overeducated compensate the lack of formal education with this training.

1.2. Level of acquired education

Workers with higher formal education are more likely to be overeducated. It is so because workers with higher education have a bigger range of jobs where they can be overeducated but a small range of jobs where they can be undereducated. The opposite happens with low educated workers.

1.3. Quality education

Following the Heterogeneous Labor Supply theory, workers with less quality education are more likely to be overeducated. In this study, quality education is approximated through public/private education since in Peru public system has less quality than the private one. One year of education in the public system has less market value than the private one. So, workers that come from public system have more probabilities to end working in jobs with less education requirement and being overeducated.

[12] Many of the determinants considered here are proposed by Wolbers (1999)

1.4 Access to social networks

According to Regional/Global labor markets approach, workers with less developed social networks are more likely to be overeducated since they have fewer information mechanisms to facilitate their incorporation to the global labor market where there are more job opportunities that fits with their level of education. The proxy used for this variable is parent's education.

2. Individual characteristics:

2.1 Gender

Women have less favorable prospects on the labor market than men. Their unemployment risk is larger, their opportunities for career mobility are smaller, their training participation is lower and their working life is more often interrupted by family obligations. The interruption of their labor life imposes a penalty in their process of human capital accumulation that would have to be compensated through other means like education in excess or, alternatively, there is a market penalty in the recognition of the number of schooling years that allocate females in jobs with lower education requirement that they actually have. Furthermore, according to Regional/Global labor market approach, women have less access to the global market which increases their probability of overeducation. In consequence, it would be more probable a higher incidence of overeducation in females.

2.2 Age

There are two opposite effects on the probability to be overeducated: 1. older workers would be more likely to be overeducated because the skills obtained in initial education may become obsolete (due to changing technology) and the relative value of vocational qualifications attended in initial education in the total amount of HC acquired decreases in the course of one's life. Because of this, old workers are assigned to jobs with lower educational requirements. 2. According to Occupational Mobility theory, young workers have more probabilities to be overeducated since workers follow a labor cycle where shortages of experience are compensated with education in excess. Besides, if it is considered the progressive deterioration of

quality education mentioned above, young workers would appear more likely as overeducated since less quality education has to be compensated with education in excess. In consequence, it is not possible to know a priori the final effect of age over overeducation.

2.3 Job tenure

Taking into account again Occupational Mobility Theory, it is expected a negative relationship between job tenure (the experience that a worker has in a specific job) and the likelihood of being overeducated.

2.4 Civil status

It is expected that married workers have higher probabilities to be overeducated since the familiar responsibilities reduces his/her time for job searching and it is less time consuming to find a job as overeducated than as just educated or undereducated worker. Besides, following Regional/Global labor market approach, married workers have also to fulfill reproductive activities in the household and their access to the global market, where there are more job opportunities in relation with the regional market, is more restricted.

3. Labor market structure determinants

3.2 Size of the firm

The probability to observe overeducated workers is smaller in bigger firms since they can provide more opportunities for individuals to find a job that matches their level of education.

3.3 Public and private sector

Educational and health workers use to have less job mismatch and, since in Peru these sectors are mainly public, the incidence of overeducation would be less in public sector.

3.4 Economic Sector

It is expected that declining sectors are more likely to have overeducated workers. As mentioned before, during the 90's the tradable sectors (mainly manufacture) declined and non tradable sectors increased. In consequence, it is expected a higher incidence of overeducation in tradable sectors, just as table 5.1 showed above.

The results[13] of the probit model are showed in Table 5.2.

Table 5.2

Probit Model

Dependent variable:	1995		1996*		1997		1998		1999	
Overeducation	Coef.	t-value	Coef.	z-value	Coef.	t-value	Coef.	t-value	Coef.	t-value
Vocational education	0.256	4.05	0.07	1.08	-0.10	-1.32	-0.13	-1.33	-0.03	-0.43
Formal education	0.151	13.41	0.19	16.24	0.25	17.26	0.19	10.89	0.20	14.07
Public education	n.a**	n.a	0.21	2.96	0.29	3.39	0.07	0.61	0.43	4.68
Parents education	n.a	n.a	-0.01	-2.15	n.a	n.a	n.a	n.a	n.a	n.a
Female	0.019	0.36	0.09	1.51	0.37	5.17	-0.04	-0.41	0.13	1.74
Age	-0.02	-4.9	-0.01	-1.72	-0.02	-4.56	-0.01	-3.11	-0.01	-3.90
Job tenure	-0.01	-2.92	-0.01	-2.63	-0.01	-1.91	-0.01	-0.75	-0.01	-1.22
Married	0.074	1.13	-0.06	-0.96	0.20	2.41	0.09	0.83	n.a	n.a
Size of the firm (>10)	-0.49	-7.1	-0.17	-1.96	-0.07	-0.96	-0.54	-5.11	-0.55	-5.96
Public sector	-0.71	-5.85	-0.93	-8.59	-1.25	-12.90	-0.48	-2.30	-0.67	-4.01
Tradable Sector	0.555	8.85	0.38	5.38	0.46	6.06	0.67	6.06	0.59	6.25
Constant	-1.73	-11.89	-2.39	12.02	-2.75	-13.48	-1.98	-8.02	-2.31	-10.62
F- value	43.53		n.a		40.88		16.33		27.14	
LR chi2			446.8							

*for 1996, the survey design is not available. The regression model is over the survey without weight.
** not available information

For some variables, the numeric value of the coefficients change in a relatively broad range along the period, however the signs are in most of the cases stable.

The variables *age* and *job tenure* affect negatively the probability to be overeducated which validate the predictions of Occupational Mobility theory disregarding the effect of technology changing mentioned above. Furthermore, the effect of *public education* on

[13] The complete model regressions are presented in the appendix.

overeducation is positive and also confirms the predictions of the Heterogeneous Labor Supply theory where higher incidence of overeducation is found in workers with less quality education. Finally, there are enough bases to state that females, married and workers with low schooling parents[14] (less social networks) are more likely to be overeducated which would confirm the predictions of the Regional/Global labor market theory.

In addition, the effects of formal education, tradable sector and the fact of being working in public sector and bigger firms are the expected ones in spite of the low significance of some coefficients for some years. Finally, the effect of vocational education is not clear. It could be consequence of the diverse degree of quality in the courses considered in the variable.

5.2 Effect of overeducation on wages

As explained in Chapter 3, different theories have different predictions about the returns of overeducation, required education and undereducation. Human Capital Theory – HCT suggests that productivity and earnings depends exclusively on the individual's characteristics and not in the properties of the job. Under this approach, Mincer (1974) proposes the following earning function specification:

$$Log(W) = \alpha_0 + \alpha_1 X + \alpha_2 X^2 + \alpha_3 E + \alpha_4 Z + \varepsilon_i \qquad (1)$$

Where W is wages, E is schooling years, X is labor experience and Z is a vector of personal characteristics. The coefficient of schooling years ($\alpha 3$) measures the rate of return for an additional year of education.

[14] For the case of married condition, the effect is positive and significant for 1997, positive but not significant for 1995 and 1998 and negative and not significant for 1996. For female condition, the effect is positive and significant for 1997, positive but not significant for 1995, 1996 1999 and negative and not significant for 1998.

On the other hand, Job Competition Theory – JCT states that wages depends on the characteristics of the job and not on the characteristics of the worker. The following specification is the relevant one for this approach:

$$Log(W) = \delta_0 + \delta_1 E^r + \varepsilon_2 \qquad (2)$$

Where (E^r) is the number of schooling years determined by the job characteristics.

Alternatively, Job Matching Theory integrates the previous ones suggesting that supply and demand characteristics matters to determine wages. In this case, the marginal return to each year of overeducation is positive but lower than the return to required education. To test the predictions of JMT the following specification is the appropriate one:

$$Log(W) = \beta_0 + \beta_1 E^r + \beta_2 E^0 + \beta_3 E^u + \beta_4 X + \varepsilon_3 \qquad (3)$$

Where W refers to labor earnings and the number of schooling years has been split into years required by the job (E^r), years of overeducation (E^o) and years of undereducation (E^u)[15] in the next way:

$$E = E^r + E^o - E^u$$

$$E^o = E - E^r \quad if\ E>Er$$
$$E0 = 0 \quad otherwise$$

$$E^u = E^r - E \quad if\ Er > E$$
$$E^u = 0 \quad otherwise$$

X is the vector of personal characteristics

[15] For this model it will be considered additionally *"years of undereducation"* in order to be consistent with the original ORU specification and regardless the low incidence of undereducation mentioned before.

This specification is called ORU (Over, Required and Undereducation- Hartog, 1988, 2000) and allows testing not only Job Matching Theory but also three other approaches: Human Capital, Job Competition and the Motivational Approach.

Sicherman (1991) interprets the coefficients of (3) in the following way:

β_1	=	The marginal return to and additional years of required education
β_2	=	The marginal return to and additional years of overeducation
β_3	=	The marginal penalty to each year of undereducation

It is expected that $\beta_1 > 0$, $\beta_2 > 0$ and $\beta_3 < 0$

Under HCT, wages depends only on workers characteristics and the returns to each year of education are the same. In terms of (3):

$$\beta_1 = \beta_2 = -\beta_3$$

On the other hand, under JCT, only job characteristics matters and years of education above (or under) the required by the job has no economic value (or penalty). In terms of (3):

$$\beta_2 = \beta_3 = 0$$

Under JMT, wages depends on both, job and worker characteristics. In terms of (3):

$$\beta_1 > \beta_2 > 0 > \beta_3$$

Finally, under the MA, overeducation has a negative impact over productivity and wages. In terms of (3):

$$\beta_1 > 0 > \beta_2$$

Besides the educational variables ER (years of required education), EO (years of overeducation) and EU (years of undereducation), the following explanatory variables are considered in the ORU earning function model:

1.	Labor earnings	That is presented as monthly labor earnings in logarithms.
2.	Experience and squared experience	It is expected that labor experience affects positively wages but with decreasing returns (sign of squared experience negative). In the model it is considered potential experience that is equivalent to: *age - years of education - 6.*
3.	Female condition	It is expected a negative relationship between female condition and labor earnings.
4.	Size of the firm	It is expected that bigger firms are associated with higher labor earnings. The variable is a categorical one that takes two values, "0" if the firm has until 10 workers and "1" if the firm has more than 10 workers.
5.	Vocational courses	It is expected a positive relationship with earnings. This variable is taken as a categorical one: "1" or "0" in the presence or absence of vocational courses respectively.
6.	Hours of work	This variable is taken as a proxy of labor effort ad it is expected a positive relationship with labor earnings.
7.	Public education	Like in the *probit* model above, public and private education are used as proxy to quality education and let us take into account the heterogeneity of the labor supply. It is expected a negative effect of public education on labor earnings

The results of the ORU model are reported in table 5.3[16]. It is considered two specifications, the conventional OLS and the function corrected by selection bias (Heckman). The reason to include selection bias correction is that the selection process (to participate in the labor market), can be related with overeducation: if no suitable employment is available, unemployment can be chosen as a strategy to avoid overeducation.

First at all, it is worth to mention that the correction of selection bias does not change the values of the coefficients significantly[17]. Furthermore the signs of the coefficients are in most of the cases the expected ones: higher levels of income are associated positively with education, size of the firm, experience and worked hours (proxy of effort); and negatively with female condition, potential experience squared and public education.

[16] The regression models take into account the survey design for 1995, 1997, 1998 and 1999 but not for 1996 due to the lack of weighted variable for this year.

