Mexico in the 1990s: the Main Cross-Sectional Facts[*]

Chiara Binelli[†]and Orazio Attanasio[‡]

This draft: April 2009.

Abstract

We describe the main cross-sectional facts on individual and household earnings, labor supply, income, consumption and wealth in Mexico in the decade of the 1990s. We use two different data sources: the Mexican Employment Survey (ENEU) and the Mexican Income and Expenditure Survey (ENIGH). The contribution of the paper is twofold. First, we integrate the two surveys to provide a complete characterization of the changes in employment, wages, income, consumption and wealth in the 1990s. Second, we highlight some distinctive features that characterize Mexican economy in this decade. In particular, we focus on the increase in the number of informal workers in the mid 1990s and we study its relationship with the increase in wage inequality.

Key Words: Mexico, Inequality, Informality.
JEL Codes: J24, J31, O17.

1 Introduction

With a population of one hundred and four million people and per-capita GDP of around ten thousand dollars in the year 2005, Mexico is one of the countries with the highest Human Development Index (HDI). Following a steep improvement in all development indicators, Mexican HDI in 2005 scored above both the Latin American and the world average.

Notwithstanding these improvements, the distribution of income remains among the most unequal in the world. According to the CIA World Factbook, the Gini coefficient for household income in 2005 was 0.53 in Mexico, 0.45 in the US, 0.37 in the UK and 0.31 in the European Union.

The distribution of income was already skewed at the end of the 1980s, but it was in the decade of the 1990s that income inequality reached the extremely high level that we observe today. During this decade Mexico undertook a series of economic reforms that culminated in 1994 when it became a member of the Organization for Economic Cooperation and Development (OECD) and entered the North American Free Trade Agreement (NAFTA) with the US and Canada. In the same years Mexican economy was hit by a severe financial crisis that followed the peso devaluation of 1995.

In the mid 1990s all indicators for wage, income, consumption and wealth inequality increased. Also, there was a surge in the unemployment rate and in the number of workers that are not protected by labor legislation and are employed in informal firms. According to estimates based on the Mexican Employment Survey, the share of informal workers increased by around four percentage points between 1993 and 1996.

The contemporaneous shifts in the distribution of income, consumption and wealth motivates a comprehensive analysis of the changes in the level and dispersion of earnings, income, consumption and wealth across Mexican households in the 1990s.We use two different data sources: the Mexican Employment Survey (ENEU) and the Mexican Income and Expenditure Survey (ENIGH).

The paper is divided into three main parts. We start off by presenting the main cross-sectional facts about earnings, employment, income, consumption and wealth in the aggregate data and by relevant observable characteristics. We compare the trends in inequality across different variables and different indicators for the same variable in order to see which deciles of the distribution are driving the changes. When the data are available, we conduct the analysis separately for urban and for rural areas.

We find that wage inequality increased sharply in the first half of the 1990s and declined afterwards. While the rise in inequality is generalized across the distribution, the decline was concentrated in the top half. Only part of these dynamics is explained by the premium to observable characteristics that are important for wage determination, such as the premium to college education. The share of the variance of log wages that is due to the returns to unobservable characteristics is high and increasing over time.

After a steep rise in the first half of the 1990s, in the second half of the decade inequality in after-

tax household income remained almost constant. The trend in the aggregate measure was driven by a steep drop at the top and a sharp increase in inequality at the bottom of the distribution. Interestingly, the decline at the top was concentrated in urban areas while in rural areas income inequality kept on increasing both at the top and at the bottom of the distribution.

Differently from income, dispersion in log household consumption did increase throughout the decade, especially in the top half of the distribution.

Then, we estimate the changes in wage, earnings and consumption inequality over the life-cycle. We find strong evidence of time effects as opposed to cohort effects.

Finally, we highlight some distinctive features that characterize the Mexican labor market in this decade. In particular, we focus on the changes in the size of the informal sector and on its relationship with the changes in wage inequality. We find a significant positive correlation between the share of informal workers and the extent of wage inequality. The correlation is particularly strong in the mid 1990s, at the height of the pesos crisis.

The remainder of the paper is organized as follows. Section 2 describes the data used in the empirical analysis. Section 3 presents the first moments of the cross sectional distribution of wages, labor supply, income, consumption and wealth. Section 4 presents the second moments of the cross sectional distribution of wage, labor and asset income, consumption and wealth and it decomposes the trends in the overall dispersion of wages, income and consumption by relevant observable characteristics. Section 5 characterizes informal workers and documents the changes in the size of the informal sector in the 1990s. Section 6 studies the relationship between changes in informality and changes in wage inequality. Section 7 gives some concluding remarks. Appendix A presents summary statistics of the main variables of interest.

2 The data

To conduct the empirical analysis developed in this paper we make use of two different data sources: the Mexican Employment Survey, the ENEU (*Encuesta Nacional de Empleo Urbano*), and the Mexican Income and Expenditure Survey, the ENIGH (*Encuesta Nacional de Ingresos y Gastos de los Hogares*). Both Surveys are conducted by INEGI, the Mexican National Statistical Institute, but they differ significantly in coverage and structure.

The ENEU has a structure similar to the US Current Population Survey (CPS). Most importantly for the analysis conducted in this paper, for all individuals at least twelve years old the ENEU contains detailed employment information with several questions on individual's occupation status, type and characteristics of employment, characteristics and sector of main and secondary job, contract type, number of hours worked, monthly wages, unemployment status and duration, and social security contributions

paid by the worker's employer both in the private and in the public sector, which we will use to identify the workers employed in the formal and in the informal sector. We use the ENEU for the analysis of changes in the hourly wages, the number of hours worked and the employment rate. We consider all waves between 1987 and 2002.

The ENIGH has a similar structure to the Family Expenditure Survey (FES) in the UK. It is available for 1984, 1989, 1992 and every two years since then. The survey is representative at the national level and since 1992 it is also representative of rural and urban areas separately. The ENIGH considers the household as unit of observation. It is the only Mexican survey that has information on consumption, income and assets for several years.[1] It contains detailed information on assets and consumption for various categories of non-durable goods together with a wealth of demographic and labor supply variables, including wages and a detailed break down of income by source and type of generating activity. The latter are available for each income earner in the household.

