

# Earnings Inequality after Mexico's Economic Reforms

**Gladys Lopez-Acevedo (LCSPE)**<sup>1,2</sup>

The World Bank  
Universidad Panamerica- C.I.D.E  
Insurgentes 1605 Piso 24  
Torre Mural  
Col. San Jose Insurgentes  
Mexico D.F 03900  
Gacevedo@worldbank.org

JEL Codes: J21, J30

March, 2000

---

<sup>1</sup> This research was completed as part of the “Earnings Inequality after Mexico’s Economic and Educational Reforms” study at the World Bank. I am grateful to INEGI and SEP (Ministry of Education) for providing us with the data. These are views of the author, and need not reflect those of the World Bank, its Executive Directors, or countries they represent.

<sup>2</sup> I thank Lauro Ramos, Angel Salinas, Paulo Vieira, Francois J. Bourguignon, Chris Scott, Quentin Wodon, William Maloney for helpful comments and discussions. Comments received at the 2000 Conference on "Consequences of Financial Crises on Income Distribution and Poverty in Latin America" and the participants of the BBL Seminar at the World Bank. I specially thank Angel Salinas for his help in completing this paper.

# Earnings Inequality after Mexico's Economic Reforms

Gladys Lopez-Acevedo (LCSPE)

The World Bank

Universidad Panamericana and C.I.D.E

## Abstract

The second half of the eighties and the present decade are an especially meaningful period for the Mexican economy, as it encompasses a major structural change from a protected, public-sector driven economy, to a globally integrated private-sector-led one. This change has resulted in sizable economic growth but, besides being increasingly unequal, Mexico's income distribution seems resilient to both growth and public policy. Most remarkably, the level, deterioration, and policy resilience of Mexico's inequality has over the past decade co-existed with very rapid progress in education attainment, both in terms of coverage and distribution of schooling. This phenomenon, which in recent years has also been observed in other developing as well as developed countries, is somewhat surprising, given the powerful equalizing properties generally attributed to education.

This paper finds that education is by far the variable that accounts for the largest share of the variation in earnings inequality in Mexico, both in terms of its gross and marginal contributions. This indicate that as the Mexican economy progresses, education becomes even more important in determining the choices of sectors and occupations. The contribution of education to inequality in Mexico is the second highest in Latin America, next to Brazil, and that the significance of education has been increasing over time.

The increase in earnings inequality, however, does not appear to be the result of a worsening in the distribution of education, whereas the income profile, which is related to the returns to schooling, has become much steeper. Meaning that there was a shift in demand towards high skilled labor that was not met by the increase in supply probably due to the increased rate of skill-biased technological change, whose transmission to Mexico may be facilitated by the economy's increased openness.

The structure of the rates of returns indicate that the "smooth" effects of both economic sector and labor market status variables in Mexico were significantly larger in the 1992-1997 period compared to the 1988-1992 period (before the trade agreement). In other words, the degree of correlation between education and economic sector, as well as labor market status, changed. In fact, the worker's insertion into the labor market and economic sector were a consequence of skill's differentials and not solely attributed to education. The returns to education have increased in recent times, especially for higher levels of education and in the upper tail of the conditional earnings distribution. Interestingly, returns to education for higher levels of education in Mexico are significantly above the Latin American average.

## INTRODUCTION

Achieving sustainable economic growth with a more egalitarian income distribution is at the core of Mexico's development challenge. Yet, the country does not perform well in terms of equity when compared with other Latin American countries. According to a recent study developed by the IDB (1998), Mexico has the sixth most unequal overall household income distribution (and the third worst in urban areas). In the broader international context, Mexico's ratio between the income share accruing to the 10 top percent to the bottom 40 percent of the population is higher than what is observed for the high-income countries and for the vast majority of the low-income countries (see table 3 in Annex 3).

The second half of the eighties and the present decade are an especially meaningful period for the Mexican economy, as it encompasses a major structural change from a protected, public-sector driven economy, to a globally integrated private-sector-led one. This change has resulted in sizable economic growth but, besides being increasingly unequal, Mexico's income distribution seems resilient to both growth and public policy.

Most remarkably, the level, deterioration, and policy resilience of Mexico's inequality has over the past decade co-existed with very rapid progress in education attainment, both in terms of coverage and distribution of schooling.<sup>3</sup> This phenomenon, which in recent years has also been observed in other developing as well as developed countries, is somewhat surprising, given the powerful equalizing properties generally attributed to education.

This paper reviews the factors and mechanisms, which have been driving inequality in Mexico. More specifically, the observed expansion in earnings inequality in the recent period is examined with emphasis on the roles of education<sup>4</sup>, aiming at establishing an analytical framework that permits the analysis of interaction between education and labor market, and examining the evolution of earnings inequality after the macroeconomic and educational policies followed in the 80's and 90's.

The paper is organized as follows: Section I describes the evolution of total current income inequality, between 1984-96 based on the National Household Income and Expenditures Survey (ENIGH), and using the household per capita income as the unit of analysis. Section II, focuses on the evolution of individual earnings inequality, using the information of the National Urban Employment Survey (ENEU). Section III, investigates how much of the earnings inequality can be explained by education, as well as other relevant variables, both in gross and in marginal terms.<sup>5</sup> Section IV analyzes the evolution of educational attainment. Section V relates the changes in the distribution of education to the changes in earnings inequality. Section VI examines the evolution and structure of the rates of returns to education by means of ordinary least squares and quantile regressions. The last section has the concluding remarks.

## I. THE EVOLUTION OF TOTAL INCOME INEQUALITY

The first evaluation of income inequality evolution in Mexico is based on the information available in the ENIGH<sup>6</sup>. The reason for doing so is that this survey captures total current income of the households, including non-monetary income, besides earnings and other sources of monetary income. The unit of analysis is the household, and the concept of income is household per capita income.<sup>7</sup>

---

<sup>3</sup> De la Torre, Rodolfo, (1997).

<sup>4</sup> Wages are directly related to individual characteristics and do not depend upon family structure. Besides, the acquaintance with their distribution brings one most of the way to understand the distribution of welfare in society.

<sup>5</sup> Educational attainment has not only a monetary impact but can also affect other outcomes, which are important for individual's well being, but that are not necessarily measured in monetary terms. This study, however, will not consider the non-monetary impacts of education. An interesting methodology for the estimation of these impacts can be found in Wolfe, Barbara and Samuel Zuvekas (1997).

<sup>6</sup> See annex 1 for a brief description of this survey.

<sup>7</sup> This means that total current income of the household divided by its number of household members. That is, we are considering the household as a unit characterized by a flow of income transfers and disregarding aspects related to equivalence scale.

The main results of this evaluation are shown in table 1. It indicates that a very sizable deterioration in the income distribution has taken place between 1984 and 1996. While the poorest 20% of the population lost almost one seventh of their income share (0.6 percentage points), the richest 10% increased theirs by something close to one seventh (5.2 percentage points). Moreover, this last group was the only one that gained over that period, as not only the poorest, but also those in the middle lost in relative terms.

Looking at the results of this comparison, one can say that the 1984-1996 period in Mexico was marked by a series of regressive income transfers from almost the entire population spectrum to the richest stratum. Accordingly, the most commonly used inequality index points to a worsening in income inequality over this span of time. The Gini coefficient, which is more sensitive to changes in the middle of the distribution, rises from 0.473 in 1984 to 0.519 in 1996. On the other hand, the Theil T index, which is extremely sensitive to changes in the upper and lower tails, goes up from 0.411 in 1984 to 0.524 in 1996.

Even though the worsening of the distribution is indisputable, there are, nevertheless, two points that must be stressed. The first one is that, according to the ENIGH survey, most of the worsening of the income distribution happened in the mid-eighties (1984-1989). The early nineties display little variation in earnings inequality except for a small trend towards deterioration. From 1989 to 1994, the income share accruing to the 20% poorest decreased slightly (it went down from 3.9% to 3.8%), whereas the richest 10% were the only ones that increased theirs (by one percentage point), and, therefore, those in the middle also experienced losses.

**Table 1. Lorenz Curves for Total Current Income<sup>1/</sup> (Accumulated Income Share - %)**

<b>Population Share</b>	<b>1984</b>	<b>1989</b>	<b>1992</b>	<b>1994</b>	<b>1996</b>
10	1.66	1.39	1.32	1.39	1.39
20	4.47	3.88	3.68	3.76	3.89
30	8.19	7.29	6.92	6.98	7.29
40	12.85	11.65	11.09	11.08	11.63
50	18.76	17.05	16.26	16.28	17.08
60	26.15	23.78	22.83	22.79	23.86
70	35.51	32.25	31.13	31.10	32.39
80	47.64	43.12	42.14	41.93	43.44
90	64.53	58.75	58.32	57.68	59.33
92	68.79	63.06	62.81	62.03	63.61
94	73.73	68.03	68.03	67.26	68.68
96	79.38	73.82	74.47	73.70	74.95
98	86.68	81.60	82.81	82.49	83.32
100	100.0	100.0	100.0	100.0	100.0
Bottom 20%	4.5	3.9	3.7	3.8	3.9
Middle 40%	21.7	19.9	19.2	19.0	20.0
Mhigh 30%	38.4	35.0	35.5	34.9	35.5
Top 10%	35.5	41.3	41.7	42.3	40.7
Gini	0.473	0.519	0.529	0.530	0.515
Theil T	0.411	0.566	0.550	0.558	0.524

Source: Own calculations based on ENIGH.

<sup>1/</sup> Based on household per capita income.

The second fact to be emphasized is very surprising and hard to explain: the observed improvement in the income distribution between 1994 and 1996, an interval of time that entails a severe financial crisis in the Mexican economy.<sup>8</sup> Usually one would expect inequality to go up

<sup>8</sup> In 1994, current account deficit was 30 billion dollars, about 7 percent of GDP. The main effects of the financial crisis were i) GDP and domestic demand fell 6.2 percent and 14 percent respectively each; ii) the unemployment rate rose

during recessive times, as it seems plausible to admit that the rich have more ways to protect their assets than the poor do. Especially when it comes to labor which is basically the only asset of the poor (the *labor-hoarding* hypothesis). The fact, however, is that the 10% richest experienced relative losses (their income share dropped 1.6 percentage points) and, accordingly, inequality went down: the Gini coefficient came down from 0.534 in 1994 to 0.519 in 1996, whereas the drop in the Theil T was from 0.558 to 0.524. In principle, it could be argued that the richest experienced severe capital losses due to the crisis, in such a way that their total current income was affected compared to the poor. This hypothesis, however, is not supported by the data shown in table 2, as monetary income other than wages and salaries, and financial income as well, increased their share in total income in that time interim, particularly so for the urban areas. Therefore, the fall in inequality remains somewhat puzzling.

**Table 2. Share of Total Income by Source (%)**

Source	1994			1996		
	Total	Urban	Rural	Total	Urban	Rural
<b>Monetary Current Income</b>						
Total Labor Earnings	47.12	49.01	32.07	44.51	46.08	33.75
Property (Business) Income	16.96	16.23	22.75	17.74	17.11	22.07
Property Income and Rents	1.10	1.13	0.87	1.35	1.47	0.51
Income from cooperative firms	0.22	0.24	0.12	0.06	0.03	0.32
Monetary Transfers	5.44	4.72	11.23	6.55	5.89	11.11
Other Current Income	0.64	0.67	0.36	0.69	0.66	0.91
<b>No Monetary Current Income</b>						
Self-Consumption	1.44	0.81	6.46	1.20	0.69	4.72
Non Monetary Payment	1.55	1.58	1.28	2.25	2.32	1.82
Gifts	5.04	4.73	7.57	6.07	5.86	7.55
Housing Imputed Rent	16.02	16.60	11.39	13.76	14.28	10.20
<b>Financial Income</b>	4.46	4.28	5.91	5.80	5.62	7.04
<b>Total Income</b>	100.00	100.00	100.00	100.00	100.00	100.00

Source: Own calculations based on ENIGH.

In table 3 the results for the Gini and Theil T are displayed for urban and rural areas using total current income. There, one can see, for both indices that inequality in rural areas was lower than in urban areas and remarkably stable until 1992. After a small decrease in 1994, it increased in 1996, contrary to the aggregate result discussed before. In light of these outcomes, it seems pertinent to state that the leading force behind the behavior of current income distribution in Mexico is in the urban areas.

**Table 3. Inequality Measures for Total Current Income**

Year	Gini Coefficient			Year	Theil T Index		
	National	Urban	Rural		National	Urban	Rural
1984	0.473	0.442	0.448	1984	0.411	0.356	0.375
1989	0.519	0.498	0.444	1989	0.566	0.526	0.361
1992	0.529	0.498	0.434	1992	0.550	0.483	0.353
1994	0.534	0.508	0.419	1994	0.558	0.499	0.325
1996	0.519	0.493	0.452	1996	0.524	0.470	0.390

Source: Own calculations based on ENIGH.

from 3.7 percent in 1994 to 6.2 percent in 1995; and, iii) the GDP per capita decreased 7.8 percent and workers experienced a significant reduction in their real wage, nearly 17 percent in 1995.

## II. THE EVOLUTION OF EARNINGS INEQUALITY

How much of total income inequality is due to earnings inequality? Table 4 presents the results of total current income inequality for each of its components: earnings<sup>9</sup>, monetary income excluding earnings and non-monetary income by urban and rural areas (see annex 2.1 and table 1 in annex 3).<sup>10</sup> Earnings are the source of income that contributes to most of the overall inequality, being responsible for almost half of it at national level. It is clear that these figures may be affected by a possible underreporting of capital gains, but it seems valid to state that if one better understands the mechanisms that lead earnings inequality, this will be a large step towards understanding the total inequality behavior. Besides, as long as labor is the main, if not the only, asset of the poor, a better knowledge of earnings inequality is a valuable input for the assessment of poverty and welfare issues.

**Table 4. Decomposition of Total Current Income (Percentage Share in Overall Gini)**

Income Source	Earnings	Monetary income Excluding earnings	No monetary Current income	TOTAL
<b>National</b>				
1984	46.0	32.9	21.0	100.0
1989	41.0	36.0	23.0	100.0
1992	42.9	31.9	25.2	100.0
1994	50.2	25.9	23.9	100.0
1996	46.7	29.4	23.9	100.0
<b>Urban</b>				
1984	45.6	32.2	22.2	100.0
1989	38.6	37.3	24.1	100.0
1992	41.4	33.1	25.5	100.0
1994	50.0	26.0	24.0	100.0
1996	46.1	29.8	24.1	100.0
<b>Rural</b>				
1984	30.7	49.5	19.8	100.0
1989	35.7	43.5	20.8	100.0
1992	29.6	42.2	28.2	100.0
1994	31.9	43.8	24.2	100.0
1996	35.7	41.2	23.1	100.0
1996	35.7	41.2	23.1	100.0

Source: Own calculations based on ENIGH.

To examine the behavior of inequality in recent times the household survey ENEU and ENIGH were used. However, in this paper the ENEU results are analyzed, while the ENIGH results are presented in the Annex 4. The ENEU survey is collected quarterly and for the purposes of this study the third quarter of each year was chosen, in an attempt to avoid the influence of seasonal factors that could make non-comparable results to those obtained from ENIGH.<sup>11</sup>

By examining the results shown on table 5, one also reaches the conclusion that the distribution of earnings has become more unequal in recent times. The Gini coefficient jumps from 0.395 in 1988 to 0.442 in 1997, after reaching a peak in 1996 of 0.464. Similarly, the Theil

<sup>9</sup>Earnings as defined in the ENIGH survey include salaries and wages, paid over-time, tips, contract workers' earnings, Christmas or New Year gifts and other gifts, and other monetary compensations (non-regular earnings). Earnings as defined in the ENEU survey include salaries and wages, self-employed workers' earnings, contract workers' earnings, and implicit firm owners' salaries, as well as non-monetary earnings.

<sup>10</sup> Though the results are shown for the Gini coefficient, these could have been obtained for the Theil T index, as both of them satisfy the six propositions listed in Shorrocks (1982).

<sup>11</sup> In order to reduce the heterogeneity of the sample and also aspects related to self selection, the population under analysis are individuals living in urban areas, between 16 and 65 years old, working 20 hours a week or more and no seasonal workers. Also, the two highest observations were dropped from the sample, as there was clear evidence of the presence of outliers in some years.

T index went up from 0.327 in 1988 to 0.372 in 1997, with 0.474 in 1996. Another index, the R10/20,<sup>12</sup> increased from 4.48 to 6.04 over the period, reaching a maximum of 6.74 in 1996.