[17] The variables considered for the selection model are civil condition (married), number of dependent children in the household and the existence of current transfers not linked with labor. The complete model is presented in the Appendix

Table 5.3

ORU Earning Function Model

Dependent:	1995		1996		1997		1998		1999	
Labor Incomes (log)	OLS	Heckman	OLS	Heckman	OLS	Heckman	OLS	Heckman	OLS	Heckman
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
ER	0.13 *	0.12 *	0.09 *	0.08 *	0.10 *	0.09 *	0.10 *	0.10 *	0.10 *	0.10 *
EO	0.08 *	0.07 *	0.01	0.00	0.01	0.01	0.03 *	0.03 *	-0.001	0.00
EU	-0.09 *	-0.09 *	-0.08 *	-0.08 *	-0.08 *	-0.08 *	-0.07 *	-0.07 *	-0.10 *	-0.10 *
female	-0.39 *	-0.38 *	-0.25 *	-0.25 *	-0.40 *	-0.40 *	-0.30 *	-0.31 *	-0.24 *	-0.23 *
size of the firm (>10)	0.31 *	0.30 *	-0.05	-0.02	0.11 *	0.11 *	0.31 *	0.31 *	0.40 *	0.40 *
vocational education	0.08 *	0.07	0.11 *	0.1 *	0.11 *	0.10 *	0.03	0.02	0.00	0.00
potential experience	0.04 *	0.04 *	0.02 *	0.02 *	0.03 *	0.02 *	0.03 *	0.02 *	0.03 *	0.03 *
potential experience 2	-0.0005 *	-0.0005 *	-0.0004 *	-0.0003 *	-0.0004 *	-0.0004 *	-0.0004 *	-0.0004 *	-0.0004 *	-0.0004 *
worked hours	0.01 *	0.01 *	0.02 *	0.02 *	0.02 *	0.01 *	0.01 *	0.01 *	0.02 *	0.02 *
public education	n.a	n.a	-0.22 *	-0.22 *	-0.15 *	-0.16 *	-0.23 *	-0.23 *	-0.23 *	-0.23 *
_cons	3.76	3.99	4.57 *	4.81 *	4.68 *	4.84 *	4.85 *	4.95 *	4.57 *	4.59 *
R2	0.46		0.4007		0.4047		0.4041		0.4973	
F	168.29		141.06		104.44		57.09		109.28	
Wald Test of indep. equations		15.24		7.09		2.44		0.01		0.17

* Significant at 95%

5.2.1 Hypotheses Testing

Table 5.4 shows the results obtained when the two opposite theories, Human Capital (HC) and Job Competition Theory (JCT) hypothesis are tested.

In the case of HC, the hypothesis $\beta(EO) = \beta(ER)$ is rejected[18] for all the years considered. However, the hypothesis $-\beta(EU) = \beta(ER)$ can not be rejected for the entire period. In consequence, returns to years of overeducation would be different (actually less) from returns to years of required education while penalties to undereducation would be similar (in absolute value) to returns to required education.

In the case of JCT hypothesis, mixed results are obtained. First at all, for 1996, 1997 and 1999, it is not possible to reject the hypothesis $\beta(EO) = 0$ which would give support to accept JCT predictions. However, JCT also requires the coefficient of $\beta(EU)$ to be zero and this hypothesis is rejected for the whole period.

In consequence, there would be no returns to overeducation but there would be penalties for being undereducated. Furthermore, these marginal penalties are similar in absolute value to the marginal returns of required education. In other words, job characteristics would determine labor incomes, as JCT predicts, except if the worker is undereducated when his/her level of education becomes more relevant.

On the other hand, for 1995 and 1998, the hypothesis that the coefficient of EO is zero is rejected and, since it is also lower than ER's one, this outcome allow us to validate JMT predictions: both, job and workers characteristics would matter determining labor earnings.

Finally, MA hypothesis is rejected for the whole period. For any of the years it is found that years of overeducation have negative marginal returns: overeducation doesn't impose penalties to labor earnings.

---

[18] In this research, the null hypothesis is rejected if the probability of committing Type 1 error is higher than 0.05.

Table 5.4
Testing ORU Hypothesis

OLS model

Test	95	96	97	98	99
Human Capital Theory					
EO=ER	Rejected	Rejected	Rejected	Rejected	Rejected
-EU=ER	Rejected	Accepted	Accepted	Rejected	Accepted
Job Competition Theory					
EO=0	Rejected	Accepted	Accepted	Rejected	Accepted
EU=0	Rejected	Rejected	Rejected	Rejected	Rejected

Heckman Model

Test	95	96	97	98	99
Human Capital Theory					
EO=ER	Rejected	Rejected	Rejected	Rejected	Rejected
-EU=ER	Rejected	Accepted	Accepted	Accepted	Accepted
Job Competition Theory					
EO=0	Rejected	Accepted	Accepted	Rejected	Accepted
EU=0	Rejected	Rejected	Rejected	Rejected	Rejected

5.2.2 Segmented Market Analysis

The results obtained so far regarding the returns to overeducation, required education and undereducation are not uniform along the whole period. Depending on the year considered, JCT or JMT seems to be the most appropriate theories while HC theory doesn't appear as the explicative one for any of the years.

This outcome could be hiding the possible existence of segmented labor markets where different markets function under different patterns. As mentioned in Chapter 3, Doeringer (1971) and Sthincombe (1979) suggest that public and private sectors function in different ways. While private sector would function more as HC predicts, public sector would allocate workers as JCT or JSM do.

Besides, it is also possible to distinguish between formal and informal sector. There are some arguments that would lead us to think that the informal sector is more competitive such as the fact that it is a non regulated market. If that is the case, returns to overeducation would be closer to returns to education compared with the formal sector.

Finally, it is considered tradable and non tradable sectors. As was observed before, tradable sector shows a significant higher incidence of overeducation. It is tested if this difference is also present in the returns to overeducation and required education.

The main limitation for the segmented market analysis is the difficulty of correcting by selection bias since it would be necessary to predict in which of the segments would be employed the potential worker. However, the close values of the coefficients obtained before from the OLS model and the Heckman model would allow us to rely on the results of the conventional OLS.

5.2.2.1 Public and Private Sectors

Table 5.5 and 5.6 show the OLS coefficients[19] of ER, EO and EU and the results of the hypothesis testing for private and public sectors separately. The most interesting feature is presented in the case of the public sector where in four of the five years it is not possible to reject the HC hypothesis. In consequence, most of the available evidence indicates that in the public sector the returns to education are similar to the returns to overeducation which means that labor incomes depends mostly on workers characteristics. This outcome is contrary to the one expected for Segmented Market Approach but is similar to the one obtained by Van der Meer (2001) for the case of The Netherlands.

In the case of private sector, it presents a behavior closer to JMT since HC and JCT hypothesis are rejected for most of the years[20] and the returns to overeducation are positive.

[19] The tables with the complete regression models are presented in the Appendix

[20] With the exception of 1999 where the hypothesis $\beta(EO)=0$ can not be rejected

Table 5.5

ORU Earning Model for Private and Public Sector

Private Sector

	1995		1996		1997		1998		1999	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
ER	0.14	15.38	0.09	11.07	0.11	10.55	0.14	12.16	0.14	13.25
EO	0.08	5.84	0.03	2.03	0.07	5.56	0.05	2.84	0.01	0.81
EU	-0.13	-7.83	-0.12	-7.82	-0.07	-3.68	-0.11	-5.58	-0.14	-6.82

Public Sector

	1995		1996		1997		1998		1999	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
ER	0.09	5.26	0.07	3.39	0.10	6.69	0.07	2.75	0.14	5.03
EO	0.07	1.93	0.07	1.77	0.07	2.82	0.05	1.12	0.20	2.96
EU	-0.14	-5.88	-0.08	-2.61	-0.10	-4.31	-0.06	-2.67	-0.06	-1.48

Table 5.6

Comparing Private and Public Sectors

Private

Test	95	96	97	98	99
Human Capital Theory					
EO=ER	Rejected	Rejected	Rejected	Rejected	Rejected
-EU=ER	Accepted	Accepted	Accepted	Accepted	Accepted
Job Competition Theory					
EO=0	Rejected	Rejected	Rejected	Rejected	Accepted
EU=0	Rejected	Rejected	Rejected	Rejected	Rejected

Public

Test	95	96	97	98	99
Human Capital Theory					
EO=ER	Accepted	Accepted	Rejected	Accepted	Accepted
-EU=ER	Accepted	Accepted	Accepted	Accepted	Accepted
Job Competition Theory					
EO=0	Accepted	Accepted	Rejected	Accepted	Rejected
EU=0	Rejected	Rejected	Rejected	Rejected	Accepted

5.2.2.2 Formal and Informal sectors

Table 5.7 and 5.8 show the OLS coefficients of ER, EO and EU and the results of the hypothesis testing for formal and informal[20] labor markets separately.

The most interesting point to highlight is that, as expected, in the informal segment it is not possible to reject HCT hypothesis for most of the years (1995, 96 and 97) while formal segment closely reproduces the patterns showed before for the integrated market. Regardless of the necessity of further analysis about the functioning of formal and informal labor markets, the outcomes obtained are in line with the expected ones and would indicate that informal market functions closer to the competitive one compared with formal market. In the informal segment, there is a relatively higher importance of worker's characteristics determining labor earnings while job characteristics seem to be less important.

Table 5.7

ORU Earning Model for Formal and Informal Sectors

Formal

	1995		1996		1997		1998		1999	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
ER	0.11	11.34	0.06	6.48	0.08	7.08	0.10	6.55	0.10	6.22
EO	0.08	4.44	0.01	0.90	0.05	2.82	0.03	1.35	0.06	2.14
EU	-0.11	-6.48	-0.11	-6.76	-0.05	-2.54	-0.07	-2.90	-0.08	-2.41

Informal

	1995		1996		1997		1998		1999	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
ER	0.08	8.39	0.05	3.31	0.06	4.36	0.10	5.24	0.08	5.58
EO	0.06	4.47	0.02	1.30	0.04	2.72	0.04	1.92	-0.01	-0.64
EU	-0.07	-3.97	-0.02	-0.71	-0.09	-3.37	-0.09	-2.74	-0.09	-3.80

[20] It is considered that a worker is participating in the formal market if he/she has a labor contract. For 1995 this information is not available so the criterion is if the worker receives a receipt of payments.

Table 5.8
Comparing Formal and Informal Sectors

Formal

Test	95	96	97	98	99
Human Capital Theory					
EO=ER	Rejected	Rejected	Rejected	Rejected	Rejected
-EU=ER	Accepted	Rejected	Accepted	Accepted	Accepted
Job Competition Theory					
EO=0	Rejected	Accepted	Rejected	Accepted	Rejected
EU=0	Rejected	Rejected	Rejected	Rejected	Rejected

Informal

Test	95	96	97	98	99
Human Capital Theory					
EO=ER	Accepted	Accepted	Accepted	Rejected	Rejected
-EU=ER	Accepted	Accepted	Accepted	Accepted	Accepted
Job Competition Theory					
EO=0	Rejected	Accepted	Rejected	Accepted	Accepted
EU=0	Rejected	Accepted	Rejected	Rejected	Rejected

5.2.2.3 Tradable and No Tradable sectors

Table 5.9 and 5.10 show the OLS coefficients of ER, EO and EU and the results of the hypothesis testing for tradable and non tradable sectors.

It is observed that the patterns in both segments are similar each other. As in the case of the integrated market, in tradable and non tradable segments, HCT hypothesis that $\beta(EO) = \beta(ER)$ is rejected for the whole period but the hypothesis $-\beta(EU) = \beta(ER)$ can not be rejected consistently. On the other hand, testing JCT hypothesis show also, as in the integrated market, mixed outcomes. It is not possible to reject that marginal returns to overeducation are zero for all the years but, on the other hand, the penalties to years of undereducation are significant. In consequence, tradable and non tradable sectors show a similar pattern regarding the returns to education.