The structure of the ENIGH did change significantly in 1992. Only from 1992 the surveys are strictly comparable in terms of sampling frame and methodology, timing and recall periods. The 1984 wave has a very different survey design and a significantly smaller sample size. The 1989 wave still has a sample that was less than half the size of the one in 1992. Also, before 1992 it was not possible to distinguish between rural and urban areas and there were no information on financial assets and wealth.

In order to allow for consistency and comparability of the results, our main sample considers all waves between 1992 and 2002. Despite the much smaller sample size, we will also make use of the 1989 wave to complete the descriptive analysis of the trends in income and consumption that we conduct on the aggregate data. For the years between 1992 and 2002 we work with two different versions of the sample for the ENIGH. The first one includes all data, the second one restricts the sample to urban areas only, which makes it comparable with the ENEU that is only representative at the urban level.

We use the ENIGH for the analysis of the dynamics of household consumption, wealth and income. We also use the ENIGH to obtain an alternative estimate of some variables of interest. In particular, we compute the hourly wages and we compare the results with the ones obtained from the ENEU.

In all empirical analysis we use the data from the fourth quarter of each yearly wave of the ENEU that is held practically at the same time than the ENIGH, it uses the same survey questionnaire to obtain information on wages and the same sampling frame and survey methodology as in the ENIGH. Recall periods for wages are also the same in the two surveys.

Mean and standard deviation of the main variables of interest in the ENEU and in the ENIGH samples are reported in Table 4 in Appendix A.

---

[1] Detailed information on assets, consumption and income is also collected by the Mexican Family Life Survey (MxFLS). The first survey was run in 2002 and it was followed by a second wave in 2005; two additional waves are scheduled for 2009 and 2012.

3 The cross-sectional distribution of earnings, labor supply, income, consumption and wealth: means

In this section we describe the key variables of interest and we document their changes over time. We focus on real hourly wages, hours worked, employment rate, labor and asset income, consumption and wealth.

We measure all labor market variables at the individual level and income, consumption and wealth at the household level. The sample used to compute hourly wages, hours worked and the employment rate comes from the fourth quarter of the ENEU considering all individuals aged between 25 and 60 that are actively working at the time of the interview.

The choice of using the ENEU as the primary source to compute earnings and hours of work is motivated by the higher quality data on labor income collected by the Mexican Employment Survey. Also, data quality on labor income is higher for urban than for rural areas and rural activities such as agricultural self-employment involve the use of own labour and capital simultaneously, which makes it difficult to obtain a measure of income from labour net of payments from physical capital.

However, the ENIGH has been used extensively in empirical studies on wage dynamics. Therefore, it is of empirical relevance to compare the evidence on wages across the two surveys. We compute real hourly wages for all individuals aged between 25 and 60 in each ENIGH wave.

Finally, we use the ENIGH to measure household income, consumption and wealth. We consider all households headed by an adult aged between 25 and 60. We define the head of household following the ENIGH definition, which states "The head of household is defined as the person recognized as such by the household members. A head of household is considered as absent if he/she is not living in the dwelling for reasons of work, study or other since at least three months at the moment of the interview; in that case, the head is not considered a household member and no information is collected for him/her."

In the ENIGH we can identify three main types of households: couples with or without children, extended and households composed by one person only. The first category represents the vast majority of all households. In each year between 1992 and 2002 couples with or without children are between sixty-five and seventy per cent of all families. Extended families are between twenty and twenty-five per cent and one-person households are between three and six per cent of the sample. We will investigate the role of the type of household to explain the variation in income and consumption inequality.

This section ends with a comparison between the employment to population ratio, the mean monthly earnings and the per capita consumption computed from the ENEU and the ENIGH and from the Mexican National Accounts data.
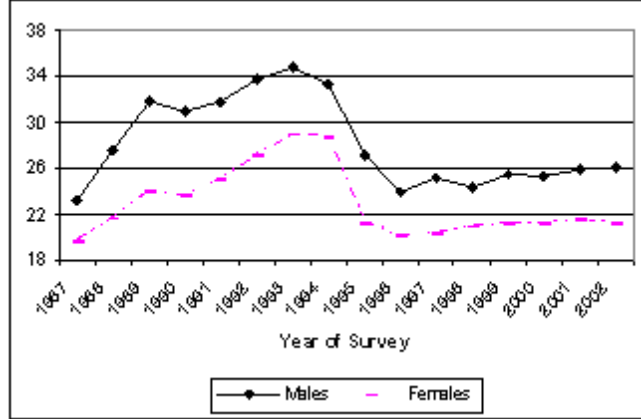
Figure 1: Mean Real Hourly Wages (Source: ENEU)

3.1 Hourly wage

The income measure that we use is the labor earnings in the primary occupation of all wage earners aged between 25 and 60. The hourly wage is computed as the ratio of monthly earnings and hours worked in the main occupation last month. We include all wage earners regardless of whether they are wage workers, self-employed, informal or formal workers. We deflate the wages using the Mexican national CPI. The values are expressed in June 2002 Mexican pesos.

We use the ENEU as the primary data source. Figure 1 presents the mean real hourly wage for males and females. As expected, for each year of the sample, wages are significantly lower for females. However, the two series follow a very similar trend. They both increase up to the mid 1990s reaching a peak in 1994 and decreasing sharply in 1995. Wages continue to decrease until the year 1996 when they start increasing again. Wage growth slows down in the year 2000: wages are almost flat between the year 2000 and 2002.