**Table 5. Inequality Indices for the Distribution of Earnings (1988-1997)**

Population Share (%)	Earnings Share (%)									
	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997
Bottom 20%	7.54	7.62	7.19	6.84	6.47	6.13	5.98	5.91	5.72	5.95
Middle 40%	25.23	24.45	23.86	23.41	23.37	22.86	22.36	22.59	22.09	23.01
MHigh 30%	33.44	34.15	33.96	33.77	33.52	33.37	32.94	33.42	33.61	35.13
Top 10%	33.78	33.78	34.98	35.98	36.64	37.63	38.72	38.08	38.58	35.91
Gini	0.395	0.398	0.414	0.426	0.434	0.447	0.458	0.455	0.464	0.442
Theil T	0.327	0.328	0.350	0.380	0.396	0.414	0.470	0.427	0.474	0.372
R <sub>10/20</sub>	4.48	4.43	4.87	5.26	5.66	6.14	6.47	6.44	6.74	6.04

Source: Own calculations based on ENEU (3<sup>rd</sup> quarter).

There are two main differences in the pattern shown by the earnings and total current income distribution. First, the gains are not limited to the richest 10%. As those in the seventh, eight, and nine tenths of the distribution also improved their relative earnings over the period by almost two percentage points; the biggest losers were the middle 40%, who lost more than two percentage points of their income share. Second, there is a clear worsening in the earnings distribution in the present decade throughout 1996. On the other hand, the inequality associated with the total current income was moderately stable in the nineties, displaying an improvement in 1996.

The different behavior between total current income and earnings inequalities from 1994 to 1996 gives support to the idea that the poor, who mostly rely on labor as a source of income, were the least able to protect themselves during the recession. However, the substantial drop in earnings inequality from 1996 to 1997<sup>13</sup> is, once more, a surprising finding. It is true that the Mexican economy as a whole had a strong and impressive performance in 1997. The aggregate growth rate was around 7%, real investment grew by 24% and exports by 17%, industrial production increased 9.7%, and the civil construction sector, which is highly intensive in less skilled labor, experienced a growth close to 11%. Under such a scenario, an improvement in distribution of earnings itself is not unlikely, but the magnitude and quickness of the recovery calls for a detailed inspection of the mechanisms responsible for it.

Three broad hypotheses are frequently advanced to explain the similar increases in earnings inequality experienced in Mexico and other countries.<sup>14</sup> These link the increase of earnings inequality to (i) the increased openness of the economy, (ii) institutional changes in the labor market, and (iii) skill-biased technological change.

The first of these hypotheses argues that as trade barriers are reduced, an economy is placed under increased competitive pressures to specialize along its lines of comparative advantage. A developed country that is relatively high skilled-abundant, like the United States, will be induced to specialize in high skill- or educational-intensive activities as its low-skilled industries come under increased competitive pressure from low skilled-abundant, low-wage countries.

Hanson and Harrison (1995) examined the impact of the Mexican trade reform on the structure of wages using information at firm level. They tested whether trade reform had shifted employment toward industries that are relatively intensive in the use of skilled labor force [the Stolper-Samuelson-Type (SST) effect]. Their main conclusion was that the wage gap was associated to changes within industries and firms, which cannot be explained by the SST effect. Thus, the increase in wage inequality was due to other factors.<sup>15</sup> Hanson's (1997) paper examined a trade theory based on increasing returns, which has important implications for regional

<sup>12</sup> This index is the ratio of the income share accruing to the 10% richest and 20% to the poorest.

<sup>13</sup> The R<sub>10/20</sub> index, for instance, was 6.74 in 1996 and went down to 6.04 in 1997.

<sup>14</sup> See, for example, the "Symposium on Wage Inequality" (1997) and the "Symposium on How International Exchange, Technology and Institutions Affect Workers" (1997).

<sup>15</sup> The Stolper-Samuelson effect was also examined under NAFTA in Burfisher, Mary E., Robinson, Sherman, and Thierfelder (1993).

economies. Hanson's conclusion is that employment and wage patterns are consistent with the idea that access to market is important for industry location.

This first hypothesis has several problems when applied to the United States, and becomes even less persuasive when applied to Mexico. Mexico greatly liberalized its trade regime since 1984. However, the reduction of its trade barriers has mostly been vis-à-vis imports from the developed countries, notably the United States and Canada, whose share in total Mexican merchandise imports increased from 68 percent in 1985 to 73 percent in 1993 and to 77.5 percent in 1996. Since Mexico is a low skilled-abundant country compared to its two northern neighbors, it would be expected that the liberalization of trade would have induced a specialization pattern that would raise the relative demand (and hence wages) for the lesser-educated members of the labor force. This did not happen. Instead, the increase in earnings inequality observed in Mexico follows the same pattern as that observed in the United States: less educated workers experienced real wage declines, while highly educated workers experienced real wage improvements. The trade-based explanation may still be relevant, however, to the extent that greater openness facilitates the transfer of ideas and technology, which is identified below as the more persuasive explanation of increase in earnings inequality. A variant on the globalization/technology nexus explanation, advanced by Feenstra and Hanson (1994), involves outsourcing behavior where multinational enterprises in the developed country relocate their lower skilled-intensive activities to the less skilled-abundant developed countries. However, what is referred to as a low skilled activity in the United States may be a high-skilled activity in Mexico, which could explain the similar evolution of earnings inequality in both countries.

The second explanation revolves around institutional changes such as reductions in the minimum wage, the decreasing strength of trade unions and the declining share of state-owned enterprises. The existence of a binding minimum wage, for example, truncates the lower end of the wage distribution. As the minimum wage is allowed to erode away, say through inflation, it becomes less binding by moving further down the low end of the wage distribution, with the result that, *ceteris paribus*, a higher share of wages will lie below the previous minimum wage level. This translates into an increased dispersion in wages and earnings. A look at institutional developments in Mexico since the early 1980s indicates that these developments have not exerted a significant influence on the earnings distribution.<sup>16</sup> The distribution of real wages, for example, does not reveal any significant distortions around the minimum wage, which suggests that it is not a binding constraint. The fact that this minimum wage has continued to erode in real value, therefore, seems to have been irrelevant. Similarly, the distribution of union wages is not significantly different from that of non-union wages, once differences in education levels are accounted for. That also renders any erosion of union power irrelevant for the distribution of earnings. In conclusion, while the influence of institutional factors cannot be rejected entirely, it also does not appear to be the principal cause of the increase in earnings inequality.

A persuasive explanation, both for the United States and Mexico, seems to be one which links earnings inequality to skill-biased technological changes that raise the relative demand for higher-skilled labor. Cragg and Epelbaum (1996) examined the demand shift in Mexico. They pointed out that the major source of rising inequality is a biased demand shift rather than a uniform demand growth when there are different labor supply elasticities. Meza (1998) also investigated demand shifts. The author's hypothesis is that the demand shifts for more educated labor force, "within" economic sector, explains the increase in their premium when compared to the demand shift for less educated workers "between" economic sector. Tan and Batra (1997) studied the skill-biased technical change hypothesis as a plausible explanation of wage inequality using data at the firm level for Colombia, Mexico, and Taiwan. They obtained the following results: i) firm investments in technology have the largest impact on wage size distribution for skilled workers. ii) It had the smallest impact on wages paid to unskilled workers. iii) Lastly, a decomposition of wage effects by sources of technology revealed that wage premiums paid to skilled workers are led primarily by firm investment in R&D and training. Such conclusions seem to support the skilled-biased technological change hypothesis.<sup>17</sup> According to the typology used by Johnson

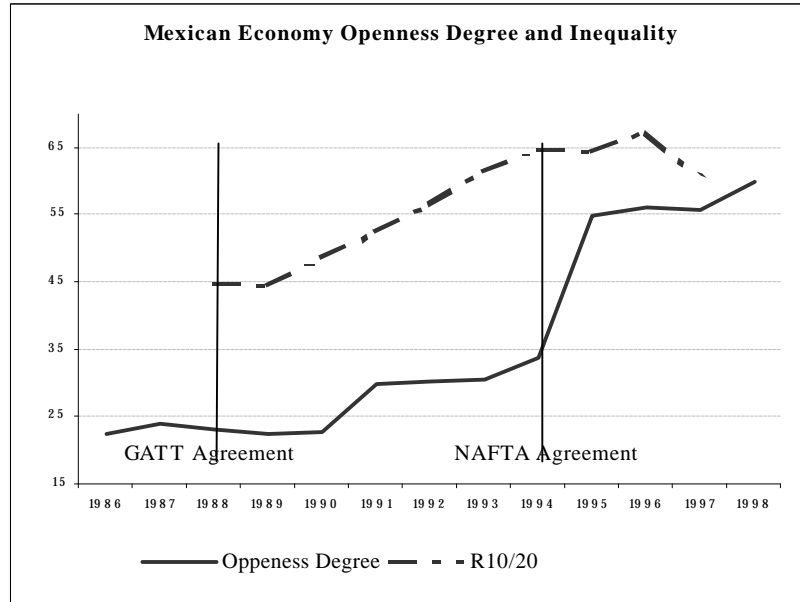
---

<sup>16</sup> See Hernandez Laos et al. "Productividad y Mercado de Trabajo en México", 1997

<sup>17</sup> Note that these results should be considered carefully, since the analysis is based on data at the firm level and just for the manufacturing industry.

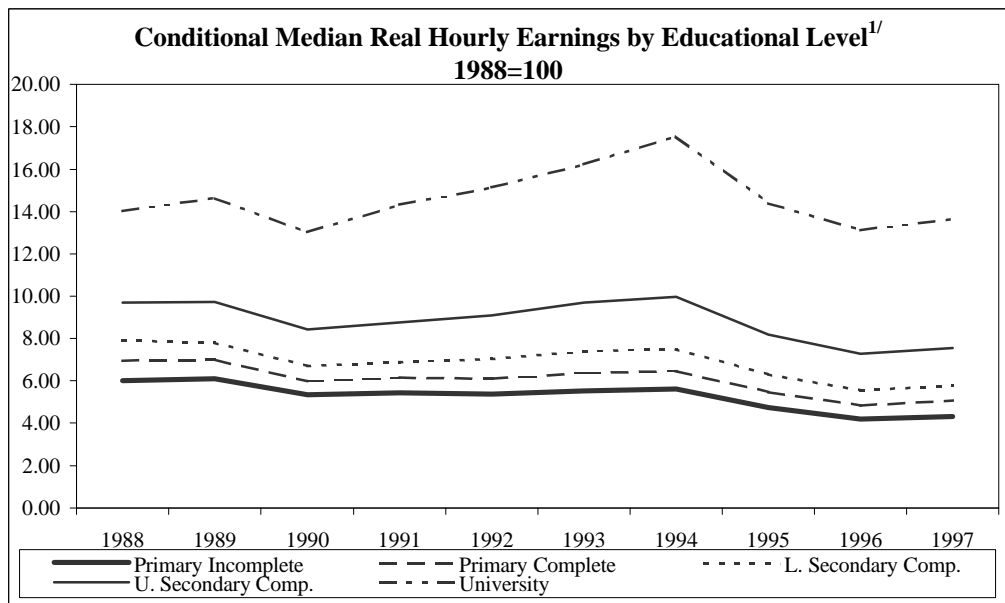
(1997), the type of technological change that drives wages up for the more highly skilled workers and drives wages down for the less skilled workers (as occurred in both the United States and Mexico) is extensive skilled-biased technological change. Under this type of technological change, skilled workers become more efficient in jobs that were traditionally performed by unskilled workers.

**Figure 1a**



Source: Own calculations based on ENIGH and INEGI data

**Figure 1b**



Source: Own estimations based on ENEU survey.

<sup>1/</sup> Medians were calculated conditional on experience, experience squared, gender, economic sector, labor market status, and region.

As it is shown in figure 1b<sup>18</sup>, all series have the same trend for all period. However, from 1990 conditional real earnings for University increased substantially, while conditional real earnings for lower educational levels remained steady, up to 1994. After that, it seems that the earnings differentials among all educational levels remained constant. This suggests that other factors rather than supply of new basic educational comers drove earnings differentials among level of schooling.

In sum, demand and supply, interacting within a context of economic modernization and globalization generate the trend towards greater wage disparity. It should be noted, however, that none of these explanations deal explicitly with the changes in the distribution of education, as well as the interaction between the educational policies that induced them and the workings of the labor market.

### III. STATIC DECOMPOSITION

This section aims at evaluating the contribution to earnings inequality in Mexico of a set of variables, either related to individual attributes, as schooling and age, or form of participation in the labor market, as number of hours worked or status, for selected years from 1988 to 1997. The idea is to measure the reduction in inequality that results from excluding the differences in average earnings among workers in different groups formed by those variables. When the exercise is conducted for a single variable, this reduction is said to be the **gross contribution** of such a variable to the overall wage inequality. When a variable is added to a model that contains all the remaining ones, the change in the gross contribution of these two models is called the **marginal contribution** of the added variable. In other words, the gross contribution can be regarded as the uncontrolled explanatory power of a given variable, and the marginal contribution as its explanatory power controlled by a set of other seemingly relevant variables.

#### *SHORT REVIEW*

Before proceeding to the decomposition exercise, it is worth to review the conclusions of other recent studies in relation to the evolution of earnings inequality and some variables that are important in the process of earnings formation.

Cragg and Epelbaum (1996) show that both average wage and education skill premium, which is defined as the percentage increase in wages over the primary schooling group, have increased substantially for more educated workers. In other words, the higher the level of education the larger the increase in average wage is, which in turn leads to an increase in inequality. They also examined whether the high demand for skilled labor is industry specific, task specific or simply general education. In order to assess the marginal contribution of other factors that are not related to education, these factors are controlled by a set of dummy variables that describe the industry and task specific effects. The authors concluded that the industry-specific effect was small and that the task-specific effect (occupational variable) explained half of the growing wage dispersion from 1987 to 1993. This conclusion, however, may not be correct, as occupation might be considered an endogenous variable, which is determined by education. As shown on table 2 in annex 3, educational level and occupational variables are highly correlated. In contrast, the correlation between education and other variables are low. Hence the occupation variable should be carefully handled in any kind of analysis.

---

<sup>18</sup> Median real hourly earnings were estimated using quantile regression models ( $\theta=0.5$ ) and conditioned on experience, gender, labor market status, economic sector and region (see annex 1 for groups definition).

## METHODOLOGY

The approach, which is used in this study, uses inequality measures known as “generalized entropy indices”. Bourguignon (1979), Cowell (1980) and Shorrocks (1980, 1984) have shown that only such measures satisfy all the desirable properties for any inequality measure and are additive decomposable<sup>19</sup>.

Assume that the population is divided into  $g$  groups (according to education, for instance). Then, a measure of inequality is said to be additive decomposable (see Shorrocks (1980)) when it can be written as:

$$I = I(\mathbf{b}_g, \mathbf{a}_g, I_g) = I_B(\mathbf{b}_g, \mathbf{a}_g) + \sum_g w(\mathbf{b}_g, \mathbf{a}_g) I_g \quad (1)$$

where  $\mathbf{b}_g$  is the fraction of the labor force employed in the group  $g$ ,  $\mathbf{a}_g$  is its relative mean income, and  $I_g$  represents the wage dispersion within this group as measured by the index  $I$ .

The term  $I_B$  on the right side of (1) corresponds to the inequality *between* groups (i.e. the amount of inequality that would be observed in the case of an earnings redistribution within each group, in such a way that, at the end, all workers in a group would receive the same earnings). The second term in the right-hand side (from here on called  $I_W$ ) reflects the inequality *within* groups; i.e., the share of overall inequality that is associated to factors other than those involved in the particular partition under study. It represents the degree of inequality that would be observed if all groups had the same average earnings. Notice that  $I_W$  is a weighted average of the internal inequalities, the weights,  $w(\mathbf{b}_g, \mathbf{a}_g)$ , being a function of the population share and average earnings of each group.

One can thus estimate the contribution of a given variable(s) to the overall earnings inequality at a given point in time as the fraction of this inequality that would be eliminated if the average wage of all groups formed by that (those) variable(s) were equalized, while keeping the internal dispersions unchanged. The *rationale* behind this exercise is that the effect of this (these) variables is captured by differences in average earnings at group level.

Amongst the most commonly used inequality indices, the Theil T is one of the few that is additive decomposable.<sup>20</sup> The general statistics needed for the decomposition by age, sector, level of schooling, hours worked and status from 1988 to 1997 are shown on table 6a and table 6b.