Table 5.9

ORU Earning Model for Tradable and Non Tradable Sectors

Tradable

	1995		1996		1997		1998		1999	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
ER	0.13	12.91	0.10	8.62	0.15	8.05	0.13	6.30	0.12	5.55
EO	0.06	4.44	0.03	1.98	0.06	3.08	0.04	1.62	0.01	0.47
EU	-0.07	-4.01	-0.07	-3.14	-0.10	-2.98	-0.07	-1.93	-0.11	-2.96

No tradable

	1995		1996		1997		1998		1999	
	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value	Coeff.	t-value
ER	0.13	17.17	0.08	13.41	0.09	13.61	0.10	9.09	0.09	10.12
EO	0.08	8.47	0.00	-0.26	0.004	0.35	0.04	2.60	0.00	0.20
EU	-0.09	-5.79	-0.09	-7.15	-0.09	-5.23	-0.07	-3.83	-0.10	-5.88

Table 5.10
Comparing Tradable and Non Tradable Sector

Tradable

Test	95	96	97	98	99
Human Capital Theory					
EO=ER	Rejected	Rejected	Rejected	Rejected	Rejected
-EU=ER	Rejected	Accepted	Accepted	Accepted	Accepted
Job Competition Theory					
EO=0	Rejected	Rejected	Rejected	Accepted	Accepted
EU=0	Rejected	Rejected	Rejected	Accepted	Rejected

No Tradable

Test	95	96	97	98	99
Human Capital Theory					
EO=ER	Rejected	Rejected	Rejected	Rejected	Rejected
-EU=ER	Rejected	Accepted	Accepted	Accepted	Accepted
Job Competition Theory					
EO=0	Rejected	Accepted	Accepted	Rejected	Accepted
EU=0	Rejected	Rejected	Rejected	Rejected	Rejected

In conclusion, the results of the analysis of educational returns are much clearer when the segmented approach is introduced. In this sense, HC seems to be the most appropriate theory explaining the functioning of public and informal sectors while JMT is more powerful explaining the performance of private sector. Finally, tradable and non tradable sectors do not seem to function in different ways regarding the retribution to education and overeducation; both of them perform in most of the cases as JMT predicts. Table 5.11 summarizes the findings regarding ORU earning function model and the hypotheses testing.

Table 5.11

	Human Capital Theory	**Job Competition Theory**	**Job Matching Theory**	**Motivational Approach**
Integrated Market	Never Accepted	Accepted in some of the cases	Accepted in some cases	Never Accepted
Private Sector	Never Accepted	Almost never Accepted	Accepted in most of the cases	Never Accepted
Public Sector	Accepted in most of the cases	Almost never Accepted	Almost never Accepted	Never Accepted
Formal Sector	Never Accepted	Almost never Accepted	Accepted in most of the cases	Never Accepted
Informal Sector	Accepted in most of the cases	Accepted in some cases	Never Accepted	Never Accepted
Tradable Sector	Never Accepted	Accepted in some of the cases	Accepted in most of the cases	Never Accepted
Non Tradable Sector	Never Accepted	Accepted in some of the cases	Accepted in most of the cases	Never Accepted

5.3 Implications of the results obtained

The outcomes obtained from the empirical analysis can give support to draw some policy implications regarding labor and educational markets.

If Job Competition theory is accepted, it would imply that the high incidence of overeducation reflects the existence of a relative excess of supply of high educated

workers. In this case, the skills that workers have acquired during their education will be wasted[21]. One policy response to such a situation would be to reduce the investment in education and the numbers of students receiving higher education since the economy is producing too many of them. According to the results obtained this could be the case for some particular economic sectors where effectively it seems to be an excess of labor supply. However there is some additional evidence that lead us to think that it is not a generalized phenomenon. As observed in Chapter 4, the wage inequality between high educated and low educated workers in Peru increased during the 90's. This pattern is contrary to what would be expected if there is an excess supply of educated workers.

On the other hand, if Human Capital Theory is the valid one then all the educational background and skills of the workers would be fully utilized in the market and overeducation would be only an apparent phenomenon that hides the lack of other forms of human capital, such labor experience, and intrinsic or acquired differences in worker's quality. In this case there wouldn't be so much concern about active labor policies. The results obtained show that, in effect, overeducation is highly linked with labor experience and low quality workers. However; the results of the earning function do not validate the predictions of Human Capital for most of the sectors of the economy: there is an effective economic penalty for being overeducated.

Finally, the acceptance of Job Matching theory implies that overeducation reflects a degree of mismatch in the labor market. Workers would not be assigned to the most appropriate jobs due to problems in the access to information or rigidities in the labor market. Restrictions in the access to labor information make it difficult to find the best fitted jobs for worker's skills and rigidities may prevent them the acceptance of these jobs. In this case, the policy recommendation would be oriented to the improvement of the inefficiencies of the labor market providing information mechanisms that help agents (mostly young people) to decide the level and kind of investment in education they should do and implementing policies oriented to promote the incorporation of certain population groups (females and married, for instance) to best fitted jobs. In general it

[21] Without considering the non pecuniary returns of education.

would be important to consider not only policies oriented to the increase of the population's human capital but also to develop educational (sectoral) planning policies. The outcomes obtained from the empirical evidence give enough support to accept the Job Matching Theory as the most appropriate framework to understand the functioning of the Peruvian labor market

Conclusions

Overeducation refers to a situation in which individuals have schooling level in excess of the educational level required by their job. The measurement of overeducation is not straightforward and there are many approaches that try to approximate this phenomenon. The method of Job Analysis, based on the analysis of occupational codes, is recognized as the most appropriate one. This research used the measurement proposed by the Peruvian Ministry of Labor that is built under the Job Analysis method.

The present study analyzed the phenomenon of overeducation for the Peruvian case during the period 1995-1999. It was found that during this period, between 30.6% and 38.8% of the Peruvian working force was overeducated while the incidence of undereducated workers was much lower: between 10.5% and 13.7%. From the data available, it was not found any systematic pattern between overeducation and economic activity. However further analysis are necessary to get more conclusive conclusions about this relationship.

Two fields often present in the overeducation literature were studied: the socioeconomic and demographic determinants of overeducation and the economic returns to overeducation.

In the case of the analysis of overeducation determinants, a *probit* model was used to evaluate the predictions of three theories about the effect of certain variables on the probability of being overeducated. The Occupational Mobility Theory predicts a higher incidence of overeducation in younger and less experienced workers. The Heterogeneous Labor Supply Theory postulates differences in the quality of the workers beyond the quantity of education received: workers that receive less quality education (like those that study in Peru in the public system compared with the private one) would be more likely to be overeducated. Finally, the Regional/Global labor market approach postulates that the restricted access to regional markets increases the probability of being overeducated. Usually, females, married and workers with less developed social networks are the ones

that have more problems accessing to the global market and would show higher incidence of overeducation. The predictions of the three theories were confirmed empirically. Besides, other characteristics such size of the firm, the fact of being working in the public sector and in the non tradable one reduces the probability to be overeducated.

In the case of the economic return to overeducation, the analysis was done using the so called ORU earning function and in the frame of four theories that have different explanations about the functioning of the labor market and the returns to education: Human Capital Theory, Job Competition Theory, Job Matching Theory and the Motivational Approach. The objective was to determine which theory fitted better for the Peruvian case.

The econometric model yielded mixed results: The Human Capital Theory and the Motivational Approach predictions were rejected for all the years considered while the hypotheses of the Job Competition ad Job Matching theories were confirmed in most of the cases.

The mixed results obtained for the integrated labor market motivated a segmented market analysis where it was argued that each segment functions under different logic. There were considered three segments: Private and public, formal and informal and tradable and non tradable sectors.

The outcomes of the segmented market analysis were more conclusive than the previous one. It was found that the private and the formal sectors function as the Job Matching Theory predicts and that the public and the informal sector function closer to the Human Capital approach. On the other hand, tradable and non tradable sector do not seem to have differences in the retribution to overeducation. In most of the cases analyzed they function as the Job Matching Theory predicts.

Finally, some possible policy recommendations were drawn. The high incidence of overeducated workers is an inefficient use of the human resources of the economy where there is an economic penalty for being overeducated. To increase the efficiency of the market, it would be recommendable to establish information mechanisms that help agents to decide their investment in education. It would be also important to implement policies

to promote the incorporation of certain population groups to best fitted jobs and to develop educational-sectoral planning policies.

References

ALBA RAMÍREZ, Alfonso
1993 "Mismatch in the Spanish Labor Market: Overeducation?", Journal of Economic Resources, 28(2): 259-278.

ARREGUI, Patricia
1993 "Empleo, ingresos y ocupación de los profesionales y técnicos en el Perú", in Notas para el Debate, 9: 9-48. Lima: GRADE.

BECKER, Gary
1980 "Human Capital". Chicago: University of Chicago Press.

BUCHEL, F.
2000 "The Effects of Overeducation on Productivity in Germany: The Firms' Viewpoint". IZA DP # 216. Institute for the Study of Labor – IZA. Bonn

BURGA, Cybele and MORENO, Martin
2001 "Existe Subempleo profesional en el Peru urbano?" Investigaciones Breves 17. Lima, Grade.

CLOGG, Clifford and James W. SHOCKEY
1984 "Mismatch Between Occupational and Schooling: A Prevalence Measure, Recent Trends and Demographic Analysis", Demography, 21: 235-257.

DIAZ, Juan Jose, SAAVEDRA, Jaime and TORERO, Maximo
2000 "Liberalizacion de la Balanza de Pagos, Efectos sobre el Crecimiento, el Empleo y Desigualdad y Pobreza. El caso de Peru". Lima, Grade

DUNCAN, Greg J. and Saul HOFFMAN
1981 "The Incidence of Wage Effects of Overeducation", Economics of Education Review, 1(1): 57-86.

FREEMAN, Richard B.
1976 The Overeducated American. New York: Academic Press.

GREEN, F., MCINTOSH, S and VIGNOLES, A.
1999 "Overeducation and Skills: Clarifying the Concepts". Centre for Economic Performance.

HALABY, Charles
1994 "Overeducation and Skills Mismatch". Sociology of Education, 67(1): 47-59

HARTOG, Joop

2000 "Overeducation and earnings: where we are, where should we go?, Economic of Education Review, vol 19: 131-148

MOULLET, S.
1999 "Overeducation, Undereducation and allocation on the French labour market". IDEP-LEST

RUMBERGER, R.
1981 "The rising incidence of overeducation in the US labor market", in Economics of Education Review, 1, (3), 293-314

RUMBERGER, R.
1987 "The Impact of Surplus Schooling on Productivity and Earnings", Journal of Human Resources, 22(1): 25-50.

SAAVEDRA, Jaime and MURUYAMA, Eduardo
1999 "Los retornos a la educación y a la experiencia en el Peru"
in "Pobreza y Economia Social: analisis de una encuesta ENNIV-1997". Lima, Instituto Cuanto.

SATTINGER, M.
1993 "Assignment models of the distribution of Earnings", Journal of Economic Literature, XXXI, 831-880

SICHERMAN, Nachum
1991 "Overeducation in the Labor Market". Journal of Labor Economics, vol. 9(2): 101-122.

SPENCE, M.
1973 "Job Market signaling" in Quarterly Journal of Economics , n 87, pp355-374

TINBERGEN, J.
1956 "On the Theory of Income Distribution", Weltwirtschaftliches Archiv, 77: 156-175.

TSANG, M.
1987 "The impact of underutilization of education on productivity: a case study of the US Bell Companies" in Economics of Education Review, vol. 6, n 3, pp 239-254.

TSANG, M; RUMBERGER, R. and LEVIN, H.
1991 "The impact of surplus schooling on work productivity" in Industrial Relations, 30, 209-228

VAN DER MEER, P. and GLEBBECK, A.

2001 "Bureaucracy, overschooling and credentials inflation". Paper presented in: 'Market Expansion, Welfare State Retrenchment and their Impact on Social Stratification', April 26-28, 2001, Manheim, Germany.

VERDUGO, Richard R. and Naomi TURNER VERDUGO
1989 "The Impact of Surplus Schooling on Earnings: Some Additional Findings", Journal of Human Resources, 24(4): 629-643.

WOLBERS, M.
2002 "Job mismatches and their labour market effects among school-leavers in Europe". Research Centre for Education and the Labour Market (ROA), Maastricht University.