The wage trend is a reflection of the turbulent decade of the 1990s. The peso crisis of 1994 resulted into a massive devaluation of the national domestic currency. In December 1994, following a drain on its foreign exchange reserves, the peso was devalued by forty per cent. Between 1994 and 1996 Mexican GDP decreased by seven per cent a year. The international response to the crisis was immediate and assistance was promptly provided by the US and the IMF. As part of the rescue package, in March 1995 the Mexican government released a new economic plan to address the economic requirements set by the US and the IMF. The recovery was rather quick and by the end of 1995 Mexico had reentered the international capital markets.[2]

As an interesting comparison, Figure 2 presents the mean real hourly wages for male and female

[2]For a detailed description of the Mexican peso crisis, see Whitt (1996).
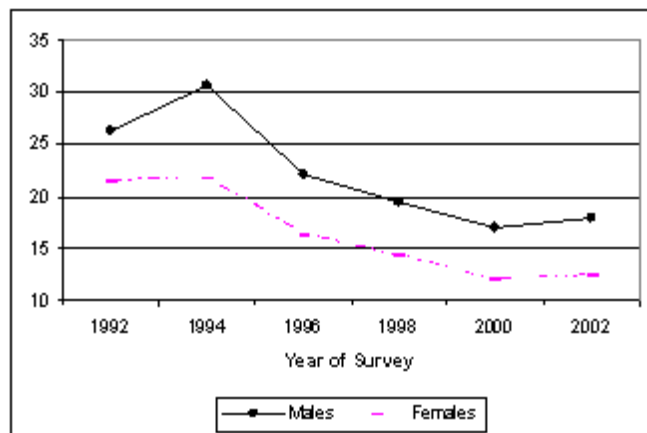
6

Figure 2: Mean Real Hourly Wages (Source: ENIGH, Urban Data)

workers aged between 25 and 60 computed using the urban sample of the ENIGH. The trend of the wages is similar to the one in the ENEU: the wages reach a peak in 1994, decrease afterwards and stabilize between the year 2000 and 2002. However, for each year of the sample mean wages in the ENIGH are significantly lower than in the ENEU with a gap of between ten and twenty per cent for each year between 1992 and 1998 and of around thirty per cent in 2000 and 2002.

One possible reason of discrepancy in the level of wages in the ENEU and in the ENIGH samples could be the different way in which the information on income is collected in the two Surveys. The ENIGH contains a detailed breakdown of all income sources including income from labor, entrepreneurial rents, interest income, property rents, transfers, and non-monetary income. In order to construct the measure of wage income, we use the survey questionnaire to identify the income obtained specifically as remuneration to labor earnings during the previous month and divide it by the number of self-reported worked hours. On the contrary, the ENEU does not distinguish between different sources of income.

3.2   Labor supply

We consider three measures of labor supply: the number of annual hours worked and the fraction of the working age population that works part time and full time. All measures are computed by considering all individuals that report a positive number of worked hours. We use the ENEU and we conduct the analysis separately for males and females.

Figure 3 presents the number of annual hours worked in the main occupation by males and females that are between the age of 25 and 60. For both males and females the series follows an increasing trend over time with a growth rate of around six per cent. The results have been obtained by considering all individuals that report a positive number of worked hours. As expected, the female curve lies well below
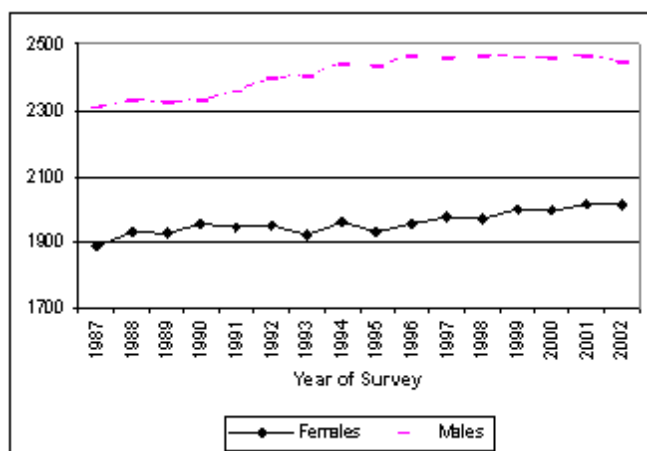
7

Figure 3: Yearly Hours Worked (Source: ENEU)

the male one: women do work on average four hundred and fifty hours per year less than men. For both males and females the average amount of annual worked hours is between ten and fifteen per cent higher than the corresponding figure commonly computed with US data.

Figure 4 and 5 present the fraction of the male and female adult population that is working full and part time. We define as working full time all individuals aged between 25 and 60 working at least thirty hours per week and as part time workers those reporting between one and twenty-nine hours of work per week.

The employment rate for males is at around ninety per cent for all years. Interestingly, it decreases by around two percentage points between 1994 and 1995 reaching the lowest value of around eighty-seven per cent in the years of the peso crisis.

In each year of the sample the employment rate for females is at a much lower level. However, consistently with the steep increase in the number of worked hours, it increases monotonically over time: from a value of thirty-five per cent in 1987, it reaches a value of around forty-six per cent in 2002.

Finally, as expected, many more females than males are employed in part time jobs. The proportion of adult females working part time is around twenty-seven per cent in 1987 and it reduces to approximately twenty per cent in 2002. The incidence of part time work is also decreasing for males: from a value of around seven per cent of the male working population in 1987, it drops to around five per cent in 2002.

For both males and females, the share of the adult population in part-time work increases by around two percentage points between 1994 and 1995. For males, this increase is offsetting the drop in the full-time work experienced in the turbulent years of the mid 1990s.
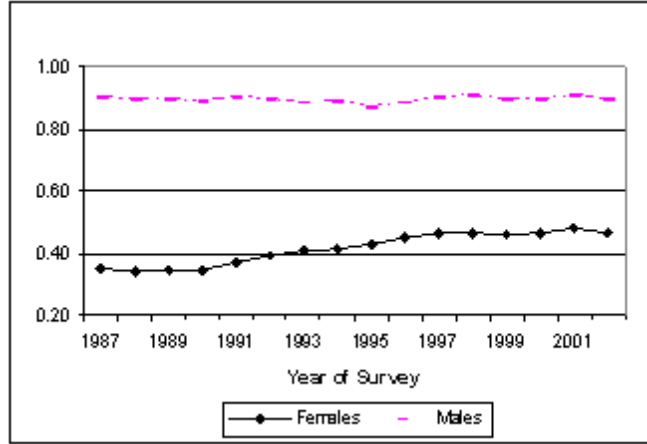
8