**Table 6a General Statistics for the Static Decomposition**

Variable	1988			1992			1996			1997		
	Beta	Alfa	Theil	Beta	Alfa	Theil	Beta	Alfa	Theil	Beta	Alfa	Theil
<b>Schooling</b>												
Primary Incomplete	0.185	0.70	0.220	0.147	0.65	0.234	0.129	0.57	0.283	0.127	0.57	0.207
Primary Complete	0.277	0.81	0.257	0.259	0.72	0.207	0.244	0.65	0.270	0.237	0.67	0.207
Lower Secondary Comp	0.241	0.88	0.228	0.264	0.80	0.281	0.257	0.74	0.264	0.263	0.76	0.229
Upper Secondary Comp	0.189	1.09	0.234	0.205	1.07	0.300	0.216	1.04	0.278	0.221	1.05	0.259
University Complete	0.107	2.10	0.343	0.124	2.32	0.359	0.154	2.30	0.430	0.151	2.22	0.289
<b>Total</b>			<b>0.327</b>			<b>0.395</b>			<b>0.464</b>			<b>0.372</b>
<b>Age</b>												
16-25	0.320	0.74	0.202	0.323	0.68	0.201	0.280	0.64	0.239	0.282	0.66	0.217
26-34	0.278	1.07	0.259	0.276	1.07	0.334	0.279	1.02	0.332	0.274	1.05	0.320
35-49	0.282	1.17	0.364	0.293	1.24	0.441	0.323	1.26	0.541	0.327	1.21	0.374
50-65	0.119	1.13	0.475	0.108	1.14	0.521	0.119	1.08	0.589	0.117	1.13	0.496
<b>Total</b>			<b>0.327</b>			<b>0.395</b>			<b>0.464</b>			<b>0.372</b>

Source: Own calculations based on the ENEU (3<sup>rd</sup> quarter).

Note: The sample includes only those who reported information on schooling level, age, economic sector, and labor market status simultaneously.

<sup>19</sup> Annex 2.4 has a review of several decomposition analysis methods.

<sup>20</sup> For the decomposition of the Theil T, see Ramos (1990) and annex 2.2.

**Table 6b General Statistics for the Static Decomposition**

Variable	1988			1992			1996			1997		
	Beta	Alfa	Theil	Beta	Alfa	Theil	Beta	Alfa	Theil	Beta	Alfa	Theil
<b>Sector</b>												
Primary Sector	0.019	0.99	0.508	0.016	0.99	0.667	0.014	1.20	0.976	0.012	1.20	0.621
Manufacturing Industry	0.274	0.97	0.323	0.242	0.96	0.379	0.221	0.94	0.559	0.227	0.92	0.371
Non Manuf. Industry	0.058	0.91	0.224	0.064	1.06	0.409	0.060	0.91	0.382	0.057	0.88	0.331
Commerce	0.178	1.01	0.415	0.196	0.92	0.415	0.188	0.90	0.484	0.180	0.89	0.407
Finance Services/Rent	0.030	1.39	0.230	0.027	1.77	0.384	0.024	1.90	0.407	0.027	1.79	0.332
Transp./communication	0.066	1.12	0.191	0.069	1.12	0.310	0.064	1.03	0.344	0.068	1.06	0.255
Social Services	0.253	1.10	0.280	0.261	1.12	0.380	0.294	1.23	0.373	0.293	1.25	0.317
Other Services	0.122	0.73	0.385	0.125	0.70	0.291	0.136	0.58	0.274	0.136	0.58	0.269
Total			0.327			0.395			0.464			0.372
<b>Hours Worked</b>												
20 – 39	0.201	0.89	0.278	0.174	0.87	0.391	0.174	0.84	0.399	0.172	0.86	0.333
40 – 48	0.581	0.96	0.280	0.566	0.95	0.332	0.525	0.96	0.421	0.540	0.98	0.331
49 - +	0.218	1.20	0.438	0.260	1.20	0.483	0.301	1.16	0.535	0.288	1.13	0.444
Total			0.327			0.395			0.464			0.372
<b>Status</b>												
Employer	0.046	2.32	0.549	0.048	2.44	0.463	0.048	2.18	0.561	0.046	2.15	0.428
Self employed	0.158	0.97	0.338	0.149	0.89	0.354	0.174	0.75	0.377	0.167	0.79	0.340
Informal salaried	0.122	0.58	0.210	0.140	0.54	0.158	0.147	0.47	0.174	0.150	0.48	0.175
Formal salaried	0.609	0.99	0.240	0.602	1.03	0.342	0.558	1.14	0.412	0.567	1.13	0.311
Contract	0.064	0.98	0.230	0.062	0.96	0.297	0.072	0.77	0.302	0.070	0.79	0.268
Total			0.327			0.395			0.464			0.372

Source: Own calculations based on the ENEU (3<sup>rd</sup> quarter).

Note: The sample includes only those who reported information on schooling level, age, economic sector, and labor market status simultaneously.

## RESULTS

The results for the exercise of static decomposition are shown on table 7.<sup>21</sup> Education (the result of the interaction between demand and supply) is by far the variable that accounts for the largest share of earnings inequality in Mexico, both in terms of its gross and marginal contributions. The gross contribution, i.e., its explanatory power when it is considered alone, amounts to one fifth of total inequality in 1988 and one third in 1997.<sup>22</sup> The marginal contribution, i.e., the increase in the explanatory power when it is added to a model that already has the other variables, is remarkably stable and meaningful, staying around 21% throughout the whole period. It is worth pointing out that the difference between the two contributions has been increasing over time, indicating that the degree of correlation and other variables has been going up, i.e., the “indirect” effects are becoming more important.

<sup>21</sup> Since this exercise is very intensive in the number of observations (which constitutes its main handicap) the variable “hours worked” was dropped in order to avoid the problems with cells with too few observations. The decision was made through the comparison among different combinations of variables, where hours worked ended up being the least relevant.

<sup>22</sup> In most earnings equations for any country, the set of measurable observable variables explains at most 60% of the total variance. In the United States, education accounts for 10% of the total variance.

**Table 7. Contribution to the Explanation of Earnings Inequality(%)**

Variables	1988		1992		1996		1997	
	Gross	Marginal	Gross	Marginal	Gross	Marginal	Gross	Marginal
Education	20.2	20.8	26.9	21.6	29.3	21.2	32.6	21.2
Age	5.4	8.3	7.2	6.1	6.6	6.2	7.3	5.4
Economic Sector	2.3	8.1	4.0	5.2	6.8	5.2	8.6	4.4
Status	12.8	11.2	13.7	8.9	13.7	7.4	15.6	7.5

Source: Own calculations based on ENEU.

The other variables considered seem to be much less important. The three of them, and particularly the economic sector and status in the labor market, display an upward trend in their gross contribution, and a declining one in their marginal contribution. This can be interpreted, as evidence that the interaction between these variables and education has become more intense. That is, the workers' skills are becoming increasingly more relevant to the determination of their type of participation in the labor market, as well as for their position across different economic segments of the economy. Note that the same pattern holds when hours worked instead of sector is considered (see table below).

**Table 8. Contribution to the Explanation of Earnings Inequality (%)**

Variables	1988		1992		1996		1997	
	Gross	Marginal	Gross	Marginal	Gross	Marginal	Gross	Marginal
Education	20.2	20.2	26.9	22.3	29.3	22.6	32.6	24.5
Age	5.4	5.4	7.2	4.8	6.6	4.6	7.3	4.5
Hours worked	1.7	3.8	1.9	3.3	1.3	4.0	1.2	3.5
Status	12.8	7.6	13.7	7.1	13.7	6.0	15.6	6.5

Source: Own calculations based on the ENEU (3<sup>rd</sup> quarter).

The analysis of these results leads to the conclusion that education is a key variable for the understanding of earnings inequality in Mexico.<sup>23,24</sup> Even though this is to some extent a remarkable finding, it comes as no surprise in the Latin American context. The results for some countries in the region, where similar exercises were carried out, are reported on table 9. Mexico stays on the average range for Latin American countries, and displays a situation close to that observed in Colombia and Peru. However, education seems to be more important for inequality in Brazil, and much less important in Argentina and Uruguay. It is important to stress the fact that this is a comparison in relative terms. Given that in Colombia and Peru, where education has a similar explanatory power, there is a lower degree of inequality compared to Mexico, the absolute contribution of education is higher in Mexico. As a matter of fact, in absolute terms, the contribution of education to inequality in Mexico is the second highest in Latin America, next only to Brazil. Moreover, what seems to be particularly interesting in the Mexican experience is the fact that the significance of education has been increasing over time. Therefore, the inspection of the evolution of the educational distribution and the income profile associated to it, as well the link between changes in this distribution and changes in earnings inequality will be addressed in the next sections.

<sup>23</sup> Additional evidence is that the explanatory power of the complete model was 42.5% in 1988, 45.0% in 1992, 45.5 in 1996, and 48.3% in 1997. This means that the marginal contribution of education is almost equal to the joint contribution of age, economic sector, and status in the labor market.

<sup>24</sup> Miguel Székely (1995) applied the static decomposition of the Theil to the ENIGH for the years 1984, 1989 y 1992. He used education, occupation, region, economic sector, and job status as control variables. The main finding is that this set of variables explain 55%, 58% y 64% of income dispersion for each respective year with education and job status being the relevant variables.

**Table 9. Contribution of Education to Earnings Inequality. International Comparison**

Country	Author(s)	Period	Gross Contribution (%)
<b>Latin America</b>	Altimir and Piñera (1982)	1966/74	17-38
<b>Argentina</b>	Fiszbein (1991)	1974/88	16-24
<b>Brazil</b>	Ramos and Trindade (1992)	1977/89	30-36
	Vieira (1998)	1992/96	30-35
<b>Colombia</b>	Reyes (1988)	1976/86	29-35
	Moreno (1989)	1976/88	26-35
<b>Costa Rica</b>	Psacharopoulos et alt. (1992)	1981/89	23-26
<b>Peru</b>	Rodríguez (1991)	1970/84	21-34
<b>Uruguay</b>	Psacharopoulos et alt. (1992)	1981/89	10-13
<b>Venezuela</b>	Psacharopoulos et alt. (1992)	1981/89	23-26

#### IV. THE EVOLUTION OF EDUCATIONAL ATTAINMENT

Educational attainment levels increased rapidly in most developing countries since the 1950s (Schultz (1988)). While Mexico also partook in that development, earlier studies had identified a significant lag in its educational indicators. Londoño (1996) for example, points to an “education deficit”, according to Latin American countries in general, and Mexico in particular, have approximately two years less of education than would be expected for their level of development.<sup>25</sup> Elias (1992) found that education was the most important source of labor quality improvement in Latin America between 1950 and 1970, but points out that such improvements did not take place to the same extent in Mexico as in other countries in the region. This changed dramatically in the 1980s, figure 2 shows that although Mexico’s educational attainment increased steadily since the 70’s, it continued to remain below the international trend line<sup>26</sup>. In the 1980s, however, the growth of educational attainment in Mexico accelerated, permitting it to catch up with international standards by 1990; where its placement in figure 2 is slightly above the trend line.

In the closure of Mexico’s education gap vis-à-vis the rest of the world was hastened in part by the country’s economic stagnation. Mexico’s real GDP per capita in the mid-1990s was roughly the same as it had been in the first half of the 1980s. Nevertheless, the preceding observation should not detract from the remarkable increase in schooling that occurred during the 1980s. While the level of average schooling in Mexico increased by roughly a year per decade during 1960-1980 (from 2.76 to 4.77 years), it increased by two years in the 1980-1990 decade. The acceleration in schooling during the 1980s, in turn, was the product of concerted efforts to increase basic education coverage combined with advances made in the reduction of primary school repetition and dropout rates.

<sup>25</sup> On the other hand, Behrman (1987) classifies Mexico as an overachiever in what comes to the relation between economic development and educational progresses in the context of developing countries.

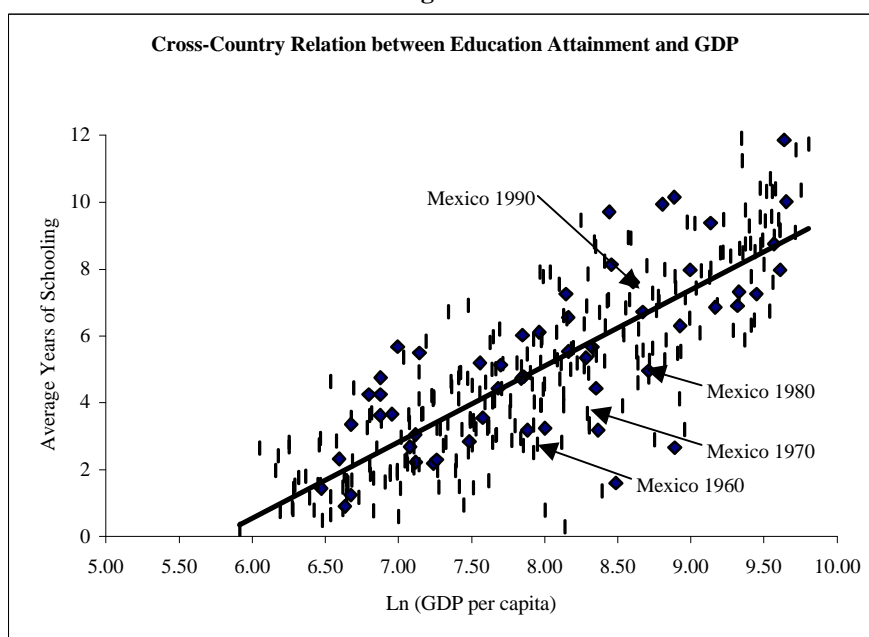
<sup>26</sup> The scatter diagram is based on 317 observations from five different years. The trend line represents the least square regression line given by:

$$S = -13.17 + 2.28 \ln(GDPcap) \quad \text{Adj. } R^2 = 0.68$$

(-18.7)      (26.0)      t-values in parentheses

The application of Ramsey’s RESET test to this regression equation failed to detect a specification error; unlike with the alternative specification of type:  $S = a + bX + cX^2$ .

Figure 2



The observations pertaining to Mexico, ordered by date, are as follows:

Year	Average Schooling (years)	Ln (GDP per capita; 1980 US\$)
1960	2.76	7.95
1970	3.68	8.29
1980	4.77	8.71
1985	5.20	8.63
1990	6.72	8.67

With respect to changes in the distribution of schooling by socioeconomic groups, there are several aspects to be considered. In particular, three of them are examined here: the changes in this distribution related to gender, economic sector and age.

Table 10 shows the schooling distribution by gender from 1988 to 1997. There one can see that, even though there were clear improvements for both males and females, which translates to an upgrade of educational attainment, women achieved a better performance during that period, especially at the top of the distribution. Improvements for males, on the other hand, were more evenly spread over the entire distribution. Nevertheless, in 1997 it is possible to state that women were undoubtedly more educated than men, as their cumulative distribution dominates that of men (see figure 3).<sup>27</sup>

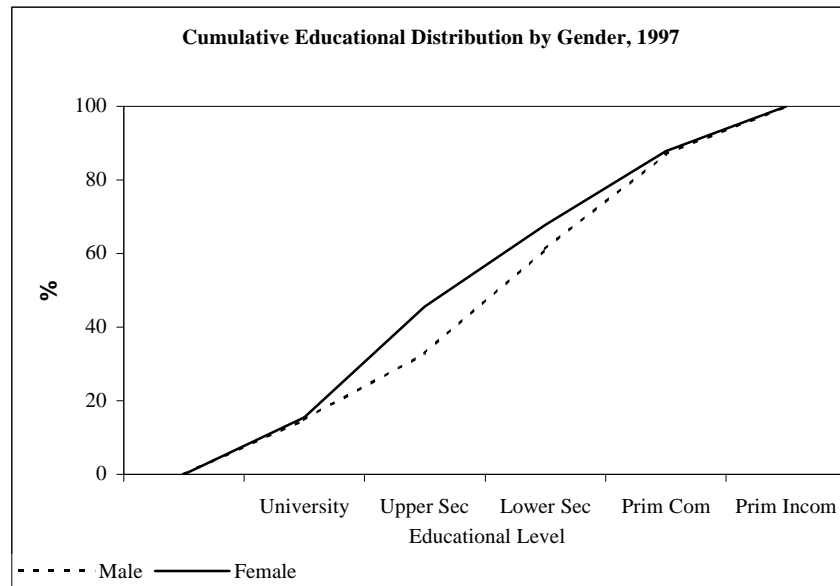
Table 10. Evolution of Educational Distribution by Gender (%)

Educational Group	Primary Incomplete	Primary Complete	Lower Secun. Complete	Upper Secun. Complete	University Complete
<b>1988</b>					
Male	19.0	30.1	24.5	14.6	11.8
Female	17.3	22.2	23.2	29.1	8.2
Total	18.5	27.7	24.1	18.9	10.7
<b>1997</b>					
Male	13.0	25.7	28.4	18.0	14.9
Female	12.2	20.0	22.3	30.1	15.5
Total	12.7	23.7	26.3	22.1	15.1

Source: Own calculations based on the ENEU survey (3rd quarter).

<sup>27</sup> This remark is true for the 1997 overall distribution relative to the 1988 one.