APPENDIX

PROBIT MODEL

1995

Survey probit regression

pweight: pesoper
Strata: upm
PSU: usm

Number of obs = 2825
Number of strata = 1
Number of PSUs = 193
Population size = 1996713.1
F(9, 184) = 43.53
Prob > F = 0.0000

overeducat~n	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
voc. edu.	.256404	.0632952	4.05	0.000	.1315607	.3812472
edu. formal	.1507478	.0112395	13.41	0.000	.1285791	.1729165
female	.0187416	.0522001	0.36	0.720	-.0842176	.1217008
age	-.0150713	.0030732	-4.90	0.000	-.0211328	-.0090097
tenure	-.0129454	.0044315	-2.92	0.004	-.0216861	-.0042047
married	.0743895	.0656539	1.13	0.259	-.0551061	.2038851
size2	-.4948407	.0697235	-7.10	0.000	-.6323631	-.3573182
public	-.7068496	.1207669	-5.85	0.000	-.9450497	-.4686495
tradable	.5546487	.0626369	8.85	0.000	.4311038	.6781936
_cons	-1.729757	.1455362	-11.89	0.000	-2.016812	-1.442702

1996

Probit estimates

Number of obs = 2195
LR chi2(11) = 446.11
Prob > chi2 = 0.0000
Log likelihood = -1202.8271 Pseudo R2 = 0.1564

overeducat~n	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
voc. edu.	.0657824	.0610942	1.08	0.282	-.0539601	.185525
edu. formal	.1979864	.012133	16.32	0.000	.1742061	.2217667
public_sch	.2084127	.0723896	2.88	0.004	.0665318	.3502936
eduparents	-.0065109	.0031257	-2.08	0.037	-.0126371	-.0003846
female	.0854631	.0621486	1.38	0.169	-.036346	.2072722
age	-.0045964	.003355	-1.37	0.171	-.011172	.0019793
tenure	-.0144041	.0049468	-2.91	0.004	-.0240996	-.0047086
married	-.0624399	.0743763	-0.84	0.401	-.2082148	.0833351
size2	-.1743889	.0652711	-2.67	0.008	-.302318	-.0464598
public	-.9291464	.1089431	-8.53	0.000	-1.142671	-.7156218
tradable	.380381	.0707955	5.37	0.000	.2416243	.5191377
_cons	-2.39992	.1966659	-12.20	0.000	-2.785378	-2.014462

1997

Survey probit regression

```
pweight:  pesoper                                 Number of obs    =      2107
Strata:   <one>                                   Number of strata =         1
PSU:      <observations>                          Number of PSUs   =      2107
                                                  Population size  =   2097797
                                                  F(  10,   2097)  =     40.88
                                                  Prob > F         =    0.0000

------------------------------------------------------------------------------
overeducat~n |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   voc. edu. |  -.1008476   .0765221    -1.32   0.188    -.2509143    .0492191
 edu. formal |   .2489096   .0144208    17.26   0.000     .2206291    .2771901
      female |   .3668948   .0710235     5.17   0.000     .2276112    .5061783
         age |  -.0164822   .0036113    -4.56   0.000    -.0235642   -.0094001
      tenure |  -.0114981   .0060047    -1.91   0.056    -.0232738    .0002776
       cives |   .1991972   .0825951     2.41   0.016     .0372206    .3611738
       size2 |  -.0694095   .0724448    -0.96   0.338    -.2114803    .0726614
      public |  -1.252904   .0971027   -12.90   0.000    -1.443331   -1.062477
    tradable |   .4644543   .0766968     6.06   0.000     .3140449    .6148637
  public_sch |   .2936969   .0866138     3.39   0.001     .1238394    .4635544
       _cons |  -2.746769      .2037   -13.48   0.000    -3.146243   -2.347294
------------------------------------------------------------------------------
```

1998

Survey probit regression

```
pweight:  pesoper                                 Number of obs    =      1153
Strata:   <one>                                   Number of strata =         1
PSU:      <observations>                          Number of PSUs   =      1153
                                                  Population size  = 2108015.5
                                                  F(  10,   1143)  =     16.33
                                                  Prob > F         =    0.0000

------------------------------------------------------------------------------
overeducat~n |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   voc. edu. |   -.125787   .0948991    -1.33   0.185    -.3119815    .0604075
 edu. formal |   .1896423   .0174161    10.89   0.000     .1554714    .2238132
      female |   -.036838   .0907536    -0.41   0.685    -.2148989     .141223
         age |  -.0149572   .0048082    -3.11   0.002     -.024391   -.0055234
      tenure |  -.0061993   .0082792    -0.75   0.454    -.0224434    .0100448
       cives |   .0854131   .1035091     0.83   0.409    -.1176744    .2885005
       size2 |  -.5416393    .106088    -5.11   0.000    -.7497865    -.333492
      public |  -.4805055   .2092242    -2.30   0.022    -.8910087   -.0700023
    tradable |   .6709867   .1107236     6.06   0.000     .4537443    .8882292
  public_sch |   .0680472   .1122974     0.61   0.545    -.1522831    .2883776
       _cons |  -1.974985   .2464099    -8.02   0.000    -2.458447   -1.491522
------------------------------------------------------------------------------
```

1999

Survey probit regression

```
pweight:  pesoper                                 Number of obs    =      1468
Strata:   <one>                                   Number of strata =         1
PSU:      <observations>                          Number of PSUs   =      1468
                                                  Population size  = 2314869.5
                                                  F(   9,   1459)  =     27.14
                                                  Prob > F         =    0.0000
```

overeducat~n	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
voc. edu.	-.0341596	.0791455	-0.43	0.666	-.18941	.1210907
edu. formal	.1968837	.0139964	14.07	0.000	.1694286	.2243388
female	.1330809	.0763835	1.74	0.082	-.0167517	.2829134
age	-.0127961	.0032837	-3.90	0.000	-.0192374	-.0063548
tenure	-.0083898	.0068621	-1.22	0.222	-.0218505	.0050708
size2	-.5453122	.0915209	-5.96	0.000	-.7248379	-.3657864
public	-.6655303	.1658394	-4.01	0.000	-.9908379	-.3402227
tradable	.5912227	.0946081	6.25	0.000	.4056411	.7768043
public_sch	.4306492	.0919652	4.68	0.000	.2502518	.6110466
_cons	-2.305079	.2170594	-10.62	0.000	-2.730859	-1.879299

EARNING FUNCTION

1995

Survey linear regression

pweight: pesoper
Strata: upm
PSU: usm

Number of obs	= 2763
Number of strata	= 1
Number of PSUs	= 193
Population size	= 1953188.2
F(9, 184)	= 168.29
Prob > F	= 0.0000
R-squared	= 0.4682

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.1272645	.0065294	19.49	0.000	.114386	.140143
EO	.0764671	.0077908	9.82	0.000	.0611007	.0918336
EU	-.086141	.012759	-6.75	0.000	-.1113069	-.0609751
female	-.3917503	.0333387	-11.75	0.000	-.4575075	-.3259931
size2	.308447	.0319667	9.65	0.000	.2453961	.371498
voc. edu.	.0803637	.0289108	2.78	0.006	.0233402	.1373872
expot	.0438566	.0035621	12.31	0.000	.0368308	.0508824
expot2	-.0005117	.0000716	-7.15	0.000	-.0006529	-.0003706
hours	.010411	.0007653	13.60	0.000	.0089015	.0119204
_cons	3.757025	.0917815	40.93	0.000	3.575995	3.938055

1996

Source	SS	df	MS		
Model	654.33138	10	65.433138	Number of obs	= 2121
Residual	978.789935	2110	.463881486	F(10, 2110)	= 141.06
				Prob > F	= 0.0000
Total	1633.12131	2120	.770340243	R-squared	= 0.4007
				Adj R-squared	= 0.3978
				Root MSE	= .68109

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0859897	.0053136	16.18	0.000	.0755693	.0964101
EO	.0066837	.0081205	0.82	0.411	-.0092413	.0226087
EU	-.078596	.0104428	-7.53	0.000	-.0990752	-.0581168
female	-.2492589	.0309412	-8.06	0.000	-.3099373	-.1885806
size2	-.010888	.0326428	-0.33	0.739	-.0749034	.0531273
voc. edu.	.1059381	.0303874	3.49	0.000	.0463457	.1655304
expot	.0244262	.0033907	7.20	0.000	.0177767	.0310757
expot2	-.0003554	.0000686	-5.18	0.000	-.00049	-.0002208
hours	.0184667	.0007433	24.84	0.000	.017009	.0199243
public_sch	-.2209641	.0371629	-5.95	0.000	-.2938438	-.1480844
_cons	4.570468	.0943022	48.47	0.000	4.385533	4.755403

1997

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs = 2028
Number of strata = 1
Number of PSUs = 2028
Population size = 2019776.5
F(10, 2018) = 104.44
Prob > F = 0.0000
R-squared = 0.4047

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0994867	.0063397	15.69	0.000	.0870536	.1119198
EO	.0146146	.0097218	1.50	0.133	-.0044511	.0336804
EU	-.0847186	.0151303	-5.60	0.000	-.1143912	-.055046
female	-.3970492	.035497	-11.19	0.000	-.4666635	-.3274348
size2	.1056296	.0372986	2.83	0.005	.0324821	.1787771
voc. edu.	.1094614	.03744	2.92	0.003	.0360366	.1828862
expot	.0257866	.0036328	7.10	0.000	.0186621	.0329111
expot2	-.0004303	.0000702	-6.13	0.000	-.0005681	-.0002925
hours	.0154031	.000979	15.73	0.000	.0134832	.017323
public_sch	-.1519354	.046948	-3.24	0.001	-.2440067	-.0598641
_cons	4.68385	.1097814	42.67	0.000	4.468553	4.899146

1998

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs = 1119
Number of strata = 1
Number of PSUs = 1119
Population size = 2044260
F(10, 1109) = 57.09
Prob > F = 0.0000
R-squared = 0.4041

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.102225	.0093939	10.88	0.000	.0837933	.1206568
EO	.0341618	.0120242	2.84	0.005	.0105693	.0577544
EU	-.0688356	.01663	-4.14	0.000	-.1014652	-.0362061

female	-.3016071	.0466737	-6.46	0.000	-.393185	-.2100291
size2	.3103614	.0538158	5.77	0.000	.2047701	.4159528
voc. edu.	.0298009	.0496697	0.60	0.549	-.0676555	.1272572
expot	.0257411	.0051343	5.01	0.000	.0156672	.0358151
expot2	-.0003786	.000103	-3.67	0.000	-.0005807	-.0001764
hours	.011764	.0013398	8.78	0.000	.0091351	.0143929
public_sch	-.2270415	.0629741	-3.61	0.000	-.3506022	-.1034807
_cons	4.852915	.1702629	28.50	0.000	4.518845	5.186986

1999

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 1416
Number of strata	= 1
Number of PSUs	= 1416
Population size	= 2237802.1
F(10, 1406)	= 109.28
Prob > F	= 0.0000
R-squared	= 0.4973

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0974276	.0082796	11.77	0.000	.081186	.1136692
EO	-.0007496	.0113077	-0.07	0.947	-.0229313	.021432
EU	-.1031449	.0159213	-6.48	0.000	-.1343768	-.071913
female	-.2366492	.0390056	-6.07	0.000	-.3131643	-.1601342
voc. edu.	.004688	.0442065	0.11	0.916	-.0820294	.0914054
size2	.3987769	.0470459	8.48	0.000	.3064898	.4910641
expot	.0321709	.0052437	6.14	0.000	.0218846	.0424572
expot2	-.0004273	.0001135	-3.76	0.000	-.0006499	-.0002046
hours	.0188704	.0011091	17.01	0.000	.0166948	.0210461
public_sch	-.2290244	.0517666	-4.42	0.000	-.330572	-.1274769
_cons	4.572173	.1309936	34.90	0.000	4.31521	4.829135

EARNING FUNCTION CORRECTED BY SELECION BIAS (HECKMAN)

1995

Heckman selection model
(regression model with sample selection)

Number of obs	= 4319
Censored obs	= 1583
Uncensored obs	= 2736
Wald chi2(9)	= 1768.29
Prob > chi2	= 0.0000