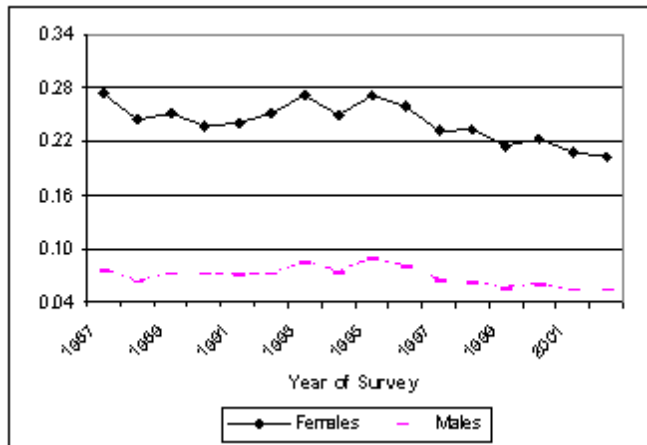Figure 4: Full Time Employment Rate (Source: ENEU)

Figure 5: Part Time Employment Rate (Source: ENEU)

3.3 Labor and asset income

We measure income at the household level and we exploit the richness of the information collected in the ENIGH to define different measures of both labor and asset income.

We consider two definitions of labor income. The first one is total labor income that is the sum of the earnings of all working household's members. The second one is total labor income plus private transfers received from all individuals that are residents in Mexico (alimony, child support, transfers from relatives, etc.) and income from retirement plans. Transfers received from individuals that live abroad are not included and will be considered separately. All measures use income net of taxes and contributions paid on labor earnings since the ENIGH (as well as the ENEU) does not report information on gross earnings. We then compute the equivalized version of both measures. To this purpose, we use the OECD equivalence scale that assigns a weight of one to the household's head, a weight of 0.7 to each additional individual aged 17 or older and a weight of 0.5 to each individual aged 16 or younger.[3]

We consider two measures of asset income. The first one is net financial income defined as the sum of dividends on stocks, interests on bonds and bank accounts net of the interest paid on household financial debt. For the self-employed we add the asset part of business income. The second one is net total asset income defined as the sum of net financial income and the net rents from all owned real estate property, that is the imputed rent for the owned primary residence and the rental income from any additional owned property. The sum of our second measure of labor income and net financial income gives a measure of total household income net of paid taxes.

The sum of our second measure of labor income and net financial income gives a measure of total after-tax household income. Ideally, we would like to define a measure of total disposable income by adding the amount of transfers received by the household from the government (unemployment insurance, social security benefits, welfare payments, etc.) net of paid taxes. The ENIGH does report information on income from public transfers and social programs. However, the information is difficult to compare between successive waves since the type of programs considered changed over time and there is no disaggregate information on the amount of income received from each program. Therefore, we use total after-tax income as our most comprehensive measure of household income and we present separately the amount of transfers received from the government. As discussed above, our measure of after-tax income only includes transfers from individuals that are residents in Mexico. We separately report the total amount of transfers received from abroad.

For each available year, graph A and B in Figure 6 present, respectively, the mean of the household equivalized after-tax labor income with and without income from private transfers in the aggregate data

---

[3]For a discussion on equivalence scales for Mexico see Rubalcava and Teruel (2004) and Rubalcava, Santana and Teruel (2005). In this paper we use the OECD scale that has been widely used for many developed countries, which allows for an easier cross-country comparability of the results.

and separately for urban and rural areas. Separate information on urban and rural areas is available from 1992; therefore, for 1989 we only report the statistics computed on the aggregate data. In the overall sample as well as in the urban and rural one the series for the two labor income measures do follow a very similar trend until the year 2000. They both decline with the exception of an increase between 1992 and 1994 in the urban sample and, as a reflection, in the aggregate data. Private transfers increase in the year 2002, which results into an increase of the second measure of labor income between the year 2000 and 2002.

Graph C in Figure 6 presents the average total amount of per household transfers received from the government and from abroad. Government transfers did not counteract the steep decline in household income and instead did follow a very path similar to the one observed for after-tax labor income. On the contrary, remittances do show a non-linear trend with a steep jump in the years of the peso crisis: in the mid 1990s migrants did give more help to their families facing a deep recession. In the years of the peso crisis both the number of households that received transfers from abroad and the average amount of remittances per household increased: between 1994 and 1996 the proportion of households with positive remittances increased by more than two percentage points and the proportion of remittances increased from around fifteen to over twenty-five per cent of total household income.[4]

Graph D in Figure 6 shows the evolution of equivalized net financial income in the aggregate data and separately for urban and rural areas. The information on net rents from all owned real estate property is only available for few households and, when available, it represents a small proportion of total income. Therefore, the series for net financial income and the one for total asset income do almost coincide. We only report the results for net financial income.

In the aggregate data equivalized net financial income decreases monotonically between 1992 and 1996 and it is almost constant since then. It follows a similar trend in both the urban and the rural sample until the year 2000 while it diverges in the last year of the sample when it increases (decreases) in the rural (urban) sample to reach a value of around two hundreds pesos per equivalent unit in both urban and rural areas.

---

[4] All results are available from the authors upon request.