**Figure 3**



Source: Own calculations based on ENEU data

With respect to the distribution of schooling by economic sector, table 11 shows that there has been a significant upgrade from 1988 to 1997. Three points, nonetheless, deserve to be stressed. First, financial and social services industries became relatively more intensive in the use of high-skilled labor. Second, the primary sector, together with non-manufacturing industry and other services, were characterized by more intensive use of low-skilled labor. Third, in a surprising way, the manufacturing industry, in contrast to what seems to be the common wisdom, cannot be characterized as a sector that intensively uses high-skilled labor.

**Table 11. Evolution of Educational Distribution by Economic Sector (%)**

<b>Educational Group</b>	<b>Primary Incomplete</b>	<b>Primary Complete</b>	<b>Lower Secon. Complete</b>	<b>Upper Secon. Complete</b>	<b>University Complete</b>
<b>1988</b>					
Primary Sector	41.1	21.0	13.3	14.3	10.3
Manufacturing Industry	16.2	33.3	27.8	14.7	8.0
Non Manufacturing Industry	36.6	28.5	14.7	9.0	11.2
Commerce	18.0	28.7	28.8	18.7	5.8
Finance Services/Rent	4.8	6.1	19.5	47.1	22.5
Transportation/communication	14.4	35.7	26.0	18.9	5.0
Social Services	11.3	17.6	21.7	28.2	21.2
Other Services	32.8	36.6	20.2	8.1	2.3
Total	18.5	27.7	24.1	18.9	10.7
<b>1997</b>					
Primary Sector	28.1	27.4	17.7	10.9	15.9
Manufacturing Industry	11.0	29.5	32.7	18.2	8.7
Non Manufacturing Industry	28.6	31.7	18.4	10.0	11.4
Commerce	12.4	23.4	30.6	24.1	9.5
Finance Services/Rent	2.7	5.4	16.1	40.3	35.6
Transportation/communication	9.1	26.8	32.2	23.9	8.0
Social Services	6.0	13.2	21.1	29.6	30.0
Other Services	26.2	35.7	24.6	11.1	2.4
Total	12.7	23.7	26.3	22.1	15.1

Source: Own calculations based on the ENEU (3<sup>rd</sup> quarter).

Another relevant observation is that the age groups also experienced upgrades in their educational attainment, as the distribution by educational level in 1997 is above the one in 1988 (table 12). In an attempt to reach a better understanding of this event, it is interesting to contrast the time and cohort effects.<sup>28</sup> In order to do this, one can look at the first age groups, 16-25 and 26-34, like synthetic cohorts. Namely, the 26-34 age group in 1997 can be directly compared to the 16-25 in 1988, and, in to a lesser extent, the 35-49 in 1997 to the 26-34 in 1988. From 1988 to 1997 the percentage of those in the primary incomplete level decreased, this reduction was higher than that experienced by the 16-25 age group (later being the 26-34 in 1997). The opposite took place for the highest level of instruction. In other words, it seems that the improvements throughout the educational process in Mexico are significant, both for those entering the system (higher coverage) and for those already in there (higher efficiency).

**Table 12. Evolution of Educational Distribution by Age Groups (%)**

<b>Educational Group</b>	<b>Primary Incomplete</b>	<b>Primary Complete</b>	<b>Lower Sec. Complete</b>	<b>Upper Sec. Complete</b>	<b>University Complete</b>
<b>1988</b>					
16-25	8.5	26.5	36.7	23.7	4.6
26-34	12.6	23.7	23.1	22.5	18.2
35-49	24.0	33.3	16.8	14.3	11.6
50-65	46.1	27.2	9.9	9.0	7.8
Total	18.5	27.7	24.1	18.9	10.7
<b>1997</b>					
16-25	5.8	23.8	38.7	25.5	6.2
26-34	6.9	19.5	28.1	27.0	18.5
35-49	14.8	25.8	19.5	19.1	20.7
50-65	37.3	27.6	11.5	10.6	13.0
Total	12.7	23.7	26.3	22.1	15.1

Source: Own calculations based on the ENEU (3rd quarter).

Also concerning the interaction between age and education, one can argue that the effect of developments in the educational system is more important for the new generations than for the elderly. To investigate this, it is necessary to contrast the behavior of inequality between different age groups to that of inequality within synthetic cohorts and in relation to education. As seen above, the younger cohorts are in fact better educated. At the same time the “within” income dispersion for the youngest cohorts seems to increase over time, compared the internal Theil in 1997 and 1988 (see table 6). Thus, it becomes easier to understand why the gross contribution of age to inequality has been going up and at the same time its marginal contribution has been decreasing. In other words, differences in both educational attainment and distribution among cohorts have become pronounced in recent times, leading to a higher correlation (negative) between education and age.

## **V. THE DYNAMIC DECOMPOSITION**

### *AN STYLIZED VIEW*

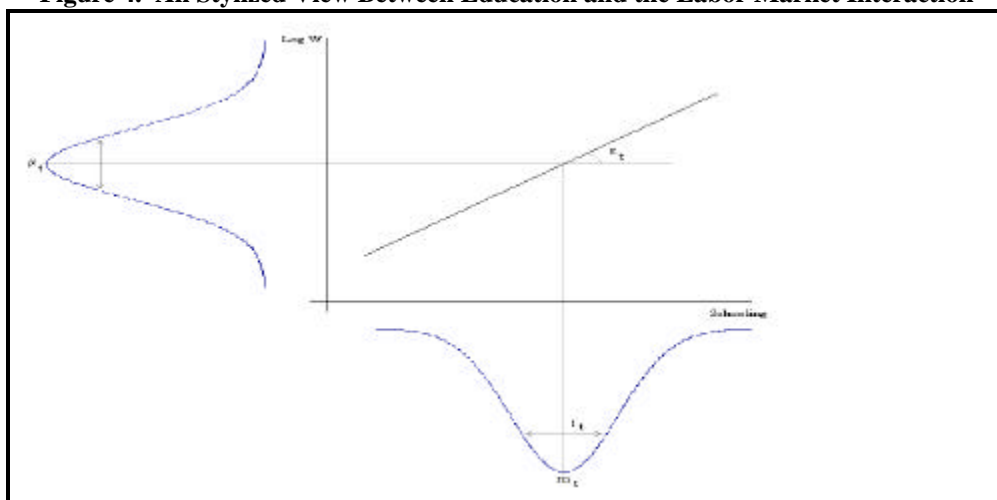
In order to address the relationship between education (the result of the interaction between supply and demand) and earnings inequality it is necessary to explain the role of the labor market, since the way it works determines the earnings differentials among workers with different educational attributes. Thus, this relationship can be viewed as being determined by two elements: (i) the distribution of education itself; and (ii) the way the labor market rewards educational attainment. The first element reflects a pre-existing social stratification that already entails some

<sup>28</sup> The time effect refers to the comparison of the same age group in two different points of time.

inequality, due to reasons other than the workings of the labor market itself. The second is associated to the degree of growth of this pre-existing inequality into earnings inequality due to the performance of the labor market (i.e. demand behavior).

The diagram below shows the distribution of education in the horizontal axis ( $m_t$  is an indicator of the average schooling of the labor force and  $i_t$  represents its dispersion) while the vertical axis has the distribution of earnings. The first quadrant depicts the interaction between the pre-existing conditions (the distribution of education) and the workings of the labor market, through the steepness  $s_t$  of the income profile related to education. Therefore, at a point of time: (i) the higher  $m_t$  is, the larger the average earning will be; (ii) the lower  $i_t$  is, the smaller the earnings inequality will be; and (iii) the higher  $s_t$  is, the bigger the growth of pre-existing disparities, and, accordingly, the higher the earnings inequality will be. As these indicators change over the time, there are going to be alterations in the income distribution induced by them: changes in  $i_t$ , assuming  $s_t$  constant, will change earnings inequality due to changes in the composition of the labor force (the so-called allocation/population effect), whereas changes in  $s_t$  will produce alterations in the earnings differentials (the income effect).

**Figure 4. An Stylized View Between Education and the Labor Market Interaction**



Barros and Reis (1991) developed three synthetic measures for the indicators  $m_t$  (average schooling),  $i_t$  (schooling inequality), and  $s_t$  (income profile), based directly on the definition of the Theil T index (see annex 2.3). The figures for Mexico from 1988 to 1997 are presented in the table below. As it can be seen, there was some improvement on average schooling, but the inequality of the distribution of education has deteriorated over the period studied, whereas the income profile, which is related to the returns to schooling, has become much steeper. Meaning that, there was a shift in demand towards high skilled labor that was not met by the increase in supply probably due to the increased rate of skill-biased technological change, whose transmission to Mexico may be facilitated by the increased openness of the economy. The same pattern observed for the overall sample holds for the 16-25 years old age group: the  $m_t$  goes up from 0.561 to 0.574 in 1988 through 1997; the  $i_t$  increases from 0.0196 to 0.0218, whereas the  $s_t$  doubles going from 0.0196 to 0.0383.

**Table 13. Synthetic Indicators of Schooling Distribution and Income Profile**

Year	1988	1992	1996	1997
$m_t$	0.476	0.491	0.511	0.510
$i_t$	0.066	0.069	0.076	0.075
$s_t$	0.066	0.102	0.122	0.111

Source: Own calculations based on the ENEU survey (3rd quarter).

## METHODOLOGY

The dynamic decomposition analysis is a suitable tool for translating this stylized view in quantitative results, giving one a better understanding of the socioeconomic transformations responsible for changes in the earnings distribution. Besides permitting the identification of the relevant individual variables, it also helps in understanding the nature of the contribution of each variable to the evolution of earnings inequality over time.

It has been shown by Ramos (1990), following Shorrocks (1980), that it is possible to break down the change in inequality between two points in time. This is done according to whether it can be attributed to changes in the socioeconomic groups relative to incomes, to group sizes or in their internal inequalities, through the use of the Theil T index. In generic terms, as shown before in a slightly different way, for a given partition of the population, the inequality indices of this class can be written as:

$$I = I(\mathbf{a}_g, \mathbf{b}_g, I_g)$$

where:  $\mathbf{a}_g$  is the ratio between the average income of group  $g$  and the average income of the whole population,  $\mathbf{b}_g$  is the proportion of the population in group  $g$ , and  $I_g$  is the internal dispersion of incomes in group  $g$ .

Of course the  $\mathbf{a}$ 's are related to the indicator  $s_i$  in the previous picture, as well as the  $\mathbf{b}$ 's refer to  $m_i$  and  $i_i$ . In this context, the *population or allocation effect* corresponds to the variation induced in the inequality index  $I$  by modifications in the allocation of the population among the groups (changes in the  $\mathbf{b}$ 's), with no direct changes in the group's relative incomes ( $\mathbf{a}$ 's).<sup>29,30</sup> The *income effect* corresponds to the changes in  $I$  induced by changes in group incomes ( $\mathbf{a}$ 's), without changing the group population shares ( $\mathbf{b}$ 's), and the internal effect is the change in the inequality caused only by modifications in the dispersions at group level (the  $I_g$ 's).<sup>31</sup> The expressions corresponding to the Theil T index are derived in annex 2.2.

## RESULTS

The results of the decomposition of the variations in the Theil T index for different intervals of time are shown in the table below. The first point to highlight is the fact that, when the variables are considered alone, education has the highest gross contribution to the explanation of changes in earnings distribution. Second, both the allocation and the income effect were positive in all periods. This means that the changes in the distribution of education and in the relative earnings among educational groups were always in phase with the alterations in the earnings distribution. Namely, when the income profile related to education became steeper and the inequality of education increased, the earnings distribution worsened (as in the 1988-1992, 1992-1996, and 1988-1997 periods), and vice-versa (as in the 1996-1997 period).

Third, the income effect is always the prevalent one. If one considers, for instance, the 1988-1997 period, the changes in the relative earnings among educational groups alone, would have generated deterioration in the earnings distribution higher than the one observed. To lesser extent, the same holds true for the other periods.<sup>32</sup> Even the decrease in inequality observed between 1996 and 1997 is partially explained by the changes in relative earnings (it is possible to see that in table 13 the income profile related to education became less steep in this period). Therefore, it seems reasonable to conclude that the income effect is the leading force underlying

<sup>29</sup> The difference between this and what Knight and Sabot (1983) call the "compression" effect is that in the present exercise we are including the indirect change induced in  $I$  through the variation in the weights of the  $I_g$ 's.

<sup>30</sup> Of course the individuals  $\alpha$ 's change as the  $\beta$ 's change, since the overall average income is altered. This indirect impact is also computed in the composition effect (see the annex 2.2).

<sup>31</sup> The methodology applied by Fields (1996) and Bouillon, et. al. (1998) makes important assumptions as it was pointed out in section III.2. On the other hand, Székely (1995), in order to explain the inequality changes between two points of time, applied a methodology that differs drastically from the dynamic decomposition since he does not control for the effects that arise from changes in the population distribution and from changes in the relative income groups earnings considered in the partition of the population (see annex 2.4).

<sup>32</sup> Of course the explanation for such a phenomenon is that the changes in the other variables worked in the direction of attenuating the changes in the rewards to education.

the increase in inequality, and that, in turn, suggests that the workings of the labor market, and its interaction with the educational policies, should be thoroughly examined.

Fourth, it is worth pointing out that the significance of changes in the distribution of education remains high even when one controls for changes in other relevant variables.<sup>33,34</sup> As a matter of fact, with the exception of the 1996-1997 transitional period, the marginal contribution of age, economic sector and status in the labor market are usually negative. This means that the changes in these variables contributed to reduce the effects induced by changes related to education, as most of the time they work in the direction of reducing inequality after the influence of education is accounted for.

**Table 14. Results of the Dynamic Decomposition**

Period	Variable	Allocation	Income	Gross	Marginal
1988-1992	Education	<b>11.4</b>	<b>58.8</b>	<b>70.2</b>	<b>30.5</b>
	Age	-1.8	21.9	20.2	-5.2
	Sector	-0.6	7.8	7.1	-17.7
	Status	3.9	15.1	19.0	-7.4
1992-1996	Education	<b>23.9</b>	<b>32.8</b>	<b>56.7</b>	<b>27.6</b>
	Age	11.1	10.5	21.6	10.5
	Sector	-5.4	25.4	20.0	10.5
	Status	1.2	12.4	13.6	-4.2
1996-1997	Education	<b>2.2</b>	<b>15.5</b>	<b>17.7</b>	<b>24.2</b>
	Age	-0.4	5.9	5.5	12.5
	Sector	0.4	1.0	1.4	18.4
	Status	1.4	6.1	7.5	7.8
1988-1997	Education	<b>35.8</b>	<b>108.4</b>	<b>144.1</b>	<b>33.7</b>
	Age	7.4	32.7	40.1	-19.9
	Sector	-6.6	43.2	36.6	-40.6
	Status	9.0	20.2	29.2	-35.6

Source: Own calculations based on the ENEU (3<sup>rd</sup> quarter).

The last period, from 1996 to 1997, deserves special comment. First because inequality was substantially reduced. Secondly because, once more, there were alterations associated with education, now working in the other direction, and such alteration appear to be the main factor responsible for the reduction in inequality. As it can be seen from the synthetic indicators, there was a small improvement in the distribution of schooling during the period and, a sizable decrease in the steepness of income profile related to education. All other variables, as observed for other periods, also contributed to an improvement in earnings inequality.

The next table shows the results of the same kind of decomposition for Brazil, Argentina and Peru. The significance of education as an explanation of changes in inequality seems to be a common pattern in Latin American countries. Moreover, the relevance of the income effect over the allocation (population) effect is also a trait shared by all countries where a similar analysis was carried out. Interestingly, in the Mexican case the figures are above those for other countries (in a shorter period of time length, one should stress). That means that the changes in the structure of supply and demand for labor, which are greatly affected by the educational and macroeconomic policies followed by the country and/or their interaction with the workings of the labor market, were particularly relevant for the earnings distribution.

**Table 15. Education and Inequality Variation: Brazil, Argentina and Peru**

<sup>33</sup> Székely (1995) concludes that for the 1984-1989 period, the variables that highly contributed to explaining inequality were education and economic sector, while education and job status were significant in the 1984-1992 period. The selected variables were education, occupation, region, economic sector, and job status.

<sup>34</sup> Bouillon, et al. (1998), applying Bourguignon methodology to the ENIGH, found that the return effect to the household characteristics (age/gender, education/age, assets) explained 49% of the increase in the Gini between 1984 and 1994, education being the most important explanatory variable. The region effect (urban/ rural) was 9%, the south effect 15%. And the population effect 23%.

Country	Author(s)	Period	Explanatory Power (%)*	Income Effect (%)
Brazil	Ramos and Trindade (1992)	1977/1989	6-20	10-17
Argentina	Fiszbein (1991)	1974/1988	54-56	38-46
Peru	Rodríguez (1991)	1970/1984	32-47	34-43

\*The explanatory power is the income plus the allocation/population effect.