Log pseudo-likelihood = -2807715

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
logingtot						
ER	.1228088	.004875	25.19	0.000	.1132539	.1323637
EO	.0723429	.0077952	9.28	0.000	.0570646	.0876212
EU	-.0853254	.0121448	-7.03	0.000	-.1091288	-.061522
female	-.3753884	.0287353	-13.06	0.000	-.4317085	-.3190683

size2	.3024931	.0313392	9.65	0.000	.2410693	.3639169
voc. edu.	.0738549	.02833	2.61	0.009	.0183291	.1293807
expot	.041453	.0033929	12.22	0.000	.034803	.0481029
expot2	-.0004682	.0000699	-6.70	0.000	-.0006052	-.0003313
hours	.0076952	.0009912	7.76	0.000	.0057525	.0096379
_cons	3.989828	.0922958	43.23	0.000	3.808931	4.170724
select						
ER	.0606946	.0107546	5.64	0.000	.0396161	.0817732
EO	.0465745	.0194939	2.39	0.017	.0083672	.0847819
EU	-.0040892	.019879	-0.21	0.837	-.0430514	.0348729
female	-.4577497	.0616849	-7.42	0.000	-.5786499	-.3368495
size2	-.1945131	.062741	-3.10	0.002	-.3174832	-.071543
voc. edu.	.0145152	.0609499	0.24	0.812	-.1049444	.1339747
expot	.0400911	.0074775	5.36	0.000	.0254355	.0547467
expot2	-.0007121	.0001318	-5.40	0.000	-.0009705	-.0004537
hours	.0632201	.0037268	16.96	0.000	.0559157	.0705245
married	-.2284004	.0764582	-2.99	0.003	-.3782557	-.0785451
i_other	-.4357025	.0659094	-6.61	0.000	-.5648825	-.3065225
num_hij_dep	-.0915624	.0343376	-2.67	0.008	-.1588629	-.0242618
_cons	-1.246153	.1701965	-7.32	0.000	-1.579732	-.9125735
/athrho	-.2852956	.073073	-3.90	0.000	-.4285159	-.1420752
/lnsigma	-.3541955	.0192676	-18.38	0.000	-.3919592	-.3164318
rho	-.2777991	.0674337			-.4040803	-.1411269
sigma	.7017378	.0135208			.6757317	.7287447
lambda	-.1949421	.048296			-.2896005	-.1002838

Wald test of indep. eqns. (rho = 0): chi2(1) = 15.24 Prob > chi2 = 0.0001

1996

Heckman selection model
(regression model with sample selection)

Number of obs	= 2173
Censored obs	= 90
Uncensored obs	= 2083
Wald chi2(10)	= 1073.85
Prob > chi2	= 0.0000

Log likelihood = -2455.086

	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
logingtot						
ER	.0796326	.0054975	14.49	0.000	.0688576	.0904076
EO	.0045229	.0082997	0.54	0.586	-.0117442	.02079
EU	-.0772622	.0107218	-7.21	0.000	-.0982767	-.0562478
female	-.254051	.0318675	-7.97	0.000	-.3165103	-.1915918
size2	-.0224747	.0331206	-0.68	0.497	-.0873898	.0424404
voc. edu.	.0894226	.0312313	2.86	0.004	.0282103	.1506349
expot	.0183979	.0035576	5.17	0.000	.0114251	.0253706
expot2	-.0002651	.0000708	-3.74	0.000	-.0004039	-.0001263
hours	.0171083	.0007772	22.01	0.000	.015585	.0186317
public_sch	-.2227913	.0379072	-5.88	0.000	-.2970881	-.1484946
_cons	4.813554	.1020632	47.16	0.000	4.613514	5.013594
select						
ER	.0889665	.0244209	3.64	0.000	.0411024	.1368305
EO	.0038511	.0284582	0.14	0.892	-.0519259	.059628
EU	.0059841	.0469325	0.13	0.899	-.0860019	.0979702

female	-.4028458	.109799	-3.67	0.000	-.6180478	-.1876438
size2	.0163005	.1284707	0.13	0.899	-.2354974	.2680983
voc. edu.	.3104515	.1087958	2.85	0.004	.0972157	.5236873
expot	.0603942	.0138917	4.35	0.000	.033167	.0876214
expot2	-.0009123	.0002425	-3.76	0.000	-.0013876	-.0004369
hours	.00951	.0027405	3.47	0.001	.0041387	.0148813
public_sch	-.0032028	.146301	-0.02	0.983	-.2899474	.2835419
married	-.2057175	.1451088	-1.42	0.156	-.4901256	.0786905
i_other	.5024738	.1479908	3.40	0.001	.212417	.7925305
num_hij_dep	.0003347	.0011523	0.29	0.771	-.0019237	.0025932
_cons	.0888337	.3744683	0.24	0.812	-.6451107	.8227782
/athrho	-.6871353	.1621918	-4.24	0.000	-1.005025	-.3692453
/lnsigma	-.3678559	.0181969	-20.22	0.000	-.4035212	-.3321907
rho	-.5961385	.1045519			-.7636966	-.3533314
sigma	.6922169	.0125962			.6679639	.7173505
lambda	-.4126572	.0765658			-.5627234	-.262591

LR test of indep. eqns. (rho = 0): chi2(1) = 7.09 Prob > chi2 = 0.0078

1997

Heckman selection model
(regression model with sample selection)

Number of obs = 2105
Censored obs = 109
Uncensored obs = 1996

Wald chi2(10) = 816.03
Log pseudo-likelihood = -2636460 Prob > chi2 = 0.0000

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
logingtot						
ER	.0948852	.0065677	14.45	0.000	.0820126	.1077577
EO	.0109294	.0098913	1.10	0.269	-.0084573	.0303161
EU	-.0796149	.0148076	-5.38	0.000	-.1086373	-.0505925
female	-.4012143	.0363119	-11.05	0.000	-.4723843	-.3300442
size2	.1144818	.0377166	3.04	0.002	.0405586	.1884051
voc. edu.	.0988289	.0377131	2.62	0.009	.0249126	.1727452
expot	.0224902	.0038103	5.90	0.000	.0150222	.0299582
expot2	-.0003899	.0000727	-5.36	0.000	-.0005323	-.0002474
hours	.0145436	.0010285	14.14	0.000	.0125278	.0165593
public_sch	-.1592834	.0471093	-3.38	0.001	-.2516161	-.0669508
_cons	4.838536	.1186318	40.79	0.000	4.606021	5.07105
select						
ER	.0769968	.0293425	2.62	0.009	.0194866	.1345069
EO	.0188097	.0254551	0.74	0.460	-.0310814	.0687008
EU	.002456	.0454693	0.05	0.957	-.0866623	.0915742
female	-.443692	.1122344	-3.95	0.000	-.6636674	-.2237166
size2	.017041	.1174975	0.15	0.885	-.2132499	.2473318
voc. edu.	.1170472	.1344443	0.87	0.384	-.1464587	.3805531
expot	.046928	.0115532	4.06	0.000	.0242841	.0695719
expot2	-.0006442	.0001818	-3.54	0.000	-.0010005	-.0002878
hours	.0139025	.0029157	4.77	0.000	.0081879	.0196171
public_sch	.1180167	.1340223	0.88	0.379	-.1446621	.3806956
cives	-.1588228	.1205487	-1.32	0.188	-.395094	.0774484

i_other	-.2586744	.1631892	-1.59	0.113	-.5785193	.0611705
num_hij_dep	.0002063	.0053153	0.04	0.969	-.0102115	.0106241
_cons	.1856707	.4080151	0.46	0.649	-.6140242	.9853656
/athrho	-.2239662	.1434449	-1.56	0.118	-.5051129	.0571806
/lnsigma	-.2745256	.0212885	-12.90	0.000	-.3162503	-.232801
rho	-.220295	.1364835			-.4661287	.0571184
sigma	.7599325	.0161778			.728877	.7923113
lambda	-.1674094	.1046905			-.3725989	.0377802

Wald test of indep. eqns. (rho = 0): chi2(1) = 2.44 Prob > chi2 = 0.1184

1998

Heckman selection model
(regression model with sample selection)

Number of obs	1158
Censored obs	51
Uncensored obs	1107
Wald chi2(10)	514.51
Prob > chi2	0.0000

Log pseudo-likelihood = -2514518

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
logingtot						
ER	.100168	.0094208	10.63	0.000	.0817035	.1186325
EO	.0300668	.0120694	2.49	0.013	.0064114	.0537223
EU	-.0706782	.0167272	-4.23	0.000	-.1034629	-.0378934
female	-.3052023	.0475976	-6.41	0.000	-.3984919	-.2119126
size2	.305315	.0547489	5.58	0.000	.1980091	.412621
voc. edu.	.0214405	.0496756	0.43	0.666	-.0759218	.1188028
expot	.0239494	.0052176	4.59	0.000	.0137231	.0341757
expot2	-.0003553	.0001038	-3.42	0.001	-.0005587	-.000152
hours	.011134	.0013693	8.13	0.000	.0084502	.0138177
public_sch	-.2284791	.0632822	-3.61	0.000	-.3525099	-.1044484
_cons	4.94557	.1739694	28.43	0.000	4.604596	5.286544
select						
ER	.0427928	.028027	1.53	0.127	-.0121391	.0977247
EO	.0381263	.0412169	0.93	0.355	-.0426574	.11891
EU	-.0623196	.0455513	-1.37	0.171	-.1515985	.0269594
female	-.5154015	.1605702	-3.21	0.001	-.8301132	-.2006898
size2	.7880005	.2072224	3.80	0.000	.3818522	1.194149
voc. edu.	.0591266	.1853728	0.32	0.750	-.3041974	.4224506
expot	.0356073	.0185777	1.92	0.055	-.0008043	.0720188
expot2	-.0005902	.0003449	-1.71	0.087	-.0012662	.0000857
hours	.0152492	.0036438	4.18	0.000	.0081075	.0223909
public_sch	-.0344776	.1952274	-0.18	0.860	-.4171163	.3481611
cives	.2995895	.1972081	1.52	0.129	-.0869313	.6861103
i_other	.0711757	.1529183	0.47	0.642	-.2285386	.37089
num_hij_dep	-.0737162	.0253477	-2.91	0.004	-.1233969	-.0240355
_cons	.6403334	.4853325	1.32	0.187	-.3109008	1.591568
/athrho	-.0124821	.1292106	-0.10	0.923	-.2657303	.240766
/lnsigma	-.3234683	.0309355	-10.46	0.000	-.3841008	-.2628358
rho	-.0124815	.1291905			-.2596474	.2362192
sigma	.7236349	.022386			.6810628	.7688681
lambda	-.009032	.0934587			-.1922078	.1741437

Wald test of indep. eqns. (rho = 0): chi2(1) = 0.01 Prob > chi2 = 0.9230

1999

Heckman selection model
(regression model with sample selection)

Number of obs = 1456
Censored obs = 61
Uncensored obs = 1395

Wald chi2(10) = 645.68
Log pseudo-likelihood = -2766958 Prob > chi2 = 0.0000

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
logingtot						
ER	.0975698	.0083648	11.66	0.000	.081175	.1139645
EO	.0021533	.0116113	0.19	0.853	-.0206044	.0249109
EU	-.1042512	.0166324	-6.27	0.000	-.1368501	-.0716523
female	-.2289107	.0402856	-5.68	0.000	-.3078691	-.1499524
size2	.4011569	.0564686	7.10	0.000	.2904805	.5118332
voc. edu.	-.0011903	.0451146	-0.03	0.979	-.0896133	.0872327
expot	.0317126	.0059681	5.31	0.000	.0200154	.0434098
expot2	-.000415	.0001216	-3.41	0.001	-.0006534	-.0001766
hours	.0186357	.001216	15.33	0.000	.0162524	.021019
public_sch	-.2341039	.052116	-4.49	0.000	-.3362494	-.1319585
_cons	4.587044	.1610435	28.48	0.000	4.271404	4.902684
select						
ER	.0390051	.0283312	1.38	0.169	-.016523	.0945333
EO	-.0280328	.0332131	-0.84	0.399	-.0931293	.0370637
EU	-.0353199	.0477068	-0.74	0.459	-.1288235	.0581837
female	-.2056884	.1401148	-1.47	0.142	-.4803084	.0689316
size2	.9060145	.2195735	4.13	0.000	.4756584	1.336371
voc. edu.	.1290655	.1525206	0.85	0.397	-.1698693	.4280003
expot	.0625156	.0153729	4.07	0.000	.0323854	.0926458
expot2	-.0009533	.0002787	-3.42	0.001	-.0014996	-.000407
hours	.010041	.0033969	2.96	0.003	.0033833	.0166988
public_sch	.1258357	.1792557	0.70	0.483	-.225499	.4771705
i_other	.1533729	.148733	1.03	0.302	-.1381384	.4448843
num_hij_dep	.0311744	.1100821	0.28	0.777	-.1845824	.2469313
_cons	.2932797	.4598425	0.64	0.524	-.607995	1.194554
/athrho	-.1984599	.4775294	-0.42	0.678	-1.1344	.7374806
/lnsigma	-.3016445	.0259883	-11.61	0.000	-.3525807	-.2507083
rho	-.1958947	.4592044			-.8125199	.6276206
sigma	.7396009	.019221			.7028719	.7782493
lambda	-.1448839	.3414452			-.8141042	.5243364