Figure 6: Household Equivalized Labor Income, Government Transfers, Remittances and Net Financial Wealth (Source: ENIGH)
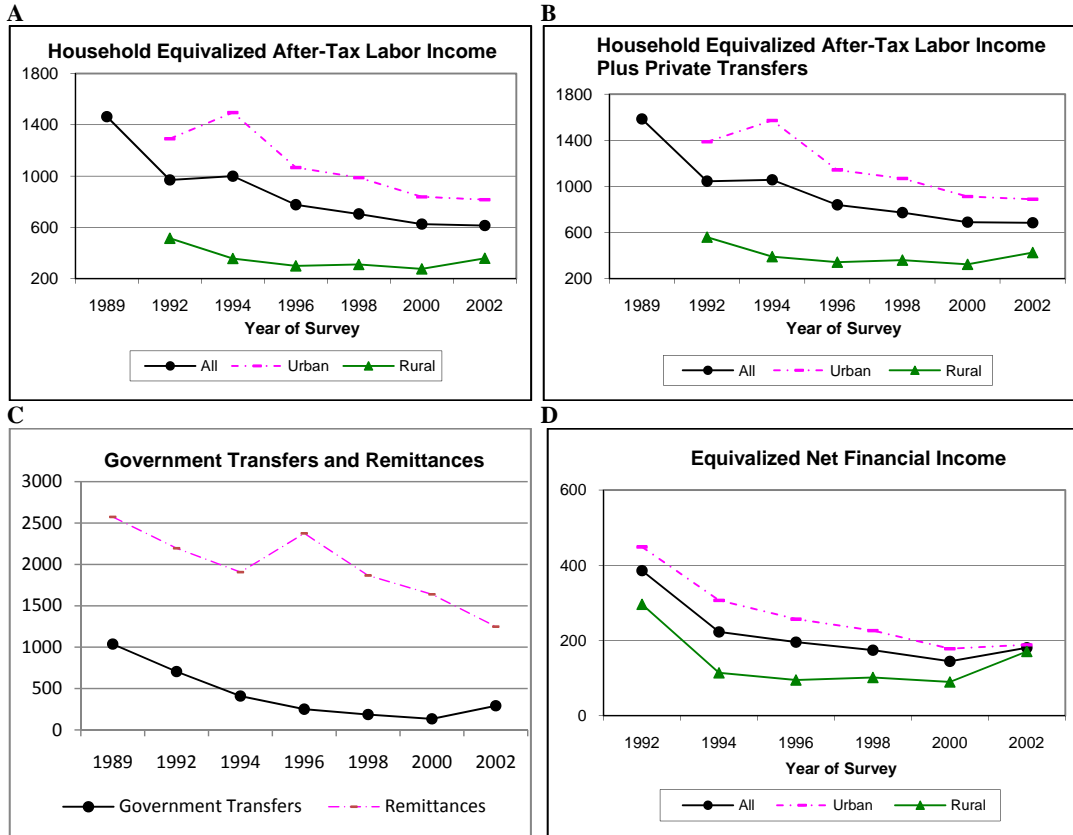
3.4 Consumption and wealth

As the CEX in the US and the FES in the UK, the ENIGH includes detailed questions on household's expenditures on several types of non-durable consumption goods. We consider two measures of non-durable consumption. The first measure is expenditure on non-durable goods such as food, alcohol, tobacco, personal care items, fuel, utilities and public services, public transportation, gasoline, apparel, entertainment, maintenance and repair of vehicles. We compute two different versions of this first measure, respectively, by first including and then excluding education and out of pocket health expenditures.

The second measure of consumption is given by non-durable expenditures plus the services from housing (rent paid for tenants, and imputed rent for homeowners). As we did for the first measure, we compute two different versions of this measure by including and by excluding education and out of pocket health expenditures.

We exclude all households reporting less than ten pesos of consumption expenditures per month. As we did for the wage data, we deflate each observation by the national Mexican CPI and we convert non-durable consumption into adult equivalent units by using the OECD equivalence scale.

Graph A in Figure 7 presents the results for the first measure of equivalized non-durable consumption when we include education and out of pocket health expenditures. The results do not change when we exclude education and health expenditures. Since the information on the services from housing is only available for very few households and, when available, it represents a very small proportion of total household consumption, the series for the second measure of consumption does follow a very similar trend. Therefore, we only report the results on the first measure of consumption.

The peso crisis of 1994 resulted into a significant decline in household consumption. Consumption decreased between 1992 and 1996 when it started increasing slightly. The decrease was concentrated in urban areas while in rural areas consumption remained essentially flat throughout the sample period.

As a measure of wealth we use total financial wealth within a household, that is the sum of financial wealth hold by all household's members. We compute two measures of household wealth: net financial wealth and net total wealth. Net financial wealth is defined as the sum of financial assets (such as checking/saving accounts, bonds, stocks, private pension funds and cash) net of liabilities (such as credit card debts and consumer loans). Net total wealth is defined as total financial wealth plus the market value of all owned residential real estate minus the value of any outstanding debt on mortgage and home equity lines.

As we did for the wage and consumption data, we convert both wealth measures into adult equivalent units by using the OECD equivalence scale. In the aggregate as well as in rural and urban areas the two series move very close together since the information on mortgages and home equity lines is only available for very few households. Therefore, we only report the results obtained for net financial wealth.

Graph B in Figure 7 presents the series for the aggregate data and separately for urban and rural areas. All series reach a peak in 1996, which is very pronounced in the urban sample. As it was the case for net financial income, they tend to converge to a common value in 2002.

As for wages, income and consumption, financial wealth shows the impact of the peso crisis. Interestingly, it is counter-cyclical with a peak in 1996. Financial wealth is the sum of investments in fixed-rate, low risk assets such as savings accounts and government bonds, and investments in risky assets, such as stocks and shares. Mexican households hold predominantly low-risk assets, which are driving the trend of net financial wealth. The amount invested in savings accounts and government bonds increased between 1994 and 1996 and declined afterwards following a sharp increase in the interest rate on fixed-rate investments in 1996 and a steep fall afterwards. Between 1994 and 1996 the three-month real interest rate on Mexican government bonds increased from around nine point five per cent to a value of more than forty per cent. In 1998 it was back to a value of around ten per cent.[5]

Graphs C and D in Figure 7 show the wealth-income ratios computed by dividing net financial wealth and net total wealth over total disposable income across all households in the sample. Consistently with the trends of net financial wealth, the two series exhibit a pronounced spike in 1996. In each year of the sample both series remain below one: this is a much smaller value that the one commonly computed with US data.

---

[5] The increase in financial wealth in 1996 that we find using the ENIGH data is confirmed by the evidence from macro data. The statistics reported in the "Report on Mexican Monetary Policy" that is released annually by the Bank of Mexico show that between 1994 and 1996 aggregate savings as a share of the national GDP increased by around one percentage point.