## VI. THE EVOLUTION AND STRUCTURE OF THE RATES OF RETURNS TO EDUCATION: AN APPLICATION OF QUANTILE REGRESSION

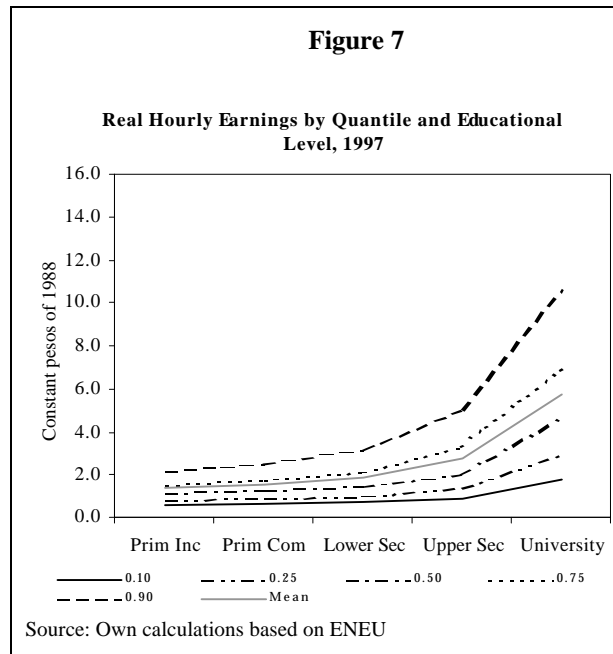
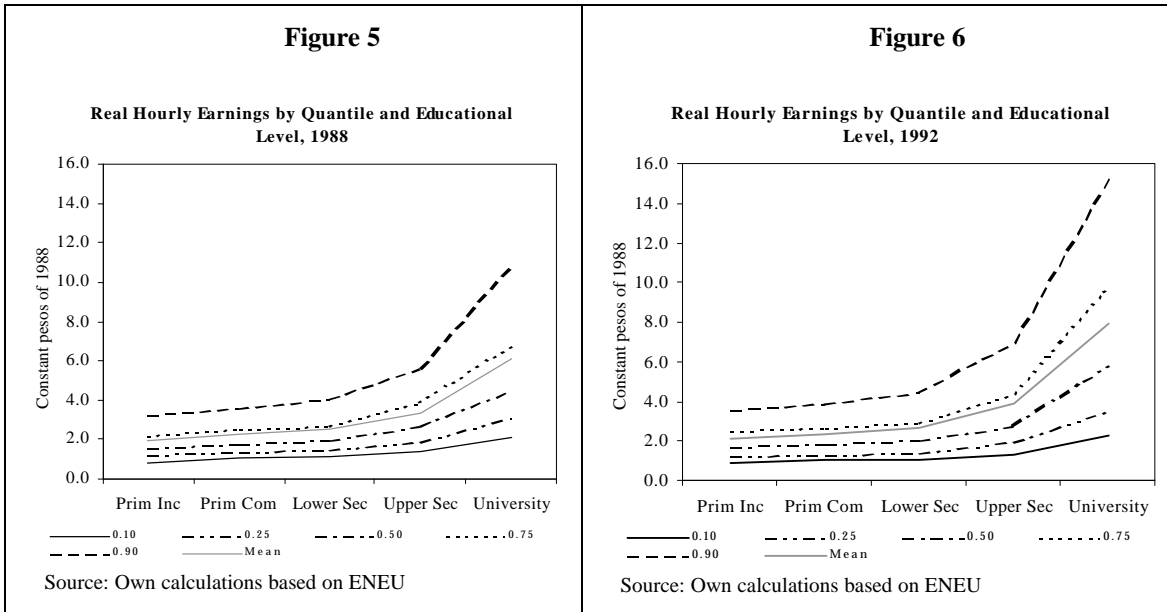
The increase in earnings inequality is not the result of a worsening in the distribution of education, whereas the income profile, which is related to the returns to schooling, has become much steeper. In light of this evidence, this section analyzes the structure and evolution of the rates of return to education. Though this is a common procedure, this is an important caveat, as the international comparison becomes cumbersome because the structure of the educational process in Mexico is different to other countries.

### *QUANTILE ANALYSIS*

Before the rate of returns to education estimation, it is necessary to take a preliminary look at the relationship between the earnings distribution and the educational attainment in Mexico. For this purpose, real hourly earnings by quantile (0.10, 0.25, 0.50, 0.75 and 0.90) and the mean are computed<sup>35</sup>.

As it can be seen from figure 5 through 7 the curves do not cross each other for all educational categories and for all periods. This suggests that there is a strict dominance of the education variable throughout the earnings distribution. In other words, there is a positive relation between educational level and hourly earnings throughout the distribution. Those figures also show that the difference among quantiles (i.e. from 10<sup>th</sup> to 25 percentile, from 25<sup>th</sup> to 50<sup>th</sup> percentile, etc.) change throughout educational level (the greater the level of education, the larger the difference among quantiles of hourly earnings). In addition, the difference between quantiles also changes through time. These patterns may provide empirical evidence that there are differences in the increase of real hourly earnings throughout educational distribution and time. The quantile analysis provides a complete assessment of the impact of many variables (education, age, gender, economic sector, labor market status, region, etc) throughout the earnings distribution. Finally, for all educational categories real average hourly earnings is greater than the median; then, the hourly earnings distribution is always right-skewed.

<sup>35</sup> The third quarter of the ENEU data for 1988, 1992, and 1996 were used. The sample is described in the appendix.



Summarizing, the above results suggest that a quantile method of estimation is needed for providing a better understanding of the rate of returns to education.

### ***METHODOLOGY***

One concise way to estimate the returns to schooling is by using conventional Mincerian earnings equations based on the human capital paradigm, with controls for other characteristics (individual attributes or labor market traits) that might influence the differentials. This approach allows one to disentangle the association between individual earnings and levels of education from the joint influence of other variables on earnings.

In this study an ad hoc, yet usual, specification will be used, with control variables for age (used as a rough proxy for experience), labor market status, economic sector allocation, and geographical region.<sup>36</sup> Then the earnings function can be described as follows:

<sup>36</sup> All of these variables are categorical, with the exception of age. Therefore, it is necessary to leave one category

$$\log Y_t = a_t + S_t b_t + X_t c_t + u_t, \quad t = 1988, 1992, 1996, \text{ and } 1997 \quad (1)$$

where,

$Y_t$  Vector of individual hourly earnings in time  $t$

$a_t$  Logarithm of the mean real hourly earnings of the reference group in time  $t$

$b_t$  Earnings differential associated to education in time  $t$ <sup>37</sup>

$c_t$  Vector of earnings differentials related to the control variables in time  $t$

$S_t$  Vector of educational attainment in time  $t$

$X_t$  Matrix of control variables for time  $t$

$u_t$  Vector of residual terms for time  $t$  [ $E(u_t) = 0$  and  $E(u_t u_t') = 0$ ].<sup>38</sup>

These earnings functions can be fitted using “least squares” estimation. However, a new technique of estimation has been developed recently, the quantile regression models. This technique has usually been applied to analyze the determinants of wage structure as well as rates of returns to investment in education throughout the earnings distribution. Buchinsky (1994), (1995), and (1998) applies this technique to the U.S. labor market in order to assess the wage structure and its changes. Poterba and James (1994) also use the quantile regression to study the pattern of U.S. wage differentials between state and local government employees and their private counterparts. The quantile regression analysis has also been applied to other countries: Shultz and Mwabu (1996) in South Africa, Muller (1998) in Canada, Abadie (1997) in Spain, and Montenegro (1998) in Chile. It is important to note that this methodology has never been applied in Mexico.

The quantile regression models have some desirable characteristics, especially when one is interested in the analysis of certain variable throughout its distribution. The main features of these models can be summarized as follows:

- i) The model can be used to characterize the entire conditional distribution of the dependent variable;
- ii) the quantile regression objective function is a weighted sum of absolute deviations, which gives a robust measure of location, so that the estimated coefficient vector is non sensitive to outlier observations of dependent variable;
- iii) when the error term is non-normal, quantile regression estimators may be more efficient than least squares estimators; and,
- iv) different solutions at distinct quantiles may be interpreted as differences in the responses of the dependent variable to changes in the independent variables at various points in the conditional distribution of the dependent variable (see Buchinsky (1998)).

The earnings function (1) can be rewritten as a quantile regression model. Then, we have,

$$\log Y_i = X_i \mathbf{b}_q + \mathbf{m}_q$$

$$\text{with } \text{Quant}_q(\log Y_i/X_i) = X_i \mathbf{b}_q \quad (i=1, \dots, n)$$

Where  $\mathbf{b}_q$  and  $X_i$  are  $K \times 1$  vectors, and  $X_{i1} \equiv 1$ ;

$\text{Quant}_q(\log Y/X)$  denotes the  $q$ th conditional quantile of  $Y$  given  $X$ ;

Also let  $f_{\mathbf{m}_q}(\cdot|X)$  denote the density of  $\mathbf{m}_q$  given  $X$  [it follows that  $\text{Quant}(\mathbf{m}_q|X)=0$ ].

(reference group) per variable out of regression for avoiding perfect collinearity. Primary incomplete (education), formal salaried workers (labor market status), manufacturing (economic sector), and Mexico City (region) were the selected groups left out.

<sup>37</sup> As this is a categorical variable, one has, in fact, a vector [ $b_{it}$ ] of earnings differentials, with each of its components representing the earnings differential between the  $i$ -th educational group and the reference one (primary incomplete) in time  $t$ .

<sup>38</sup> In addition, one has to assume that the residual term is not correlated with the unobserved determinants of individual earnings (family background, natural ability, etc.).

Note that the  $X_i$  vector includes the set of explanatory dummy variables as well as the controls. For an extensive review see Buchinsky (1994).

By using the regression coefficients, one can compute the differentials and marginal value related to each educational level. According to the specification of the earning functions, for least squares as well as for quantile regression, the exponential of the differential associated to the  $j$ -th category of the  $i$ -th variable,  $\exp(c_{ij})$ , corresponds to an estimate of how much higher, on average, are the earnings of an individual in that category relative to the earnings of an individual in the reference group for that variable<sup>39</sup>, all other attributes being identical. The marginal value of some educational level  $j$  in time  $t$  ( $MV_j^{educ}$ ) can be interpreted as the earnings differential for this level relative to the previous one, as follows:<sup>40</sup>

$$MV_{jt}^{educ} = b_{jt} / b_{(j-1)t} \quad \text{for } j > 1 \quad \text{and} \quad MV_j^{educ} = b_{jt} \quad \text{for } j = 1$$

### EMPIRICAL RESULTS

Both ordinary least squares and quantile regression models were estimated.<sup>41</sup> However, before analyzing the rates of return to education it is worth to investigate the role of each explanatory variable in the determination of earnings. For such purpose, several regressions were fitted adding one by one of the explanatory variables. This exercise has two advantages: i) it allows us to assess the marginal contribution of each explanatory variable; and ii) it shows the explanatory power of each variable throughout the conditional earnings distribution. Cragg and Epelbaum (1996) perform a similar exercise. Nevertheless, the occupation variable was left out of this study, since such variable, as is structured in ENEU-INEGI questionnaires, is highly correlated to the individuals' educational level. As it is shown in table 2 in Annex 3, the educational level and occupation are highly correlated, while education and the rest of the other explanatory variables are low correlated.

As it was indicated, education is the most important variable in the explanation of earnings inequality. However, we can assess the importance of other explanatory variables using the estimates of educational level differentials. If the changes of such differentials, in a given period of time, have been smoothed by other explanatory variable, then such variable is a measure of some specific-skill. For doing such assessment, the relative change in the differentials by educational level in 1988-1992 and 1992-1997 periods were computed. The estimates are presented below.

**Table 16. Change in Differentials Controlling for Economic Sector, Labor Market, Status and Region**

Education level	Controlling for none		Economic Sector		Status		Economic Sector and Status		Economic Sector, Region and Status	
	1988-92	1992-97	1988-92	1992-97	1988-92	1992-97	1988-92	1992-97	1988-92	1992-97
Period										
Primary Complete	-0.03	0.05	-0.01	0.02	-0.04	0.03	-0.02	0.02	-0.02	0.01
Lower-Secondary Complete	-0.06	0.08	-0.05	0.03	-0.06	0.03	-0.03	0.00	-0.03	0.00
Upper-Secondary Complete	-0.02	0.11	0.02	0.04	-0.02	0.04	0.01	0.00	0.01	0.00
University Complete	0.14	0.18	0.15	0.08	0.12	0.09	0.15	0.04	0.15	0.04

Source: Own calculations based on ENEU (3<sup>rd</sup> quarter).

Note: Least squares estimates. The reference group is "Primary incomplete".

Table 16 shows that the changes in earnings differentials were smoothed by the introduction of

<sup>39</sup> If the differential is close to zero then it can be interpreted as being approximately equal to the average percentage increase in earnings associated to a movement from the reference group to  $j$ -th category, *ceteris paribus*.

<sup>40</sup> Similarly, the definition applies to the results of the quantile regression approach. The only difference is that in this situation one needs an additional subscript ( $q$ ), to assign the quartile.

<sup>41</sup> The  $q$ s parameters in the quantile regression were 0.1, 0.25, 0.5, 0.75, and 0.9, following a common procedure in the literature.

the economic sector variable in the regression for 1992-1997 period, particularly for tertiary education,<sup>42</sup> while in 1988-1992 period the “smooth” effect was very small. Labor market status seems to have the same “smooth” effect pattern as the economic sector variable. These results suggest that the degree of correlation between education and economic sector, as well as labor market status, changed. In addition, table 16 also shows that the “smooth” effect of the region variable was almost zero for both periods. In other words, the inclusion of this variable in the regression only produced a very small change in the earnings differentials.

At this point, one tentative conclusion emerges: the smooth effect of both economic sector and labor market status variables were significantly larger in 1992-1997 period than in 1988-1992 period (before the trade agreement). This means that the relationship between those types of specific skills acquired through such variables and education changed in the labor market. This implies that worker’s insertion into the labor market and economic sector variables were a consequence of skill’s differentials and not solely attributed to education. Hence, in order to have a precise assessment of the marginal value to educational level the analysis must incorporate this based on the earnings regression conditioned on economic sector, labor market status, region, as well as age, age squared and gender.

Table 17 presents the marginal value of education by level. In the regression estimates, all the coefficients for education were significant at the 5% level, and the results for the marginal value of each educational level are reported in the table below.

In general the OLS estimates are quite similar to the ones obtained by the quantile regression approach for  $q=0.5, 0.75$ . It is also true, nevertheless, that the estimates through the latter technique tend to increase as one moves from the right to the left of the conditional earnings distribution, particularly for the upper levels of education. In summary, the results above have three strong implications: *i) education does play a crucial role in the process of earnings formation; ii) its effect is not the same throughout the conditional earnings distribution; and iii) the marginal value of education has not changed significantly in basic education.*

**Table 17. Marginal Value of Education by Level**

Quantile	1988						1992					
	0.10	0.25	0.50	0.75	0.90	OLS	0.10	0.25	0.50	0.75	0.90	OLS
Primary Complete	1.15	1.15	1.16	1.18	1.19	1.19	1.12	1.13	1.13	1.14	1.16	1.16
Lower-Secondary Comp	1.11	1.11	1.14	1.17	1.20	1.17	1.10	1.12	1.15	1.18	1.21	1.15
Upper-Secondary Comp	1.13	1.18	1.23	1.26	1.26	1.27	1.20	1.25	1.30	1.35	1.39	1.32
University Complete	1.34	1.39	1.44	1.46	1.52	1.49	1.46	1.54	1.66	1.70	1.69	1.69
Quantile	1996						1997					
	0.10	0.25	0.50	0.75	0.90	OLS	0.10	0.25	0.50	0.75	0.90	OLS
Primary Complete	1.14	1.14	1.15	1.17	1.20	1.15	1.15	1.16	1.17	1.18	1.18	1.18
Lower-Secondary Comp	1.12	1.13	1.15	1.18	1.20	1.16	1.11	1.12	1.14	1.18	1.22	1.14
Upper-Secondary Comp	1.21	1.25	1.31	1.40	1.48	1.34	1.20	1.25	1.31	1.39	1.47	1.32
University Complete	1.60	1.71	1.80	1.78	1.70	1.74	1.63	1.76	1.80	1.77	1.70	1.75

Source: Own calculations based on ENEU (3<sup>rd</sup> quarter).

Note 1: The marginal value is with respect to the previous educational level

Note 2: The asymptotic covariance matrix of the estimated coefficient vector in quantile regression is computed using the bootstrap method

Note 3: All the coefficient are statistically significant at 5%, and conditioned to age, squared age, gender, status in the labor market, economic sector and region (North, Center, South and Mexico City).