Wald test of indep. eqns. (rho = 0): chi2(1) = 0.17 Prob > chi2 = 0.6777

1995

PRIVATE

Survey linear regression

pweight: pesoper
Strata: upm
PSU: usm

Number of obs = 1284
Number of strata = 1
Number of PSUs = 189
Population size = 906845.06
F(8, 181) = 73.80
Prob > F = 0.0000
R-squared = 0.4202

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.1375406	.0089421	15.38	0.000	.1199008	.1551803
EO	.0803507	.0137471	5.84	0.000	.0532323	.1074692
EU	-.130407	.0166553	-7.83	0.000	-.1632623	-.0975516
female	-.2632976	.0456251	-5.77	0.000	-.3533004	-.1732947
voc. edu.	.1366684	.03679	3.71	0.000	.0640941	.2092426
expot	.0509432	.0054173	9.40	0.000	.0402566	.0616297
expot2	-.0004998	.0001244	-4.02	0.000	-.0007451	-.0002545
hours	.0094248	.0012957	7.27	0.000	.0068688	.0119807
_cons	3.803945	.1417738	26.83	0.000	3.524273	4.083617

PUBLIC

Survey linear regression

pweight: pesoper
Strata: upm
PSU: usm

Number of obs = 309
Number of strata = 1
Number of PSUs = 128
Population size = 221673.5
F(8, 120) = 9.78
Prob > F = 0.0000
R-squared = 0.2944

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0930027	.0176906	5.26	0.000	.0579961	.1280093
EO	.0680193	.035247	1.93	0.056	-.0017281	.1377667
EU	-.135426	.0230461	-5.88	0.000	-.1810302	-.0898219
female	-.362166	.0750879	-4.82	0.000	-.5107514	-.2135806
voc. edu.	-.0093973	.076356	-0.12	0.902	-.1604922	.1416975
expot	.042559	.011388	3.74	0.000	.0200242	.0650938
expot2	-.0005001	.0002551	-1.96	0.052	-.0010049	4.61e-06
hours	.0056895	.0023124	2.46	0.015	.0011138	.0102653
_cons	4.542171	.3021167	15.03	0.000	3.944337	5.140005

1996

PRIVATE

Source	SS	df	MS

Number of obs = 1007
F(9, 997) = 59.86

```
      Model |  233.425586     9  25.9361762           Prob > F      =  0.0000
   Residual |  432.010229   997  .433310159           R-squared     =  0.3508
------------+------------------------------           Adj R-squared =  0.3449
      Total |  665.435814  1006  .661467012           Root MSE      =  .65826

------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
------------+-----------------------------------------------------------------
          ER |   .0914546   .0082618    11.07   0.000      .075242    .1076672
          EO |   .0260041   .0127932     2.03   0.042     .0008994    .0511088
          EU |  -.1165968   .0149011    -7.82   0.000    -.1458378   -.0873557
      female |  -.2840398   .0459818    -6.18   0.000    -.3742721   -.1938076
   voc. edu. |   .1551224   .0429606     3.61   0.000     .0708189     .239426
        expot |    .029661   .0050594     5.86   0.000     .0197328    .0395893
       expot2 |  -.0003573   .0001122    -3.18   0.002    -.0005775    -.000137
        hours |   .0121229   .0011875    10.21   0.000     .0097925    .0144532
   public_sch |  -.2251471   .0525261    -4.29   0.000    -.3282216   -.1220726
        _cons |   4.744897   .1435091    33.06   0.000     4.463282    5.026511
------------------------------------------------------------------------------
```

PUBLIC

```
      Source |       SS       df       MS              Number of obs =     231
-------------+------------------------------           F(  9,   221) =    6.97
       Model |  24.7561413     9  2.75068236           Prob > F      =  0.0000
    Residual |  87.2037134   221  .394586938           R-squared     =  0.2211
-------------+------------------------------           Adj R-squared =  0.1894
       Total |  111.959855   230  .486781977           Root MSE      =  .62816

------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ER |   .0720553   .0212767     3.39   0.001     .0301241    .1139865
          EO |   .0724501   .0408772     1.77   0.078      -.008109    .1530091
          EU |  -.0846184   .0324199    -2.61   0.010    -.1485101   -.0207266
      female |  -.2257628   .0887125    -2.54   0.012    -.4005934   -.0509321
   voc. edu. |   .1682389   .0890017     1.89   0.060    -.0071617    .3436394
        expot |   .0292357   .0136214     2.15   0.033     .0023912    .0560802
       expot2 |  -.0003635   .0003203    -1.14   0.258    -.0009947    .0002676
        hours |   .0119156   .0027087     4.40   0.000     .0065775    .0172537
   public_sch |  -.1005601   .0907488    -1.11   0.269    -.2794039    .0782837
        _cons |   4.666372   .3786462    12.32   0.000     3.920153    5.412592
------------------------------------------------------------------------------
```

1997

PRIVATE

```
Survey linear regression

pweight:  pesoper                                 Number of obs    =       603
Strata:   <one>                                   Number of strata =         1
PSU:      <observations>                          Number of PSUs   =       603
                                                  Population size  =  638657.1
                                                  F(  9,    594)   =     34.93
                                                  Prob > F         =    0.0000
                                                  R-squared        =    0.3598

------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ER |   .1130473   .0107104    10.55   0.000      .092013    .1340817
```

EO	.0666726	.0119851	5.56	0.000	.0431349	.0902103
EU	-.0748891	.0203548	-3.68	0.000	-.1148642	-.0349141
female	-.3813357	.0563954	-6.76	0.000	-.4920914	-.2705801
voc. edu.	.0793109	.05208	1.52	0.128	-.0229696	.1815913
expot	.0350866	.0057688	6.08	0.000	.0237571	.0464161
expot2	-.0004474	.0001252	-3.57	0.000	-.0006933	-.0002014
hours	.0060467	.0015387	3.93	0.000	.0030247	.0090686
public_sch	-.0557927	.0791446	-0.70	0.481	-.2112257	.0996403
_cons	4.801645	.1637166	29.33	0.000	4.48012	5.12317

PUBLIC

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 640
Number of strata	= 1
Number of PSUs	= 640
Population size	= 604941.51
F(9, 631)	= 27.79
Prob > F	= 0.0000
R-squared	= 0.3460

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.1028507	.0153734	6.69	0.000	.0726621	.1330393
EO	.0655195	.0232727	2.82	0.005	.0198193	.1112198
EU	-.0994768	.023092	-4.31	0.000	-.1448222	-.0541314
female	-.3662203	.0602353	-6.08	0.000	-.4845034	-.2479372
voc. edu.	.3829756	.0681978	5.62	0.000	.2490567	.5168944
expot	.038414	.0078916	4.87	0.000	.0229175	.0539106
expot2	-.0005107	.0001877	-2.72	0.007	-.0008792	-.0001421
hours	.0106798	.0018804	5.68	0.000	.0069873	.0143723
public_sch	-.3589985	.0592406	-6.06	0.000	-.4753283	-.2426687
_cons	4.749353	.2520339	18.84	0.000	4.254438	5.244268

1998

PRIVATE

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 651
Number of strata	= 1
Number of PSUs	= 651
Population size	= 1187602.4
F(9, 642)	= 45.06
Prob > F	= 0.0000
R-squared	= 0.4523

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.1416646	.0116525	12.16	0.000	.1187836	.1645456
EO	.0459466	.0162058	2.84	0.005	.0141246	.0777686
EU	-.1073349	.0192245	-5.58	0.000	-.1450845	-.0695852
female	-.4138563	.0620789	-6.67	0.000	-.5357556	-.291957

```
   voc. edu. |   .0357223   .0605731     0.59   0.556    -.0832204    .1546649
       expot |   .0389232   .0077627     5.01   0.000     .0236802    .0541661
      expot2 |  -.0005803   .0001746    -3.32   0.001    -.0009231   -.0002374
       hours |   .0084897   .0018733     4.53   0.000     .0048112    .0121681
  public_sch |  -.2701568   .0769713    -3.51   0.000    -.4212993   -.1190143
       _cons |   4.838851   .2120648    22.82   0.000     4.422436    5.255266
-------------------------------------------------------------------------------
```

PUBLIC

Survey linear regression

```
pweight:  pesoper                         Number of obs    =       112
Strata:   <one>                           Number of strata =         1
PSU:      <observations>                  Number of PSUs   =       112
                                          Population size  = 212117.45
                                          F(   9,    103)  =      3.14
                                          Prob > F         =    0.0022
                                          R-squared        =    0.2409
```

```
-------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+-----------------------------------------------------------------
          ER |   .0722349   .0262339     2.75   0.007     .0202506    .1242191
          EO |   .0544698   .0484214     1.12   0.263    -.0414804    .1504201
          EU |   -.064539   .0241889    -2.67   0.009    -.1124709   -.0166071
      female |  -.1530054   .0920338    -1.66   0.099    -.3353765    .0293656
   voc. edu. |   .0092139   .1133095     0.08   0.935    -.2153164    .2337441
       expot |   .0217152    .016337     1.33   0.187    -.0106578    .0540881
      expot2 |  -.0002789   .0004048    -0.69   0.492    -.0010809    .0005232
       hours |    .011386   .0037829     3.01   0.003       .00389    .0188821
  public_sch |  -.0247545   .1046978    -0.24   0.814    -.2322202    .1827112
       _cons |   5.100147   .4522946    11.28   0.000     4.203895    5.996399
-------------------------------------------------------------------------------
```

1999

PRIVATE

Survey linear regression

```
pweight:  pesoper                         Number of obs    =       787
Strata:   <one>                           Number of strata =         1
PSU:      <observations>                  Number of PSUs   =       787
                                          Population size  = 1245129.2
                                          F(   9,    778)  =     78.11
                                          Prob > F         =    0.0000
                                          R-squared        =    0.5040
```

```
-------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+-----------------------------------------------------------------
          ER |   .1394816   .0105298    13.25   0.000     .1188117    .1601514
          EO |    .012924   .0158818     0.81   0.416    -.0182519    .0440998
          EU |  -.1400506   .0205497    -6.82   0.000    -.1803893   -.0997119
      female |   -.337154   .0554861    -6.08   0.000    -.4460724   -.2282355
   voc. edu. |     .04621   .0554101     0.83   0.405    -.0625593    .1549792
       expot |    .038302   .0072506     5.28   0.000     .0240691    .0525349
      expot2 |  -.0004581   .0001674    -2.74   0.006    -.0007867   -.0001295
```

```
       hours |   .0119166   .0016309     7.31   0.000     .0087152    .0151181
  public_sch |  -.2136175   .0644025    -3.32   0.001    -.3400388   -.0871962
       _cons |    4.66273   .1710574    27.26   0.000     4.326946    4.998513
------------------------------------------------------------------------------
```

PUBLIC

Survey linear regression

```
pweight:  pesoper                                 Number of obs    =       131
Strata:   <one>                                   Number of strata =         1
PSU:      <observations>                          Number of PSUs   =       131
                                                  Population size  = 212489.67
                                                  F(   9,    122)  =      7.06
                                                  Prob > F         =    0.0000
                                                  R-squared        =    0.3299
```

```
------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ER |   .1368293   .0272082     5.03   0.000     .0830011    .1906576
          EO |     .19519   .0658929     2.96   0.004     .0648287    .3255513
          EU |  -.0551906   .0372349    -1.48   0.141    -.1288554    .0184742
      female |   -.210825   .0947359    -2.23   0.028    -.3982487   -.0234014
   voc. edu. |   .1069387   .0851273     1.26   0.211    -.0614755    .2753528
       expot |   .0295013   .0125703     2.35   0.020     .0046326    .0543701
      expot2 |  -.0003999   .0002681    -1.49   0.138    -.0009302    .0001305
       hours |   .0097904   .0027288     3.59   0.000     .0043918     .015189
  public_sch |  -.3196471   .1029442    -3.11   0.002    -.5233099   -.1159842
       _cons |   4.358135   .4710933     9.25   0.000     3.426133    5.290136
------------------------------------------------------------------------------
```