Figure 7: Household Equivalized Consumption, Financial Wealth and Wealth Income Ratios (Source: ENIGH)
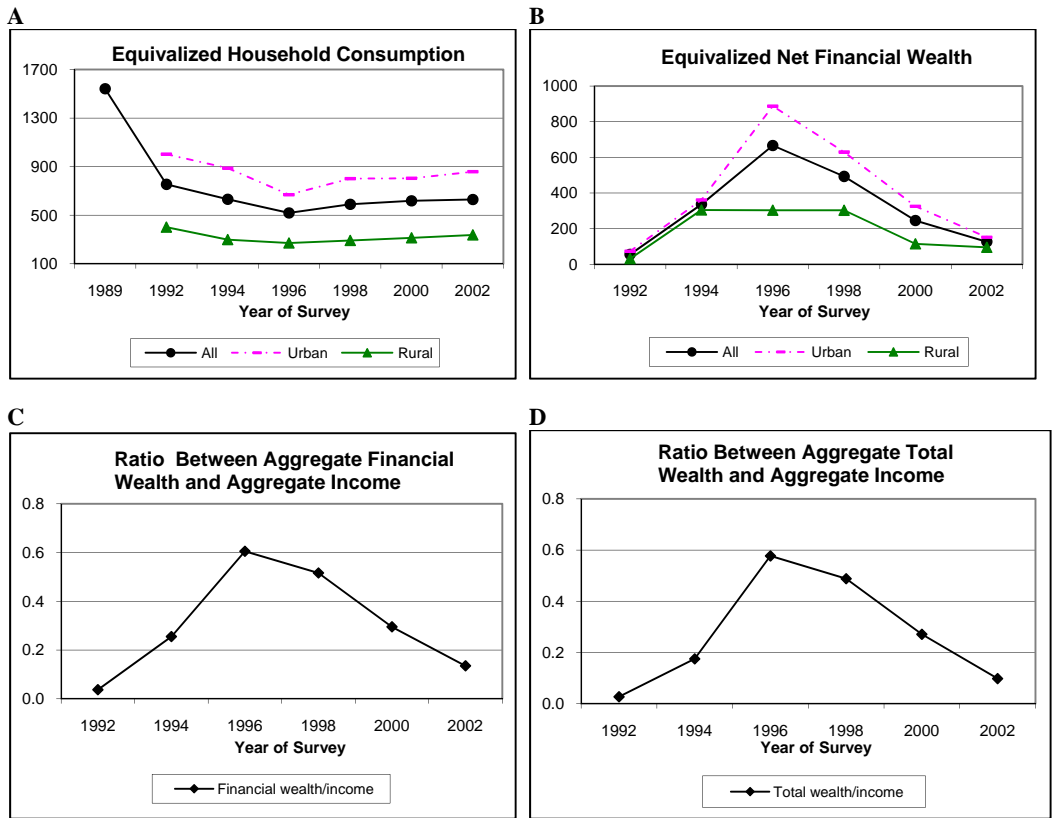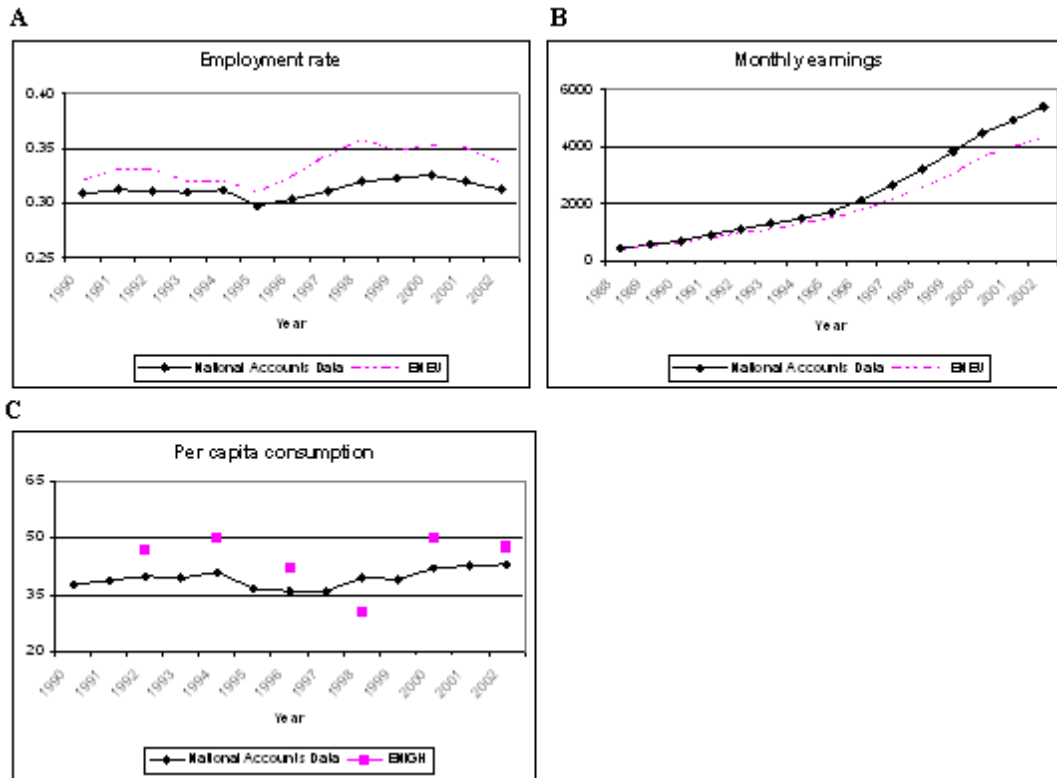
Figure 8: Comparison National Accounts with ENEU and ENIGH Data

3.5 National Accounts data

The graphs in Figure 8 compare three main variables computed from the ENEU and the ENIGH and from the Mexican National Accounts data. Graph A plots the employment to population ratio, graph B presents the mean monthly earnings[6] and graph C shows the mean per capita consumption for each available year.