Specifically, one can say that the rewards to education display a log-convexity for all years investigated. This log-convexity, however, has become pronounced in the 1988-1996 period, as the marginal value for the higher levels increased relatively more. There was a reversal in this trend in 1997, basically due to the gains associated with primary complete and losses associated with upper secondary, though in a slight way.

Regarding the accumulated changes in the marginal value of education by level (table 18), these are not significant for the primary complete and lower secondary instruction levels, along

<sup>42</sup> For a reference see Cragg and Epelbaum (1996).

the conditional earnings distribution for OLS estimates. This does not apply for upper secondary level, as the changes were substantial and very progressive across quantiles (8% at the median and 23% at the top decile). The changes were more important for the University level.

**Table 18. Percentage Change in the Marginal Value of Education, 1988-1997**

Quantile	0.1	0.25	0.5	0.75	0.9	OLS.
Primary Complete	0	1	1	0	-1	-1
Lower-Secondary Complete	1	1	0	1	2	-3
Upper-Secondary Complete	7	7	8	14	23	5
University Complete	34	45	43	36	20	30

Source: Own calculations based on ENEU.

In sum, the returns to education have increased in Mexico in recent times, especially for higher levels of education and in the upper tail of the conditional earnings distribution.

### *RATE OF RETURNS TO EDUCATION AND INEQUALITY*

The transition probabilities describe the shifts of the labor force within and across sectors by skilled versus unskilled workers. The transition probabilities is the conditional probability of finding a worker in economic sector 'k' at the end of the period given that the worker began in the sector 'j'. This probability gives us the mobility of less/high skilled workers between 'j' and 'k' economic sector.<sup>43</sup> The "between" probability is the mobility of unskilled/skilled workers between 'j' and 'k' economic sectors. By contrast, the "within" mobility depicts workers who move across sub-sector or occupations Tables 19a and 19b present the transition probabilities for the respective periods. On the one hand, financial service sector shows a clear trend to substitute unskilled for skilled labor force: the probability of workers changing to another economic sector is much higher for the unskilled (70%) than for the skilled labor force (21%). Primary sector follows the same trend only at the end of the 1980's. On the other hand, non-manufacturing industry is substituting skilled for unskilled workers. Finally, manufacturing industry, transportation and communications sectors do not have a clear dominance of the probabilities by skilled/unskilled workers.

**Table 19a. Transition Probabilities of Being in the Same Sector, Change Within Sector and Change Between Sector by School Level<sup>1/</sup>**

Sector	1988-1989			1992-1993			1996-1997		
	Not Change	Sector Change		Not Change	Sector Change		Not Change	Sector Change	
		Within	Between		Within	Between		Within	Between
<b>Incomplete upper secondary or lower</b>									
Primary Sector	46.4	14.4	39.3	28.5	46.8	24.7	47.8	13.6	38.6
Manufacturing Industry	52.7	24.5	22.7	53.6	23.1	23.4	56.9	24.5	18.6
Non Manufacturing Industry	45.6	11.8	42.6	50.3	7.6	42.1	46.9	9.4	43.7
Commerce	48.0	16.2	35.7	60.3	14.0	25.6	53.6	17.7	28.7
Financial Services/Rent	25.7	4.1	70.2	35.6	8.0	56.4	67.6	2.7	29.6
Transportation/communication	65.1	6.7	28.2	71.7	8.0	20.2	72.5	8.4	19.0
Social Services	59.0	21.2	19.8	61.6	18.3	20.2	66.9	14.7	18.4
Other Services	59.7	8.8	31.5	65.7	4.9	29.4	60.5	2.3	37.2
<i>Weighted average</i>	54.2	17.5	28.3	58.7	15.3	26.0	59.2	14.6	26.1

Source: Own calculations based on the ENEU survey (3rd quarter).

The length of time is one year. The sample includes those in the labor force and in the panel

1/ Sector change within sector was defined according to change between sub-sector

**Table 19b. Transition Probabilities of Being in the Same Sector, Change Within Sector and Change**

<sup>43</sup> Skilled workers are those individuals with more than 12 years of schooling.

**Between Sector by School Level<sup>1/</sup>**

Sector	1988-1989			1992-1993			1996-1997		
	Not Change	Sector Change		Not Change	Sector Change		Not Change	Sector Change	
		Within	Between		Within	Between		Within	Between
<b>Complete upper secondary or higher</b>									
Primary Sector	41.3	0.0	58.7	42.3	6.1	51.6	35.4	21.2	43.4
Manufacturing Industry	42.3	29.9	27.8	50.7	24.7	24.6	53.2	27.4	19.3
Non Manufacturing Industry	57.7	8.1	34.1	51.0	15.4	33.6	54.0	6.5	39.4
Commerce	41.9	13.6	44.4	52.2	15.4	32.4	51.6	15.4	33.0
Financial Services/Rent	77.0	1.3	21.6	69.2	6.1	24.7	66.2	18.6	15.1
Transportation/communication	50.9	18.9	30.2	74.7	8.3	17.0	69.6	6.1	24.4
Social Services	55.1	34.0	10.9	63.8	23.2	13.0	71.5	18.2	10.3
Other Services	45.1	4.4	50.4	56.5	0.5	43.0	56.5	1.4	42.0
<i>Weighted average</i>	51.1	25.1	23.8	59.1	19.1	21.8	64.0	17.4	18.7

Source: Own calculations based on the ENEU survey (3rd quarter).

The length of time is one year. The sample includes those in the labor force and in the panel

1/ Sector change within sector was defined according to change between sub-sector

Using both “within” and “between” economic sector shifts, one can explore the effect of these shifts on the relative wage of skilled and unskilled workers. Tables 19a and 19b also show that for all periods considered the “between” probability for skilled versus unskilled labor force is substantially higher; conversely, the “within” probability for the skilled is significantly lower compared to the unskilled. Therefore, one might infer that the relative wage of unskilled labor relative to skilled labor increased, derived from within economic sector shifts. However, this effect might have been partially offset by the decrease in relative wages of unskilled labor relative to skilled labor, derived from the between economic sector shifts. Therefore and given the results of rates of returns to education, it is plausible to infer that the relative demand shifts within economic sectors dominated the relative demand shifts between sectors.

With the goal of putting the rates of returns results in perspective, table 20 shows the percentage of earnings differentials for other Latin American countries. Mexico is above the average, only after Brazil (the country that has the highest inequality in Latin America). Once more, this is an indication that educational policies must be at the core of any effort aimed at reducing inequality, and by extension poverty, in Mexico.

**Table 20. Percent Earnings Differentials by Country**

	Latin America	Mexico	Brazil	Argentina	Peru
Primary Complete	50%	100%	100%	35%	40%
Upper Secondary Complete	120%	170%	170%	80%	80%
University Complete	200%	260%	280%	160%	145%

Reference group: non-schooling

Source: IDB (1998).

## VII. CONCLUSIONS

Even though the educational attainment levels expanded very rapidly, Mexico also experienced a pronounced increase in the degree of income inequality over the period of analysis. Most of the worsening of income distribution happened in the mid-eighties and early nineties, display little variation in earnings inequality except for a small trend towards deterioration.

There are two main differences in the pattern shown by the earnings and total current income distribution. First, the gains are not limited to the richest 10%, as those in the seventh, eight, and nine tenths of the distribution who also improved their relative earnings over the period by almost two percentage points. Second, there is a clear worsening in the earnings distribution in the present decade throughout 1996. On the other hand, the inequality associated with the total current income was moderately stable in the nineties, displaying an improvement in 1996.

The different behavior between total current income and labor earnings inequalities from 1994 to 1996 gives support to the idea that the poor, who rely the most on labor as a source of income, were the least able to protect themselves during the recession. However, the substantial drop in earnings inequality from 1996 to 1997 is a surprising finding.

This paper shows that education is by far the variable that accounts for the largest share of earnings inequality in Mexico, both in terms of its gross and marginal contributions. In addition, the gross contribution of age to inequality has been going up and at the same time its marginal contribution has been decreasing. In other words, differences in both educational attainment and distribution among cohorts have become pronounced in recent times, leading to a higher correlation (negative) between education and age. The education contribution to inequality in Mexico is the second highest in Latin America, next to Brazil. Moreover, what seems to be particularly interesting in the Mexican experience is the fact that the significance of education has been increasing over time.

The increase in earnings inequality, however, does not appear to be the result of a worsening in the distribution of education, whereas the income profile, which is related to the returns to schooling, has become much steeper. Meaning that there was a shift in demand towards high skilled labor that was not met by the increase in supply probably due to the increased rate of skill-biased technological change, whose transmission to Mexico may be facilitated by the economy's increased openness.

The rates of returns to education analysis showed that there is a strict dominance of education throughout the earnings distribution. That means that there is a positive relation between educational level and earnings throughout the entire conditional earnings distribution. It is also shown that there are differences in the increase of earnings throughout education distribution and time. Finally, real average hourly earnings are greater than their respective medians for all educational categories; that is, the hourly earnings distribution is always right-skewed.

The analysis of the rates of returns to education indicated that: i) the changes in earnings differentials were smoothed by the introduction of the economic sector variable in the regression for 1992-1997 period, particularly for tertiary education, while in 1988-1992 period the "smooth" effect was very small. Labor market status seems to have the same "smooth" effect pattern as the economic sector variable. These results suggest that the degree of correlation between education and economic sector, as well as labor market status, changed. In addition, it was shown that the "smooth" effect of the region variable was almost zero for both periods. In other words, the inclusion of this variable in the regression only produced a very small change in the earnings differentials. ii) Education does play a crucial role in the process of earnings formation and iii) the returns to education have increased in Mexico in recent times, especially for higher levels of education and in the upper tail of the conditional earnings distribution.

Using both "within" and "between" economic sector shifts, it is observed that the "between" probability for skilled versus unskilled labor force is substantially higher. Conversely, the "within" probability for the skilled is significantly lower compared to the unskilled. Therefore, one might infer that the relative wage of unskilled labor relative to skilled labor increased, derived from within economic sector shifts. However, this effect might have been partially offset by the decrease in relative wages of unskilled labor relative to skilled labor, derived from the

between economic sector shifts. In conclusion and given the results of rate of return to education, it is plausible to infer that the relative demand shifts within economic sectors dominated the relative demand shifts between sectors.

## ANNEX I. DATA SOURCES

The National Household Income and Expenditure Survey (ENIGH) and the National Urban Employment Survey (ENEU) were used in this study.

### 1.1 ENIGH

The National Household Income and Expenditures Survey is collected by the Instituto Nacional de Estadística, Geografía e Informática (INEGI). This survey is available for 1984, 1989, 1992, 1994 and 1996<sup>44</sup>. Each survey is representative at the national level, urban and rural areas. For 1996, the ENIGH is also representative for the states of Mexico, Campeche, Coahuila, Guanajuato, Hidalgo, Jalisco, Oaxaca and Tabasco.

For each year the survey design was stratified, multistage and clustered. The final sampling unit is the household and all the members within the household were interviewed. In each stage, the selection probability was proportional to the size of the sampling unit. Then, it is necessary to have the use of weights<sup>45</sup> in order to get suitable estimators.

The table below shows the sample size for each year.

Year	Number of households	Number of persons
1984	4,735	23,756
1989	11,531	56,727
1992	10,530	50,378
1994	12,815	59,835
1996	14,042	64,359

The available information can be grouped into three categories:

- Income and consumption: the survey has monetary, no monetary and financial items.
- Individual characteristics: social and demographic, i.e., age, schooling attendance, level of schooling, position at work, sector, etc.
- Household characteristics.

### Category Selection

For the purpose of the analysis, the individuals in the sample were classified according to their educational level, position in occupation, sector of activity and geographical region in the following categories:

#### a) Educational level

- i) Primary incomplete: no education and primary incomplete (one to five years of primary)
- ii) Primary complete: primary complete and secondary incomplete (one or two years)
- iii) Secondary complete: secondary complete and preparatory incomplete (one or two years)
- iv) Preparatory complete: preparatory complete and university incomplete
- v) University complete: university complete (with degree) and postgraduate studies

#### b) Position in occupation

- i) Worker or employee

<sup>44</sup> The sample in a given year is independent from another.

<sup>45</sup> The weights should be calculated according to the survey design and corresponds to the inverse of the probability inclusion.

- ii) Employer
  - iii) Self employed
- c) *Sector of activity*
- i) Agriculture
  - ii) Manufacturing
  - iii) Construction
  - iv) Commerce
  - v) Services
  - vi) Other (utilities, extraction, transports, financial services, communications, etc)
- d) *Geographical regions*
- i) North: Baja California, Baja California Sur, Coahuila, Chihuahua, Durango, Nuevo Leon, Sinaloa, Sonora, Tamaulipas and Zacatecas
  - ii) Center: Aguascalientes, Colima, Guanajuato, Hidalgo, Jalisco, Mexico, Michoacan, Morelos, Nayarit, Puebla, Queretaro, San Luis Potosi and Tlaxcala
  - iii) South: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz and Yucatan
  - iv) Distrito Federal.

### Group Selection

The labor force was limited to individuals who are:

- i) working as employee, employer or self employed<sup>46</sup>;
- ii) between 12 and 65 years old;
- iii) living in urban areas;
- iv) working 20 hours or more per week;
- v) with positive income;
- vi) having the attributes of interest defined.

The number of persons in the survey that belong to the labor force is shown in the next table.

**Table 2. Sample size for the labor force**

Year	Number of persons	% of the total sample
1984	3,892	16.4
1989	10,401	18.3
1992	8,752	17.4
1994	10,982	18.4
1996	12,996	20.2

According to the groups mentioned we have that,

<sup>46</sup> The respective categories: workers without payment and cooperative members were excluded because of the sample size.

**Table 3. Sample size by variable and year**

Variable	1984	1989	1992	1994	1996
<b>Education Level</b>					
Primary Incomplete	1,246	1,951	1,879	2,387	2,736
Primary Complete	1,299	3,006	2,501	2,975	3,411
Secondary Complete	803	2,875	2,489	3,014	3,734
Preparatory Complete	389	1,614	1,168	1,617	1,915
University Complete	245	955	715	989	1,200
<b>Position in Occupation</b>					
Employee	3,175	8,604	7,188	8,843	10,207
Employer	126	311	393	450	610
Self employed	681	1,486	1,171	1,689	2,179
Total	3,982	10,401	8,752	10,982	12,996

## 1.2. ENEU

The National Urban Employment Survey is also a micro-level data set collected by INEGI and contains quarterly wage and employment data over the last ten years (1987-1997). Currently, the data is representative of the 41 largest urban areas in Mexico, covering 61% of the urban population following the 2500 inhabitants or more criteria, and 92% of the population who live in metropolitan areas with 100,000 or more inhabitants. In 1985, the ENEU included 16 urban areas: Mexico City, Guadalajara, Monterrey, Puebla, Leon, San Luis Potosi, Tampico, Torreon, Chihuahua, Orizaba, Veracruz, Merida, Ciudad Juarez, Tijuana, Nuevo Laredo and Matamoros, covering 60% of the urban population for that year. In 1992, 18 more urban areas were included in the survey: Aguascalientes, Acapulco, Campeche, Coatzacoalcos, Cuernavaca, Culiacan, Durango, Hermosillo, Morelia, Oaxaca, Saltillo, Tepic, Toluca, Tuxtla Gutierrez, Villahermosa, Zacatecas, Colima and Manzanillo. In 1993 and 1994, Monclova, Queretaro, Celaya, Irapuato and Tlaxcala entered to ENEU. Finally, Cancun and La Paz joined the survey in 1996. As can be seen in the previous description, the ENEU has always covered about 60% of the national urban population. Therefore, the results deduced from this survey allow one to know and assess the socioeconomic and employment characteristics of the national urban areas.

The data is from household surveys, which fully describe family composition, human-capital acquisition, and experience in the labor market (the variables contain information about social household characteristics, activity condition, position in occupation, unemployment, main occupation, hours worked, earnings, benefits, secondary occupation, and searching for another job). As the ENIGH, the sampling design was stratified, in several stages (where the final selection unit was the household), and with proportional probability to size<sup>47</sup>. So this statistical construction allows us to make comparisons among different years. Moreover, this survey is structured to generate a panel data set which has the characteristic to conform to a rotator panel (a fifth of the total sample goes out and a new one comes in every quarter). Hence, the panel data follows the same household through out five quarters.

## Category Selection

The individuals in the sample were classified according to their educational level, age, sector of activity, position in occupation, hours worked and geographical region in the following categories:

<sup>47</sup> For this it was necessary to use weights or expansion factors.