SEGMENTED MARKET ANALYSIS: FORMAL AND INFORMAL SECTORS

1995

FORMAL MARKET

Survey linear regression

```
pweight:  pesoper                                 Number of obs    =      1007
Strata:   upm                                     Number of strata =         1
PSU:      usm                                     Number of PSUs   =       187
                                                  Population size  =  704651.9
                                                  F(   9,    178)  =     31.41
                                                  Prob > F         =    0.0000
                                                  R-squared        =    0.3084
```

```
------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ER |    .111699   .0098534    11.34   0.000     .0922602    .1311378
          EO |   .0758899    .017094     4.44   0.000     .0421669    .1096129
          EU |  -.1149703   .0177296    -6.48   0.000    -.1499472   -.0799934
      female |   -.283098   .0520686    -5.44   0.000    -.3858189    -.180377
       size2 |   .1604366   .0755925     2.12   0.035     .0113078    .3095654
   voc. edu. |   .0564267    .046892     1.20   0.230    -.0360819    .1489353
       expot |   .0349354   .0063908     5.47   0.000     .0223276    .0475431
```

expot2	-.0002617	.0001464	-1.79	0.076	-.0005506	.0000271
hours	.0095562	.0015228	6.28	0.000	.0065521	.0125604
_cons	4.185228	.1738645	24.07	0.000	3.842228	4.528228

INFORMAL MARKET

Survey linear regression

pweight: pesoper
Strata: upm
PSU: usm

Number of obs	= 757
Number of strata	= 1
Number of PSUs	= 185
Population size	= 533449.21
F(9, 176)	= 45.61
Prob > F	= 0.0000
R-squared	= 0.3751

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0834279	.0099471	8.39	0.000	.0638029	.1030528
EO	.0616199	.0137705	4.47	0.000	.0344516	.0887882
EU	-.068622	.0172868	-3.97	0.000	-.102728	-.0345161
female	-.4504336	.0536485	-8.40	0.000	-.5562789	-.3445883
size2	.3151029	.0610937	5.16	0.000	.1945687	.4356371
voc. edu.	.1966994	.0472447	4.16	0.000	.1034885	.2899104
expot	.0351618	.0060835	5.78	0.000	.0231594	.0471642
expot2	-.0003973	.0001361	-2.92	0.004	-.0006657	-.0001288
hours	.0069085	.0012335	5.60	0.000	.0044748	.0093421
_cons	4.280844	.1458105	29.36	0.000	3.993169	4.56852

1996

FORMAL

Source	SS	df	MS
Model	140.707948	10	14.0707948
Residual	354.546641	816	.434493433
Total	495.254589	826	.599581827

Number of obs	= 827
F(10, 816)	= 32.38
Prob > F	= 0.0000
R-squared	= 0.2841
Adj R-squared	= 0.2753
Root MSE	= .65916

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0637833	.0098483	6.48	0.000	.0444524	.0831142
EO	.0148049	.0164006	0.90	0.367	-.0173874	.0469972
EU	-.1092968	.0161781	-6.76	0.000	-.1410524	-.0775413
female	-.3084635	.0501749	-6.15	0.000	-.4069506	-.2099764
size2	.0978901	.0471452	2.08	0.038	.00535	.1904303
voc. edu.	.1758872	.0470836	3.74	0.000	.083468	.2683064
expot	.0284165	.0064318	4.42	0.000	.0157917	.0410413
expot2	-.0003107	.000147	-2.11	0.035	-.0005992	-.0000221
hours	.0090225	.0014446	6.25	0.000	.0061869	.0118582
public_sch	-.3037847	.0512332	-5.93	0.000	-.404349	-.2032204
_cons	5.218527	.1828329	28.54	0.000	4.859649	5.577405

INFORMAL

Source	SS	df	MS	Number of obs =	406
Model	59.7394279	10	5.97394279	F(10, 395) =	16.50
Residual	142.995105	395	.362012925	Prob > F =	0.0000
				R-squared =	0.2947
Total	202.734533	405	.500579094	Adj R-squared =	0.2768
				Root MSE =	.60168

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0462625	.0139731	3.31	0.001	.0187915	.0737335
EO	.0230181	.017717	1.30	0.195	-.0118132	.0578495
EU	-.0175249	.0248069	-0.71	0.480	-.0662948	.0312451
female	-.2552295	.066549	-3.84	0.000	-.3860641	-.1243949
size2	.0373177	.0693045	0.54	0.591	-.0989341	.1735694
voc. edu.	.1927664	.063713	3.03	0.003	.0675075	.3180254
expot	.0108872	.0068525	1.59	0.113	-.0025848	.0243592
expot2	-.0002163	.0001424	-1.52	0.129	-.0004962	.0000636
hours	.017333	.0015639	11.08	0.000	.0142584	.0204077
public_sch	-.1092435	.0856735	-1.28	0.203	-.2776766	.0591896
_cons	4.768715	.1990705	23.95	0.000	4.377345	5.160085

1997

FORMAL

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs =	656
Number of strata =	1
Number of PSUs =	656
Population size =	639463.16
F(10, 646) =	25.67
Prob > F =	0.0000
R-squared =	0.3082

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0752393	.0106234	7.08	0.000	.0543793	.0960993
EO	.0494817	.0175716	2.82	0.005	.0149782	.0839853
EU	-.0518486	.0204472	-2.54	0.011	-.0919985	-.0116987
female	-.3268977	.0569021	-5.74	0.000	-.4386302	-.2151653
size2	.1073471	.054197	1.98	0.048	.0009262	.213768
voc. edu.	.2375355	.0550181	4.32	0.000	.1295023	.3455686
expot	.0400679	.0062064	6.46	0.000	.027881	.0522548
expot2	-.0004369	.0001264	-3.46	0.001	-.0006851	-.0001887
hours	.0059702	.0016644	3.59	0.000	.0027019	.0092384
public_sch	-.303547	.0581674	-5.22	0.000	-.417764	-.18933
_cons	5.30941	.1836058	28.92	0.000	4.948883	5.669937

INFORMAL

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs =	485
Number of strata =	1
Number of PSUs =	485
Population size =	518561.92
F(10, 475) =	14.67

```
                                                Prob > F        =     0.0000
                                                R-squared       =     0.2554
------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ER |   .0564107   .0129451     4.36   0.000     .0309751    .0818463
          EO |   .0384389   .0141202     2.72   0.007     .0106944    .0661834
          EU |  -.0935992   .0278136    -3.37   0.001    -.1482496   -.0389488
      female |  -.3914989    .062272    -6.29   0.000    -.5138557   -.2691421
       size2 |   -.000281   .0574735    -0.00   0.996    -.1132094    .1126473
    oc. edu. |   .1411177   .0611183     2.31   0.021     .0210277    .2612078
       expot |   .0308835    .006319     4.89   0.000     .0184675    .0432995
      expot2 |  -.0006032   .0001511    -3.99   0.000    -.0009001   -.0003062
       hours |   .0090206   .0016596     5.44   0.000     .0057598    .0122814
  public_sch |   -.236892   .0813621    -2.91   0.004    -.3967584   -.0770255
       _cons |   5.270839   .1852118    28.46   0.000      4.90692    5.634757
------------------------------------------------------------------------------
```

1998

FORMAL

Survey linear regression

```
pweight:  pesoper                               Number of obs    =       437
Strata:   <one>                                 Number of strata =         1
PSU:      <observations>                        Number of PSUs   =       437
                                                Population size  = 764726.65
                                                F(  10,    427)  =     15.04
                                                Prob > F         =    0.0000
                                                R-squared        =    0.2597
------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ER |   .0975119   .0148762     6.55   0.000     .0682739    .1267499
          EO |   .0335121   .0248916     1.35   0.179    -.0154104    .0824346
          EU |  -.0659374   .0226991    -2.90   0.004    -.1105506   -.0213241
      female |  -.2075921   .0719683    -2.88   0.004    -.3490399   -.0661442
       size2 |   .1287326   .0925101     1.39   0.165    -.0530886    .3105538
    oc. edu. |   .0157053   .0715744     0.22   0.826    -.1249685    .1563791
       expot |     .02939   .0092569     3.17   0.002     .0111964    .0475837
      expot2 |  -.0004354   .0002195    -1.98   0.048    -.0008668   -3.95e-06
       hours |   .0088863    .002334     3.81   0.000     .0042991    .0134736
  public_sch |  -.2889105   .0791217    -3.65   0.000    -.4444177   -.1334032
       _cons |   5.251269   .2822605    18.60   0.000     4.696508    5.806029
------------------------------------------------------------------------------
```

INFORMAL

Survey linear regression

```
pweight:  pesoper                               Number of obs    =       260
Strata:   <one>                                 Number of strata =         1
PSU:      <observations>                        Number of PSUs   =       260
                                                Population size  = 515730.05
                                                F(  10,    250)  =      9.30
                                                Prob > F         =    0.0000
                                                R-squared        =    0.3231
```

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0979591	.0186941	5.24	0.000	.0611473	.1347709
EO	.0379498	.019754	1.92	0.056	-.0009491	.0768488
EU	-.0896218	.0327527	-2.74	0.007	-.1541172	-.0251264
female	-.6021047	.0923797	-6.52	0.000	-.7840157	-.4201937
size2	.0684991	.0877221	0.78	0.436	-.1042403	.2412384
oc. edu.	.0535299	.0881236	0.61	0.544	-.1200002	.2270599
expot	.0288585	.0114508	2.52	0.012	.0063099	.0514071
expot2	-.0004374	.0002654	-1.65	0.101	-.00096	.0000852
hours	.0083615	.0029261	2.86	0.005	.0025994	.0141235
public_sch	-.0999417	.1214638	-0.82	0.411	-.3391241	.1392406
_cons	5.022014	.3024041	16.61	0.000	4.42653	5.617498

1999

FORMAL

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 486
Number of strata	= 1
Number of PSUs	= 486
Population size	= 778095.55
F(10, 476)	= 23.13
Prob > F	= 0.0000
R-squared	= 0.3111

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.1007314	.0162028	6.22	0.000	.0688951	.1325677
EO	.0603373	.0281618	2.14	0.033	.005003	.1156715
EU	-.0757866	.0314901	-2.41	0.016	-.1376604	-.0139129
female	-.3626976	.0667388	-5.43	0.000	-.4938305	-.2315648
c. edu.	-.0049442	.0707918	-0.07	0.944	-.1440407	.1341523
size2	.286506	.0884912	3.24	0.001	.1126325	.4603794
expot	.0308796	.0113795	2.71	0.007	.0085203	.0532389
expot2	-.0003188	.0002695	-1.18	0.237	-.0008482	.0002107
hours	.0116632	.0021227	5.49	0.000	.0074924	.015834
public_sch	-.392269	.0713525	-5.50	0.000	-.5324673	-.2520708
_cons	5.064408	.2860247	17.71	0.000	4.502407	5.626408

INFORMAL

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 357
Number of strata	= 1
Number of PSUs	= 357
Population size	= 567870.06
F(10, 347)	= 14.63
Prob > F	= 0.0000
R-squared	= 0.3838

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0783212	.0140333	5.58	0.000	.0507226	.1059198
EO	-.0118193	.0185809	-0.64	0.525	-.0483615	.0247229
EU	-.0923744	.0242821	-3.80	0.000	-.1401289	-.04462
female	-.2935371	.0715244	-4.10	0.000	-.4342006	-.1528737
voc. edu.	-.0263199	.0719781	-0.37	0.715	-.1678757	.1152359
size2	.1775291	.0673892	2.63	0.009	.0449982	.31006
expot	.0234596	.0091842	2.55	0.011	.0053975	.0415218
expot2	-.0003476	.0002086	-1.67	0.097	-.0007578	.0000626
hours	.0160886	.0018267	8.81	0.000	.0124961	.0196811
public_sch	.1021676	.0847194	1.21	0.229	-.0644458	.2687809
_cons	4.665641	.2031369	22.97	0.000	4.266141	5.06514