The National Accounts data report the number of workers and the monthly earnings only for wage workers. Therefore, in order to make the macro and micro data comparable, we drop from the ENEU all observations on the self-employed. We then consider all individuals and households without applying any of the sample restriction that we used in the previous sections.

The employment rate is defined as the fraction of the working population over the total population and the mean per capita consumption as the ratio of total consumption divided by the total population. Since the first population Census is available for the year 2000, the series for employment and consumption start from this year. On the contrary, monthly earnings are directly reported in the National Accounts;

---

[6]We consider a measure of per capita earnings and not per capita income since the definition of earnings and salaries in the National Accounts data is directly comparable with the one in the ENEU.

therefore, the series for this variable starts from the first year for which the National Accounts data are available, which is 1988.

As shown in graph A in Figure 8, the employment rate follows a very similar trend in the ENEU and in the National Accounts. As for the levels, the employment rate tends to be overestimated in the ENEU, especially in the second half of the 1990s. Between 1990 and 1994 it is on average four per cent higher than the one computed from the National Accounts data. Between 1995 and 2002, it is on average ten per cent higher.

As shown in graph B in Figure 8, monthly earnings do also follow a very similar trend in the two data sources. However, as it was the case for the employment rate, there are some differences in the levels. On average over the sample period monthly earnings in the National Accounts are twenty per cent higher than the ones in the ENEU. One factor that could explain the differences is labor taxes. The ENEU report monthly earnings net of all labor taxes and social contributions paid in either public or private funds. On the contrary, the National Accounts report earnings net of social contributions but not of all labor taxes. Earnings are net of taxes paid directly by the employee but not of taxes paid either directly or indirectly by the employer.

As shown in graph C in Figure 8, per capita consumption follows a similar trend in the ENIGH and in the National Accounts even if the level of consumption tends to differ in the two data sources. In every year but 1998, the mean consumption value computed from the ENIGH is around twenty per cent higher than the one from the National Accounts; in 1998, the mean value computed from the ENIGH is around twenty per cent lower than the one that results from the aggregate statistics. The differences in the level of consumption could be due to measurement errors and misreporting. The ENIGH includes a very detailed list of consumption categories for non-durable goods, which could result into a more precise assessment of consumption and less measurement error than in the National Accounts data.

4 The cross-sectional distribution of earnings, labor supply, income, consumption and wealth: dispersion measures

In this section we study the changes in inequality for real hourly wages, labor and asset income, consumption and wealth. We compute different measures of cross-sectional dispersion: the variance of the logarithms, the Gini coefficient, the 90th-50th and the 50th-10th percentile ratios. We also decomposes the trends in the overall dispersion of wages, income and consumption by relevant observable characteristics and residual inequality.

In order to ensure comparability across different inequality measures, the statistics are computed on the same sample on which the variance of the logs is computed, that is the sample for which the zeros are excluded. We then assess the sensitivity of the results on the Gini coefficient and on the percentile

ratios by using the sample that includes the zeros.

The section ends with the estimation of the changes in wage, earnings and consumption inequality over the life-cycle.

4.1   Wages and hours of work

The graphs in Figure 9 present the evolution of the variance of the logarithm, the Gini coefficient, the 90th-50th and the 50th-10th percentile ratios of the hourly real wages for each ENEU wave between 1987 and 2002 for all individuals aged between 25 and 60. In order to assess the sensitivity of the results on the Gini coefficient and on the percentile ratios to the use of the sample that excludes the zeros, graphs B, C and D do also report the results obtained when the zeros are included.

As documented by several previous studies on wage inequality in Mexico[7], all inequality indicators increased sharply between the end of the 1980s and 1995 and did tend to decline afterwards. Between 1987 and 1995 the variance of the logarithm and the Gini coefficient for real hourly wages increased, respectively, by thirty-one per cent and by five percentage points. Between 1995 and 2002 the variance of the logarithm fell by around twenty per cent and the Gini coefficient decreased by around seven percentage points.

---

[7]See, among the others, Hanson and Harrison (1999), Robertson (2004), Airola and Juhn (2005) and Bosch and Manacorda (2008).