- a) *Educational level*
- i) Primary incomplete: no education and primary incomplete (one to five years of primary)
  - ii) Primary complete: primary complete and secondary incomplete (one or two years)
  - iii) Secondary complete: secondary complete and preparatory incomplete (one or two years)
  - iv) Preparatory complete: preparatory complete and university incomplete
  - v) University complete: university complete (with degree) and postgraduate studies
- b) *Age*
- i) 12 to 25 years old
  - ii) 26 to 34 years old
  - iii) 35 to 49 years old
  - iv) 50 to 65 years old
- c) *Sector of activity*
- i) Primary sector (Includes Agriculture, Forestry, Fishing, and Mining).
  - ii) Manufacturing industry
  - iii) Not manufacturing industry (Includes Construction, and Utilities)
  - iv) Commerce
  - v) Finance services and Rent
  - vi) Transportation and Communication
  - vii) Social services (Tourism, education, health, public administration, embassy)
  - viii) Other services
- d) *Labor Market Status*
- i) Employer
  - ii) Self employed
  - iii) Informal salaried: people that work in an enterprise with 15 workers or less and do not receive social security (IMSS, ISSTE, private, etc.)
  - iv) Formal salaried: people that work in an enterprise with 16 workers or more or receive social security (IMSS, ISSTE, private, etc.)
  - v) Contract
- e) *Hours worked*
- i) 20 to 39 hours per week
  - ii) 40 to 48 hours per week
  - iii) At least 49 hours per week
- f) *Geographical regions*
- i) North: Baja California, Baja California Sur, Coahuila, Chihuahua, Durango, Nuevo Leon, Sinaloa, Sonora, Tamaulipas and Zacatecas
  - ii) Center: Aguascalientes, Colima, Guanajuato, Hidalgo, Jalisco, Mexico, Michoacan, Morelos, Nayarit, Puebla, Queretaro, San Luis Potosi and Tlaxcala
  - iii) South: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz and Yucatan
  - iv) Distrito Federal.

### **Group Selection**

Analogous to the ENIGH, the sample is:

- i) between 16 and 65 years old;
- ii) living in urban areas (localities with at least 2,500 inhabitants);
- iii) regular workers (non-seasonal workers);
- iv) working 20 hours or more per week;

- v) with positive earnings<sup>48</sup>;
- vi) having the attributes of interest defined.

The table below shows the sample size and labor force.

**Table 4. Sample Size by Year**

Year	Number of persons	
	Total	Labor force
1988	124,322	45,870
1989	125,820	47,630
1990	127,387	48,109
1991	126,262	48,080
1992	235,696	91,279
1993	239,394	90,860
1994	246,906	102,105
1995	252,563	100,838
1996	262,478	108,159
1997	272,356	116,559

<sup>48</sup> In this survey an additional adjustment had to be made: if the worker got a benefit at the end of the year (“aguinaldo”), then the wage was expanded (we assumed that this benefit as equivalent to 30 days of wages a year).

## ANNEX 2. METHODOLOGICAL NOTE

### 2.1 GINI INDEX

The Gini Index is defined by

$$GI = \frac{2 \text{cov}[Y, F(Y)]}{\mathbf{m}} \quad (1)$$

where:

- $Y$  is the distribution of per capita income  $Y=(y_1, \dots, y_n)$ , where  $y_i$  is the per capita income of individual  $i$ ,  $i=1, \dots, n$
- $\mathbf{m}$  is the mean per capita income
- $F(Y)$  is the cumulative distribution of total per capita income in the sample (i.e.  $F(Y)=[f(y_1), \dots, f(y_n)]$  where  $f(y_i)$  is equal to the rank of  $y_i$  divided by the number of observations  $(n)$ <sup>49</sup>.

#### *Gini decomposition*

Equation (1) can be rewritten and expanded into an expression for the Gini coefficient that captures the “contribution to inequality” of each of the  $K$  components of income (see Leibbrandt (1996)).

$$GI = \sum_{k=1}^K R_k G_k S_k \quad (2)$$

where:

- $S_k$  is the share of source  $k$  of income in total group income (i.e.  $S_k = \mathbf{m}_k / \mathbf{m}$ )
- $G_k$  is the Gini coefficient measuring the inequality in the distribution of income component  $k$  within the group
- $R_k$  is the Gini correlation of income from source  $k$  with total income<sup>50</sup>.

The larger the product of these three components, the grater the contribution of income from source  $k$  to total inequality.

### 2.2 THEIL T INDEX<sup>51</sup>

---

<sup>49</sup> Both the covariance and cumulative distribution are computed using the household weights

<sup>50</sup>  $R_k$  is defined as:  $R_k = \frac{\text{cov}[Y_k, F(Y)]}{\text{cov}[Y_k, F(Y_k)]}$

<sup>51</sup> Theil's T is sensitive to changes at the bottom and the top tail of the distribution.

This index was calculated as follows<sup>52</sup>:

$$T = \left(\frac{1}{n}\right) \sum_{i=1}^n \left(\frac{Y_i}{\bar{Y}}\right) \ln \left(\frac{Y_i}{\bar{Y}}\right) \quad (3)$$

where:

- $Y_i$  is the income of the  $i$ -th individual
- $\bar{Y}$  is the average income
- $n$  is the population size.

#### *Static Decomposition of the Theil Index*

If the population is divided into  $G$  groups with  $n_g$  observations each, it is then possible to write (3) as:

$$T = \sum_{g=1}^G \left(\frac{1}{n}\right) \sum_{i=1}^{n_g} \left(\frac{Y_{ig}}{\bar{Y}}\right) \ln \left(\frac{Y_{ig}}{\bar{Y}}\right) \quad (4)$$

where:

- $Y_{ig}$  is the income of the  $i$ -th individual of the  $g$ -th population subgroup.

If we now define  $\mathbf{b}_g = n_g/n$  and  $Z_g = \bar{Y}_g/k$  where  $\bar{Y}_g$  is the average income of the  $g$ -th group and  $k$  is a reference income, it is possible to show, after some algebraic manipulation, that  $T$  can be expressed as:

$$T = \left(\frac{1}{k}\right) \sum_{g=1}^G \mathbf{b}_g Z_g \ln Z_g - \ln k + \left(\frac{1}{k}\right) \sum_{g=1}^G \mathbf{b}_g Z_g T_g \quad (5)$$

where:

$$k = \sum_{g=1}^G \mathbf{b}_g Z_g$$

- $T_g$  is the Theil index for the  $g$ -th group.

The first two terms on the right hand side of (5) correspond to the between group inequality, and the third one to the within group inequality.

Choosing the mean income as the reference income, i. e.,  $Z_g = \mathbf{a}_g = \bar{Y}_g/\bar{Y}$ , expression (5) simplifies to:

---

<sup>52</sup> The mathematical notation from in annex 2.2 through 2.3 2 follows Ramos (1990).

$$T = \sum_{g=1}^G \mathbf{a}_g \mathbf{b}_g \ln \mathbf{a}_g + \sum_{g=1}^G \mathbf{a}_g \mathbf{b}_g T_g \quad (6)$$

The first term in (6) is said to be the between group inequality, and the second term the within group inequality.

### *Dynamic Decomposition Analysis*

By totally differentiating (6), we have:

$$dT = \sum_{g=1}^G \frac{\partial T}{\partial \mathbf{b}_g} d \mathbf{b}_g + \sum_{g=1}^G \frac{\partial T}{\partial \mathbf{a}_g} d \mathbf{a}_g + \sum_{g=1}^G \frac{\partial T}{\partial T_g} d T_g \quad (7)$$

- The first term on the right hand side is the population allocation effect (changes in  $T$  caused exclusively by population shifts);
- The second term is the income effect (changes in  $T$  induced exclusively by changes in standardized mean incomes), and;
- The third one is the internal effect (changes in  $T$  caused by changes in internal dispersion).

It can be shown that:

$$\frac{\partial T}{\partial \mathbf{b}_g} = \mathbf{a}_g \ln \mathbf{a}_g - \mathbf{a}_g \sum_{g=1}^G \mathbf{a}_g \mathbf{b}_g (1 + \ln \mathbf{a}_g) + \mathbf{a}_g T_g - \mathbf{a}_g \sum_{g=1}^G \mathbf{a}_g \mathbf{b}_g T_g \quad (8)$$

$$\frac{\partial T}{\partial \mathbf{a}_g} = \mathbf{b}_g (1 + \ln \mathbf{a}_g) - \mathbf{b}_g \sum_{g=1}^G \mathbf{a}_g \mathbf{b}_g (1 + \ln \mathbf{a}_g) + \mathbf{b}_g T_g - \mathbf{b}_g \sum_{g=1}^G \mathbf{a}_g \mathbf{b}_g T_g \quad (9)$$

$$\frac{\partial T}{\partial T_g} = \mathbf{a}_g \mathbf{b}_g \quad (10)$$

Replacing (8), (9), and (10) into (7) and simplifying we obtain

$$dT = \sum_{g=1}^G \mathbf{a}_g (\ln \mathbf{a}_g + T_g - T - 1) d \mathbf{b}_g + \sum_{g=1}^G \mathbf{b}_g (\ln \mathbf{a}_g + T_g - T) d \mathbf{a}_g + \sum_{g=1}^G (\mathbf{a}_g \mathbf{b}_g) d T_g \quad (11)$$

The three terms on the right hand side of (10) correspond to the allocation, income, and internal effects, respectively.

For estimation purposes, equation (11) must be approximated. The convention used in the empirical exercises was to evaluate the expression at the middle points.

### **2.3 LEVEL, INEQUALITY AND THE INDICATOR OF STEEPNESS OF THE INCOME PROFILES IN EDUCATIONAL LEVEL**

Ramos (1990) used three synthetic measures for the indicators  $m_t$  (average schooling),  $i_t$  (schooling inequality), and  $s_t$  (income profile), based directly on the definition of the Theil Index.

The calculations of the principal parameters  $\mathbf{a}_g$ ,  $\mathbf{b}_g$ , and  $T_g$  (see equation 5), could determine the changes in the distribution by level of education ( $g$  groups in this category). These parameters allow us to analyze the trend in educational income differentials, the distribution of the population in each educational level and the inequality among them.

Three synthetic measures are used to summarize the changes related to education:

- $m_t$  is the average level of schooling for the year  $t$ .
- $i_t$  is the degree of inequality in the distribution of education for year  $t$ .
- $s_t$  is the variations in the income ratios associated with education for year  $t$ .

These measures can be calculated as follows:

$$m_t = \sum_g \mathbf{a}_g^* \mathbf{b}_g^t$$

$$i_t = \frac{\sum_g \mathbf{a}_g^* \mathbf{b}_g^t \log(\mathbf{a}_g^*)}{\sum_g \mathbf{a}_g^* \mathbf{b}_g^t} - \log\left(\sum_g \mathbf{a}_g^* \mathbf{b}_g^t\right)$$

$$s_t = \frac{\sum_g \mathbf{a}_g^t \mathbf{b}_g^* \log(\mathbf{a}_g^t)}{\sum_g \mathbf{a}_g^t \mathbf{b}_g^*} - \log\left(\sum_g \mathbf{a}_g^t \mathbf{b}_g^*\right)$$

where:

- $\mathbf{a}_g^*$  is the standardized income of educational category  $g$  for the reference year
- $\mathbf{b}_g^t$  is the fraction of the labor force in the  $g$ -th educational category in year  $t$
- $\mathbf{b}_g^*$  is the value  $\mathbf{b}_g$  in the reference year.

$s_t$  can be understood as an indicator of the relative steepness of the income profiles related to education. If one fixes the fraction of the labor force in each educational group, it follows that the steeper the income profile the larger the between group inequality.

$i_t$  corresponds to the Theil T index that would prevail in a population with no inequality within the educational groups, and where the group incomes were proportional to the group average incomes in the base year.

## 2.4 DECOMPOSITION ANALYSIS METHODS

The decomposition analysis is a useful tool for assessing the impact of certain factors on the evolution of the income distribution. In general, the different decomposition methods follow two definitions (Fields, 1996):

- Inequality in the population can be decomposed into different elements such as the sum of the parts is equal to total inequality;
- Inequality in the population can be decomposed as a weight sum of inequality within and between groups.

The Gary S. Fields (1996) and Bourguignon, et al (1998) papers employed the first decomposition method. Fields decomposed total population inequality in a sum of different variables or elements, each being the explanatory variable in the earnings function. This will help us to answer two questions: how much income inequality is explained by each right hand side variable in a given point in time? And, how much of the difference in inequality between groups or dates is explained by each variable? Notice that this technique assumes that we know the correct model specification.

Formally, the above methodology can be written as

$$Y = Z'B$$

where:

$Y = \ln(W)$  is the vector of the logarithm incomes  
 $Z = (1, X_1, \dots, X_J, \hat{a})$  is the matrix of explanatory variables and error term  
 $B = (\mathbf{a}, \beta_1, \dots, \beta_J, 1)'$  is the regression coefficient vector.

Then,

$$s_j = \frac{\text{cov}(\mathbf{b}_j Z_j, Y)}{\mathbf{s}^2(Y)} = \frac{\mathbf{b}_j \mathbf{s}(Z_j) \text{corr}(Z_j, Y)}{\mathbf{s}(Y)}$$

where:

$s_j$  is the relative factor weight and  $\sum s_j = R^2$  (determination coefficient)

The contribution of factor  $j$  to the change in the inequality measure  $I(\cdot)$  between time 0 and time 1 is:

$$\Delta_j [I(\cdot)] = \frac{s'_j I'(\cdot) - s_j I(\cdot)}{I'(\cdot) - I(\cdot)}$$

where:

$s_j$  is the relative weighted factor for year 0  
 $s'_j$  is the relative weighted factor for year 1.

Fields also proposes a change break down in the factor's contribution into the following: the change in the coefficient of the factor or variable, the change of the standard deviation of the variable and the change in the correlation between the variable and earnings.

Bourguignon et al. (1998) carried out a decomposition of the effects of changes in an entire distribution, rather than on a scalar summary statistic. This methodology was originally proposed by Almeida dos Reis and Paes de Barros (1991) and Juhn, Murphy and Pierce (1993) and later generalized by Bourguignon et. al.

The methodology, by means of micro simulations, decomposes the changes in income distribution into different effects. Bouillon, et. al. (1998) used this technique in the case of Mexico decomposing the change into the return effect, the population effect, the error term effect and the residual effect.

This can be expressed as follows, let  $D(y) = D(\beta, X, \mathbf{e})$  be the income distribution measure and define:

$$y = X\beta + \mathbf{e},$$

where:

$X$  is the set of demographic variables,

$\beta$  is the set of prices and

$e$  the error terms.

If  $y$  is the income in year 0 and  $y'$  in year 1, it can be show that the change in income distribution can be expressed as:

$$D = D(y') - D(y) = \beta(X', e') + X(\beta, e) + e(\beta', X') + \{e(\beta, X') - e(\beta', X')\}$$

where:

$$\beta(X', e') = D(\beta', X', e') - D(\beta, X', e') \quad \text{is the return effect}$$

$$X(\beta, e) = D(\beta, X', e) - D(\beta, X, e) \quad \text{is the population effect}$$

$$e(\beta', X') = D(\beta', X', e') - D(\beta', X', e) \quad \text{is the error term effect}$$

$$\{e(\beta, X') - e(\beta', X')\} \quad \text{is the residual effect}$$

Notice that the analysis makes the following assumptions:

- Income is correctly expressed as a linear combination;
- In order to compute  $D(\beta, X', e)$  the residuals in the second year are re-scaled to the second year of reference by a constant such that the variance in that year is the same as the variance of the residuals in the first year. This in turn implies the assumption that the distribution of  $e$  y  $e'$  just differs by the variance.

Cesar Bouillon, Arianna Legovini and Nora Lustig (1999) and Cesar Bouillon, Arianna Legovini and Nora Lustig (1998) used this methodology. In these documents, although the assumption of unchangeable dispersions, of the regression error terms, does not significantly restrict the model's results, it is questionable to use the variance instead of a proper inequality index. That means that one measure for the within inequality is used and another for the between inequality.

Miguel Székely (1995), in order to explain the inequality changes between two points in time applied the following formula:

$$C_B(\mathbf{p}) = \frac{T'_B(\mathbf{p}) - T_B(\mathbf{p})}{T' - T}$$

where:

$\mathbf{p}$  is the partition or division of the population

$T'_B(\mathbf{p})$  is the Theil index between group in year 1

$T_B(\mathbf{p})$  is the Theil index between group in year 0

$C_B(\mathbf{p})$  is the percentage of the change in inequality explained by the variables in  $\mathbf{p}$

$T'$  is the Theil index in year 1

$T$  is the Theil index in year 0.

It is important to note that this methodology does not allow us to separate the income from the allocation effect.