SEGMENTED MARKET: TRADABLE AND NON TRADABLE SECTORS

1995

TRADABLE

Survey linear regression

pweight: pesoper
Strata: upm
PSU: usm

Number of obs	= 704
Number of strata	= 1
Number of PSUs	= 175
Population size	= 507589.93
F(9, 166)	= 54.62
Prob > F	= 0.0000
R-squared	= 0.4645

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.1318346	.0102104	12.91	0.000	.1116825	.1519867
EO	.0604639	.013629	4.44	0.000	.0335644	.0873633
EU	-.0744134	.018541	-4.01	0.000	-.1110076	-.0378191
female	-.3068631	.0562782	-5.45	0.000	-.4179389	-.1957873
size2	.3206266	.0503735	6.36	0.000	.2212048	.4200484
voc. edu.	.1249579	.0536374	2.33	0.021	.0190942	.2308217
expot	.0402291	.0058461	6.88	0.000	.0286907	.0517675
expot2	-.0004649	.0001186	-3.92	0.000	-.000699	-.0002307
hours	.0093995	.0014691	6.40	0.000	.0064999	.0122991
_cons	3.878943	.1458713	26.59	0.000	3.591039	4.166848

NON TRADABLE

Survey linear regression

pweight: pesoper
Strata: upm
PSU: usm

Number of obs	= 2059
Number of strata	= 1
Number of PSUs	= 193
Population size	= 1445598.3
F(9, 184)	= 133.39
Prob > F	= 0.0000

```
                                                R-squared         =     0.4733
------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ER |   .1299953   .0075722    17.17   0.000     .1150599    .1449306
          EO |   .0778155   .0091831     8.47   0.000     .0597028    .0959283
          EU |  -.0889453   .0153679    -5.79   0.000    -.1192569   -.0586337
      female |  -.4030609   .0386046   -10.44   0.000    -.4792046   -.3269172
       size2 |   .2829462   .0429885     6.58   0.000     .1981557    .3677367
   voc. edu. |   .0511412   .0338148     1.51   0.132     -.015555    .1178374
       expot |   .0453519   .0039817    11.39   0.000     .0374985    .0532053
      expot2 |  -.0005358   .0000823    -6.51   0.000    -.0006981   -.0003736
       hours |   .0102921   .0009115    11.29   0.000     .0084943    .0120898
       _cons |   3.717972   .1087485    34.19   0.000     3.503477    3.932467
------------------------------------------------------------------------------
```

1996

TRADABLE

```
      Source |       SS       df       MS              Number of obs =     544
-------------+------------------------------           F( 10,   533) =   33.79
       Model |  160.70647     10  16.070647            Prob > F      =  0.0000
    Residual |  253.488594    533  .475588357           R-squared     =  0.3880
-------------+------------------------------           Adj R-squared =  0.3765
       Total |  414.195064    543  .762790174           Root MSE      =  .68963

------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ER |   .1037865   .0120414     8.62   0.000      .080132    .1274409
          EO |   .0300875   .0152339     1.98   0.049     .0001617    .0600133
          EU |  -.0675348    .021538    -3.14   0.002    -.1098446    -.025225
      female |  -.3235366   .0677313    -4.78   0.000    -.4565896   -.1904835
       size2 |  -.0681802   .0649886    -1.05   0.295    -.1958455     .059485
   voc. edu. |   .1443914   .0649178     2.22   0.027     .0168653    .2719175
       expot |   .0341726   .0067615     5.05   0.000     .0208901    .0474551
      expot2 |  -.0005204   .0001309    -3.98   0.000    -.0007775   -.0002632
       hours |   .0147621   .0016178     9.12   0.000     .0115841    .0179401
  public_sch |  -.2780232    .078366    -3.55   0.000    -.4319674    -.124079
       _cons |   4.464388   .1948523    22.91   0.000     4.081615     4.84716
------------------------------------------------------------------------------
```

NON TRADABLE

```
      Source |       SS       df       MS              Number of obs =    1577
-------------+------------------------------           F( 10,  1566) =  109.33
       Model |  500.146928     10  50.0146928           Prob > F      =  0.0000
    Residual |  716.403167   1566  .457473287           R-squared     =  0.4111
-------------+------------------------------           Adj R-squared =  0.4074
       Total |   1216.5501   1576  .771922649           Root MSE      =  .67637

------------------------------------------------------------------------------
   logingtot |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ER |   .0811371   .0060514    13.41   0.000     .0692675    .0930067
          EO |  -.0025109   .0096633    -0.26   0.795    -.0214653    .0164435
          EU |  -.0858205   .0120022    -7.15   0.000    -.1093626   -.0622783
      female |  -.2391766   .0353811    -6.76   0.000    -.3085759   -.1697774
```

size2	.0196768	.0378843	0.52	0.604	-.0546324	.0939861
voc. edu.	.0964793	.0347792	2.77	0.006	.0282606	.1646979
expot	.0212266	.0039245	5.41	0.000	.0135287	.0289245
expot2	-.0003002	.0000809	-3.71	0.000	-.0004589	-.0001415
hours	.0194746	.0008402	23.18	0.000	.0178266	.0211226
public_sch	-.2000148	.0424145	-4.72	0.000	-.28321	-.1168197
_cons	4.611724	.108479	42.51	0.000	4.398945	4.824504

1997
TRADABLE

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 434
Number of strata	= 1
Number of PSUs	= 434
Population size	= 452794.4
F(10, 424)	= 28.05
Prob > F	= 0.0000
R-squared	= 0.4597

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.1471341	.0182879	8.05	0.000	.1111899	.1830783
EO	.0616579	.0199988	3.08	0.002	.0223511	.1009647
EU	-.0992109	.0333006	-2.98	0.003	-.1646617	-.03376
female	-.6084932	.0819181	-7.43	0.000	-.7694997	-.4474867
size2	.2166432	.071696	3.02	0.003	.0757278	.3575585
voc. edu.	.0296684	.0742502	0.40	0.690	-.1162671	.175604
expot	.0345205	.0077872	4.43	0.000	.0192152	.0498259
expot2	-.000597	.0001636	-3.65	0.000	-.0009186	-.0002755
hours	.0103758	.0021156	4.90	0.000	.0062177	.0145339
public_sch	-.2256916	.1023707	-2.20	0.028	-.426897	-.0244863
_cons	4.425727	.2541831	17.41	0.000	3.926141	4.925313

NON TRADABLE

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 1594
Number of strata	= 1
Number of PSUs	= 1594
Population size	= 1566982.1
F(10, 1584)	= 83.20
Prob > F	= 0.0000
R-squared	= 0.4021

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0934238	.0068619	13.61	0.000	.0799645	.106883
EO	.0040027	.0113177	0.35	0.724	-.0181964	.0262017
EU	-.0896399	.0171366	-5.23	0.000	-.1232525	-.0560272
female	-.3594806	.0401237	-8.96	0.000	-.4381814	-.2807798
size2	.0827912	.0434657	1.90	0.057	-.0024648	.1680472
voc. edu.	.1322036	.0430939	3.07	0.002	.047677	.2167303
expot	.0235325	.0041161	5.72	0.000	.015459	.031606

expot2	-.0003822	.0000779	-4.91	0.000	-.0005351	-.0002294
hours	.0161752	.0010937	14.79	0.000	.0140299	.0183205
public_sch	-.1277358	.0526807	-2.42	0.015	-.2310666	-.0244051
_cons	4.714966	.1227803	38.40	0.000	4.474138	4.955794

1998

TRADABLE

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 217
Number of strata	= 1
Number of PSUs	= 217
Population size	= 428024.3
F(10, 207)	= 15.09
Prob > F	= 0.0000
R-squared	= 0.4753

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.1297124	.0206011	6.30	0.000	.0891074	.1703173
EO	.0402318	.0248137	1.62	0.106	-.0086761	.0891398
EU	-.0711221	.0368485	-1.93	0.055	-.1437508	.0015066
female	-.3752227	.0909738	-4.12	0.000	-.5545328	-.1959126
size2	.3530706	.0916492	3.85	0.000	.1724293	.5337119
voc. edu.	.1054729	.0868504	1.21	0.226	-.0657099	.2766557
expot	.0177731	.0093121	1.91	0.058	-.0005811	.0361274
expot2	-.0001323	.0001797	-0.74	0.462	-.0004864	.0002218
hours	.0093335	.002715	3.44	0.001	.0039822	.0146848
public_sch	-.0459369	.1294297	-0.35	0.723	-.3010439	.2091701
_cons	4.540418	.3813236	11.91	0.000	3.788826	5.292009

NON TRADABLE

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 902
Number of strata	= 1
Number of PSUs	= 902
Population size	= 1616235.7
F(10, 892)	= 44.76
Prob > F	= 0.0000
R-squared	= 0.3976

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0969062	.0106598	9.09	0.000	.0759852	.1178271
EO	.0363374	.0139586	2.60	0.009	.0089421	.0637326
EU	-.0686998	.0179568	-3.83	0.000	-.1039418	-.0334577
female	-.2864707	.053727	-5.33	0.000	-.3919154	-.181026
size2	.3182243	.0661659	4.81	0.000	.1883671	.4480816
voc. edu.	.009909	.0584117	0.17	0.865	-.1047299	.1245479
expot	.0294148	.005889	4.99	0.000	.0178571	.0409725
expot2	-.0004696	.0001161	-4.05	0.000	-.0006974	-.0002418
hours	.0121436	.0015274	7.95	0.000	.0091458	.0151413

public_sch	-.2605498	.06986	-3.73	0.000	-.3976571	-.1234425
_cons	4.897521	.1906361	25.69	0.000	4.523379	5.271664

1999

TRADABLE

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 261
Number of strata	= 1
Number of PSUs	= 261
Population size	= 432953.92
F(10, 251)	= 20.23
Prob > F	= 0.0000
R-squared	= 0.5464

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.119222	.0214976	5.55	0.000	.0768905	.1615536
EO	.0117726	.0252089	0.47	0.641	-.0378671	.0614122
EU	-.1075186	.0363814	-2.96	0.003	-.1791583	-.0358789
female	-.3687289	.0911983	-4.04	0.000	-.5483101	-.1891476
size2	.5208187	.1037241	5.02	0.000	.3165724	.725065
voc. edu.	.0276035	.1003343	0.28	0.783	-.1699677	.2251747
expot	.0541321	.0122694	4.41	0.000	.029972	.0782923
expot2	-.0007946	.0002489	-3.19	0.002	-.0012847	-.0003044
hours	.017299	.0030065	5.75	0.000	.0113788	.0232192
public_sch	-.0639244	.1275772	-0.50	0.617	-.3151404	.1872916
_cons	3.977637	.3257031	12.21	0.000	3.336285	4.618989

NON TRADABLE

Survey linear regression

pweight: pesoper
Strata: <one>
PSU: <observations>

Number of obs	= 1155
Number of strata	= 1
Number of PSUs	= 1155
Population size	= 1804848.2
F(10, 1145)	= 93.56
Prob > F	= 0.0000
R-squared	= 0.4886

logingtot	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ER	.0919713	.0090877	10.12	0.000	.0741411	.1098015
EO	.0024944	.0124589	0.20	0.841	-.0219503	.0269391
EU	-.1029358	.0175003	-5.88	0.000	-.1372718	-.0685998
female	-.2212194	.0429505	-5.15	0.000	-.3054893	-.1369496
size2	.4193717	.0552017	7.60	0.000	.3110647	.5276787
oc. edu.	-.0162091	.0493959	-0.33	0.743	-.1131249	.0807066
expot	.0269114	.0056412	4.77	0.000	.0158433	.0379796
expot2	-.0003362	.0001221	-2.75	0.006	-.0005758	-.0000966
hours	.0187853	.0011559	16.25	0.000	.0165175	.0210532
public_sch	-.2571392	.0555267	-4.63	0.000	-.3660838	-.1481947
_cons	4.713931	.1390192	33.91	0.000	4.441172	4.986689