**ANNEX 3. EVOLUTION OF INEQUALITY**

**Table 1. Decomposition of Total Current Income**

<b>Income Source</b>	<b>Gini coefficient by income source</b>	<b>Share in Total income</b>	<b>Gini correlation with total income rankings</b>	<b>Contribution to Gini coefficient of total income</b>	<b>Percentage share in overall Gini</b>
1984					
Earnings	0.6428	0.4688	0.7249	0.2184	46.0
Monetary income excluding earnings	0.7568	0.3191	0.6470	0.1562	32.9
No monetary current income	0.6067	0.2120	0.7750	0.0997	21.0
Total	0.4744	1.0000	1.0000	0.4744	100.0
1989					
Earnings	0.6128	0.4635	0.7562	0.2148	41.0
Monetary income excluding earnings	0.8185	0.3109	0.7410	0.1886	36.0
No monetary current income	0.6541	0.2256	0.8187	0.1208	23.0
Total	0.5242	1.0000	1.0000	0.5242	100.0
1992					
Earnings	0.6440	0.4541	0.7790	0.2278	42.9
Monetary income excluding earnings	0.8129	0.2848	0.7316	0.1694	31.9
No monetary current income	0.6079	0.2611	0.8449	0.1341	25.2
Total	0.5313	1.0000	1.0000	0.5313	100.0
1994					
Earnings	0.6690	0.4932	0.8123	0.2680	50.2
Monetary income excluding earnings	0.7948	0.2550	0.6827	0.1384	25.9
No monetary current income	0.6051	0.2518	0.8365	0.1274	23.9
Total	0.5338	1.0000	1.0000	0.5338	100.0
1996					
Earnings	0.6514	0.4725	0.7870	0.2422	46.7
Monetary income excluding earnings	0.7924	0.2802	0.6884	0.1529	29.4
No monetary current income	0.6026	0.2472	0.8325	0.1240	23.9
Total	0.5192	1.0000	1.0000	0.5192	100.0

Source: Own estimates based on ENIGH

**Table 2. Pearson Correlation among explanatory variables**

<b>1988</b>				
	Education	Occupation	Econ. Sector	Status
Education	1.00			
Occupation	0.64	1.00		
Econ. Sect.	0.08	0.10	1.00	
Status	0.05	0.06	-0.04	1.00
Spearman's rho <sup>1/</sup>	0.58			

<b>1992</b>				
	Education	Occupation	Econ. Sector	Status
Education	1.00			
Occupation	0.63	1.00		
Econ. Sect.	0.06	0.02	1.00	
Status	0.08	0.08	-0.04	1.00
Spearman's rho <sup>1/</sup>	0.60			

<b>1997</b>				
	Education	Occupation	Econ. Sector	Status
Education	1.00			
Occupation	0.64	1.00		
Econ. Sect.	0.09	0.04	1.00	
Status	0.11	0.09	-0.06	1.00
Spearman's rho <sup>1/</sup>	0.62			

Source Own calculation based on ENEU Survey.

1/ It refers to Spearman's correlation between education and occupation.

**Table 3. Ratio of Income Share of the Highest 10 Percent to the Lowest 40 Percent Household Income Distribution**

<b>Low Income Countries<sup>1/</sup></b>		<b>High Income Countries<sup>1/</sup></b>		<b>Latin American Countries<sup>2/</sup></b>	
China	1.6	Australia	1.7	Argentina	2.8
Egypt	1.3	Belgium	1.0	Bolivia	3.6
India	1.4	Canada	1.4	Brazil	5.6
Ivory Coast	1.6	France	2.1	Chile	4.4
Kenya	4.7	Germany	1.3	Costa Rica	2.5
Madagascar	2.2	Italy	1.4	Ecuador	4.9
Nigeria	2.4	Japan	1.0	El Salvador	3.5
Pakistan	1.2	New Zealand	1.8	Mexico	4.4
Sri Lanka	1.1	Spain	1.0	Panama	4.9
Tanzania	1.7	Sweden	1.0	Paraguay	5.7
Uganda	2.0	Switzerland	1.8	Peru	2.6
Vietnam	1.5	United Kingdom	1.9	Uruguay	2.2
Zimbabwe	4.6	United States	1.6	Venezuela	2.7

Sources: <sup>1/</sup> World Development Report (1996).

<sup>2/</sup> IDB (1998).

**ANNEX 4. DECOMPOSITION RESULTS USING ENIGH DATA**

**Table 1 General Statistics for the Static Decomposition<sup>1/</sup>**

	1984		1989		1992		1994		1996	
	Between	Within	Between	Within	Between	Within	Between	Within	Between	Within
<b>Education<sup>2/</sup></b>										
Primary Incomplete	-0.10	0.07	-0.07	0.04	-0.08	0.06	-0.08	0.03	-0.07	0.03
Primary Complete	-0.03	0.08	-0.06	0.09	-0.08	0.05	-0.08	0.05	-0.07	0.04
L-Secondary Complete	0.04	0.05	-0.02	0.08	-0.02	0.08	-0.04	0.09	-0.04	0.08
U-Secondary Complete	0.07	0.04	0.07	0.07	0.10	0.08	0.09	0.08	0.09	0.08
University Complete	0.11	0.04	0.20	0.11	0.24	0.09	0.30	0.11	0.26	0.11
<b>Total</b>	<b>0.08</b>	<b>0.28</b>	<b>0.12</b>	<b>0.39</b>	<b>0.16</b>	<b>0.36</b>	<b>0.20</b>	<b>0.36</b>	<b>0.17</b>	<b>0.35</b>
<b>Position in Occupation<sup>2/</sup></b>										
Employee	0.01	0.25	-0.05	0.32	-0.03	0.33	0.00	0.41	0.00	0.36
Employer	0.05	0.04	0.08	0.08	0.13	0.10	0.06	0.06	0.07	0.08
Self-employed	-0.03	0.05	-0.01	0.09	-0.04	0.04	-0.04	0.06	-0.04	0.04
<b>Total</b>	<b>0.02</b>	<b>0.34</b>	<b>0.03</b>	<b>0.48</b>	<b>0.06</b>	<b>0.46</b>	<b>0.02</b>	<b>0.53</b>	<b>0.03</b>	<b>0.48</b>
<b>Sector<sup>2/</sup></b>										
Agriculture	-0.03	0.05	-0.02	0.02	-0.01	0.06	-0.02	0.02	-0.01	0.04
Manufacturing	0.00	0.06	-0.03	0.08	-0.01	0.11	-0.03	0.09	-0.02	0.09
Construction	-0.01	0.02	-0.01	0.07	-0.02	0.03	-0.02	0.03	-0.02	0.04
Commerce	0.01	0.07	0.02	0.12	-0.01	0.09	-0.02	0.08	-0.02	0.08
Services	0.02	0.10	0.03	0.17	0.04	0.17	0.09	0.23	0.06	0.19
Others	0.03	0.04	0.02	0.04	0.03	0.05	0.04	0.07	0.03	0.06
<b>Total</b>	<b>0.01</b>	<b>0.35</b>	<b>0.01</b>	<b>0.50</b>	<b>0.01</b>	<b>0.51</b>	<b>0.03</b>	<b>0.52</b>	<b>0.01</b>	<b>0.50</b>

Source: Own estimates based on ENIGH

1/ Labor force was limited to individuals who are: i) working as employee, employer or self employed; ii) between 12 and 65 years old; iii) living in urban areas; iv) working 20 hours or more per week; v) with positive income; vi) having the attributes of interest defined.

2/ Annex 1 presents how the group categories are defined

**Table 2. Synthetic Indicators of Schooling Distribution and Income Profile**

Year	1984	1989	1992	1994	1996
m <sup>t</sup>	0.468	0.525	0.516	0.527	0.538
i <sup>t</sup>	0.083	0.073	0.075	0.075	0.072
s <sup>t</sup>	0.083	0.124	0.161	0.203	0.175

Source: Own estimates based on ENIGH

1/ Labor force was limited to individuals who are: i) working as employee, employer or self employed; ii) between 12 and 65 years old; iii) living in urban areas; iv) working 20 hours or more per week; v) with positive income; vi) having the attributes of interest defined.

2/ Annex 1 presents how the group categories are defined

**Table 3. Results of the Dynamic Decomposition**

Period	Variable	Allocation	Income	Gross	Marginal
1984-1992	Education	<b>-9.8</b>	<b>52.1</b>	<b>42.4</b>	<b>26.6</b>
	Pos. in Occupation	21.1	18.2	39.3	13.8
	Sector	-17.7	2.5	-15.2	-19.5
1984-1994	Education	<b>-3.6</b>	<b>63.0</b>	<b>59.5</b>	<b>46.0</b>
	Pos. in Occupation	6.2	-1.2	5.1	-2.8
	Sector	-7.8	11.5	3.7	-11.4
1984-1996	Education	<b>-7.4</b>	<b>59.4</b>	<b>51.9</b>	<b>34.5</b>
	Pos. in Occupation	14.6	1.3	15.9	-9.7
	Sector	-27.2	3.7	-23.5	-15.1
1994-1996	Education	<b>5.7</b>	<b>68.4</b>	<b>74.1</b>	<b>76.6</b>
	Pos. in Occupation	-15.8	-12.4	-28.2	29.5
	Sector	21.8	25.0	46.8	21.0

Source: Own estimates based on ENIGH

1/ Labor force was limited to individuals who are: i) working as employee, employer or self employed; ii) between 12 and 65 years old; iii) living in urban areas; iv) working 20 hours or more per week; v) with positive income; vi) having the attributes of interest defined.

2/ Annex 1 presents how the group categories are defined

## REFERENCES

- Altimir, O, Piñera, S. (1982), "Análisis de descomposición de las desigualdades de ingreso en la América Latina". *El Trimestre Económico*, vol. 49, num. 196, pp. 813-860.
- Barros, Ricardo and José Guilherme Almeida Reis (1991), "Wage Inequality and the Distribution of Education: A Study of the Evolution of Regional Differences in Inequality in Metropolitan Brazil". *Journal of Development Economics*, 36: 117-143.
- Bouillon, Cesar, Arianna Legovini and Nora Lustig (1999), "Can Education Explain Income Inequality Changes in Mexico?" Mimeo.
- Bouillon, Cesar, Arianna Legovini and Nora Lustig (1998), "Rising Inequality in Mexico: Returns to Household Characteristics and the Chiapas Effect". Mimeo.
- Bouillon, Cesar, Arianna Legovini and Nora Lustig (1999), "Can Education Explain Income Inequality Changes in Mexico?" Mimeo.
- Bourguignon, Francois, Martin Fournier, and Marc Gurgand (1998), "Labor Incomes and Labor Supply in the Course of Taiwan's Development, 1979-1994". Mimeo.
- Bourguignon, Francois (1979) "Decomposable Income Inequality Measures". *Econometrica*, vol. 47, num. 4, pp. 901-920.
- Burfisher, Mary E., Robinson, Sherman, and Thierfelder (1993) "Wage Changes in a US-Mexico Free Trade Area: Migration Versus Stolper-Samuelson Effects". Working paper 645, Department of Agricultural and Resources Economics, University of California, Berkeley
- Cragg, Michael and Mario Epelbaum (1996), "Why has the wage dispersion grown in Mexico? Is it the incidence of reforms or the growing demand for skills? ". *Journal of Development Economics*, vol. 51, pp. 99-116.
- Cowell, F. (1980), "On the Structure of Additive Inequality Measures". *Review of Economic Studies*, vol. 47, pp. 521-531.
- De la Torre, Rodolfo (1997), "Desigualdad, Pobreza y Polarización Social en México". Mimeo.
- Elias, Victor J., 1992, "Sources of Growth: A Study of Seven Latin American Economies". International Center for Economic Growth.
- Feenstra, Robert C. and Gordon H. Hanson, 1996, "Globalization, Outsourcing, and Wage Inequality". *AEA Papers and Proceedings*, vol. 86, num. 2, pp. 240-45.
- Fields, Gary S. (1996), "Accounting for Income Inequality and Its Change". Cornell University. Paper prepared for presentation at the American Economic Association annual Meetings, New Orleans, L.A.
- Fiszbein (1991), "A Essays on labor markets and income inequality on less developed countries". Berkeley, Ph.D. Dissertation. University of California.
- Hanson, Gordon and Ann Harrison (1995), "Trade, Technology, and Wage Inequality". Working paper, num. 5110. National Bureau of Economic Research Working papers Series.
- Hanson, H. Gordon (1997), "Increasing Returns, Trade and the Regional Structure of Wages". *The Economic Journal*, vol. 107, num. 440, pp. 113-133.
- Hernandez Laos, Enrique, Nora Garro Bordonaro and Ignacio Llamas Huitrón (1997), "Productividad y Mercado de Trabajo en México", processed.
- Knight, J. T. and Sabot, R. H. (1983), "Educational Expansion and the Kuznets effect". *American Economic Review*, 73, 1132-6.
- Johnson, George E. (1997), "Changes in Earnings Inequality: The Role of Demand Shifts". *Journal of Economic Perspectives*, vol. 11, num. 2 (Spring), pp. 41-54.
- Juhn, Chihui, Kevin Murphy and Brooks Pierce (1993), "Wage Inequality and the Rise in Returns to Skill". *Journal of Political Economy*, vol. 3, num. 3, pp.410-448.
- Leibbrandt M. V., Christopher D. Woolard and Ingrid D. Woolard (1996), "The Contribution of Income Components to Income Inequality in South Africa". Living Standards Measurement Study, Working Paper No. 125. The World Bank.

- Londoño, Juan Luis (1996), "Poverty, Inequality and Human Capital Development in Latin America, 1950-2025". World Bank Latin American and Caribbean Studies, (June).
- Lustig, Nora and Miguel Székely (1996), "Mexico: Evolucion Economica, Pobreza y Desigualdad". Mimeo
- Moreno, A (1989), "La distribución del ingreso laboral urbano en Colombia: 1976-1988". Desarrollo y Sociedad, Bogotá, num. 24.
- Pánuco-Laguette, Humberto and Miguel Székely (1996), "Income Distribution and Poverty in Mexico", in Victor Bulmer-Thomas, editor, *The New Economic Model in Latin America and its Impact on Income Distribution and Poverty*, St. Martin's Press, New York, pp. 185-222.
- Psacharopoulos, G., Morley, S., Fiszbein, A., Lee, H., Wood, B. (1992), "Poverty and income distribution in Latin America: the story of the 1980s". The World Bank.
- Psacharopoulos, George (1993), "Returns to Investment in Education: A Global Update". World Bank Policy Research Paper No. 1067, (January).
- Ramos, Lauro (1990), "The Distribution of Earnings in Brazil: 1976-1985", Ph.D. Dissertation, University of California, Berkeley.
- Ramos, L., Trindade, C. (1991), "Educação e desigualdade de salários no Brasil: 1977/89". *Perspectivas da Economia Brasileira - 1992*. IPEA, Rio de Janeiro.
- Reis, J., Barros, R. (1989), "Um estudo da evolução das diferenças regionais da desigualdade no Brasil". Rio de Janeiro, IPEA.
- Reyes, A. (1988), "Evolución de la distribución del ingreso en Colombia". *Desarrollo y Sociedad*, num.21.
- Rodríguez, J. (1991), "Desigualdad salarial y educación en Lima Metropolitana: 1970 a 1984". Masters Dissertation, PUC/RJ. Rio de Janeiro.
- Schultz, T. Paul (1988), "Education Investments and Returns", in H. Chenery and T.N. Srinivasan, editors, *Handbook of Development Economics*, vol. I, Elsevier Science Publishers B.V., pp. 543-629.
- Shorrocks, A.F. (1980), "The Class of Additively Decomposable Inequality Measures". *Econometrica*, vol. 48, num. 3, pp. 613-625.
- Shorrocks, A.F. (1984), "Inequality Decomposition by Factor Components". *Econometrica*, vol. 52, num. 6, pp. 1369-1384.
- Shorrocks, A.F. and Foster, J.E. (1985), "Transfer Sensitive Inequality Measures". *Review of Economic Studies*, 5: 105-38.
- Shorrocks, A.F. and Mookherjee, D. (1982), "A decomposition Analysis of the Trend in UK Income Inequality". *The Economic Journal*, 92, pp. 886-902.
- Székely, Miguel (1995), "Aspectos de la desigualdad en México". *El Trimestre Económico*, vol. LXII (2), num. 246, pp. 201-243.
- "Symposium Issue on How International Exchange, Technology, and Institutions Affect Workers". *The World Bank Economic Review*, 1997, vol. 11, num. 1 (January).
- "Symposium on Wage Inequality". *Journal of Economic Perspectives*, 1997, vol. 11, num. 2 (Spring).
- Tan, Hong and Geeta Batra (1997), "Technology and Firm Size-Wage Differentials in Colombia, Mexico, and Taiwan (China)". *The World Bank Economic Review*, vol. 11, num. 1, pp.59-83.
- Vieira, M.L.P. (1998), "A relação entre educação e desigualdade no Brasil". Masters Dissertation, EPGE. Rio de Janeiro.
- World Development Report (1996). The World Bank. Oxford University Press.
- World Bank (1999). *Government Programs and Poverty in Mexico*, Green Cover.
- World Bank (1998). Country Economic Memorandum, "Factor Productivity and Growth".