18

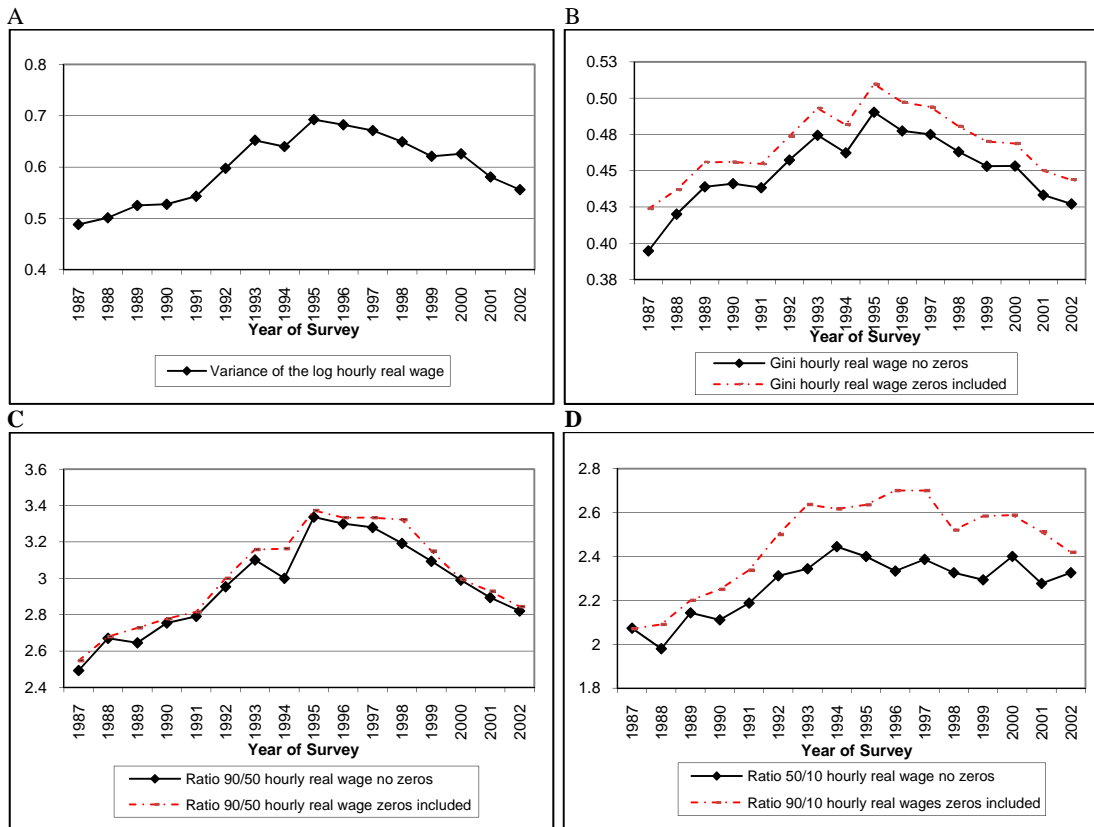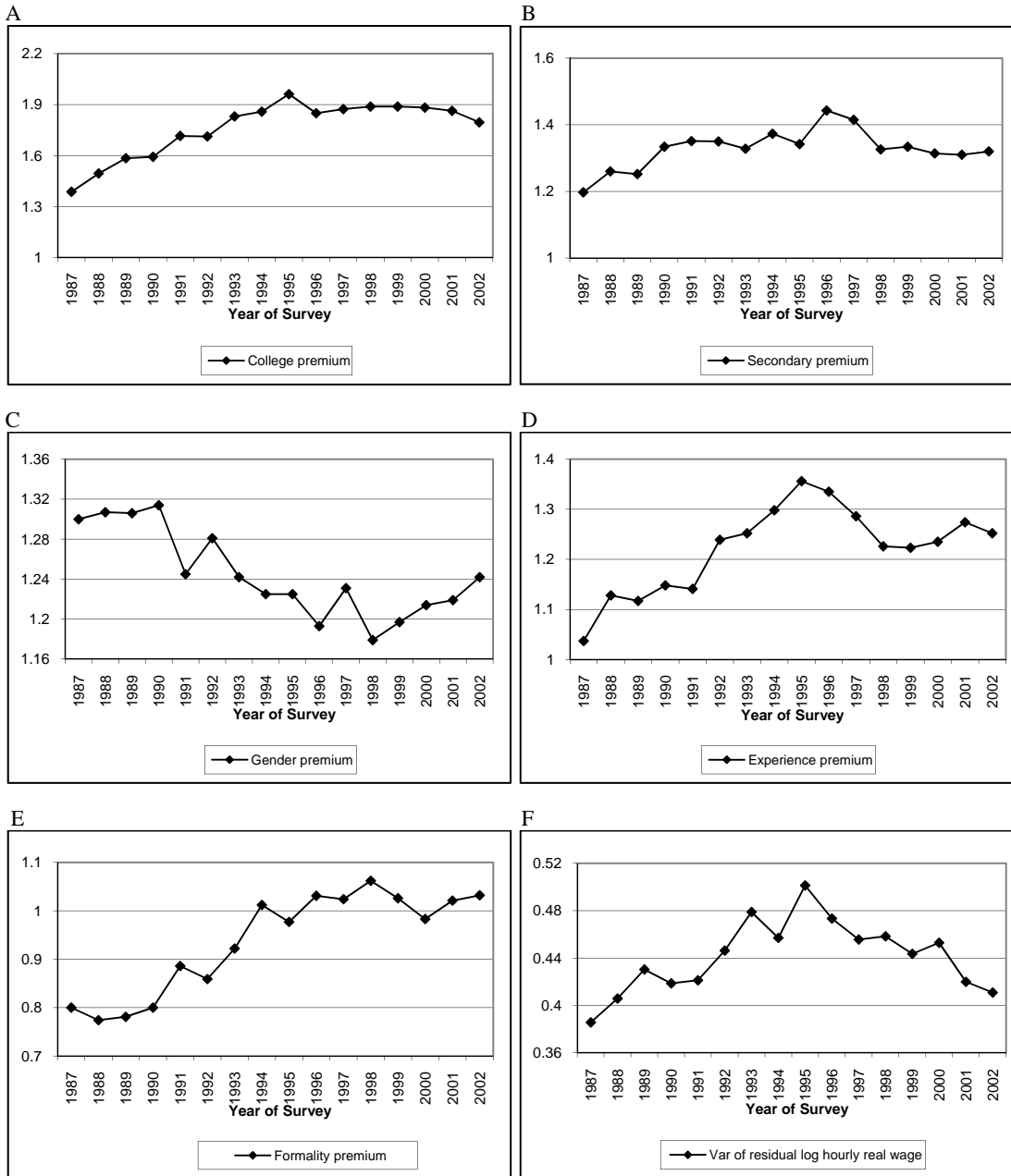Figure 9: Inequality Measures, Hourly Real Wage (Source: ENEU)

Figure 10: Returns to Education, Gender, Experience and Job Formality and Variance of Residuals of Log Hourly Real Wages (Source: ENEU)

Interestingly, the inequality decrease in the second half of the 1990s is driven by a decline in inequality at the top end of the wage distribution. As shown in graphs C and D in Figure 9, while the 50th-10th ratio of real hourly wages decreased only slightly in the second half of the 1990s, the 90th-50th declined sharply after 1995.[8] As expected, in each year of the sample, wage inequality is higher when we use the sample that includes the zeros. However, the trends remain the same: wage inequality increased in the first half of the 1990s and decreased afterwards due to a steep decline in wage dispersion at the top half of the distribution.

The decline in wage inequality between 1995 and 2000 measured with the 90th-10th ratio is more pronounced than the one measured with the variance of the logarithm.[9] This is evidence that most of the change in the variance of the log hourly wage in the second half of the 1990s is due to changes in the percentiles above the top ten per cent of the wage distribution, which suggests that inequality at the very top of the wage distribution could have declined even more than the 90th-10th and the 90th-50th ratios did.[10]

The first two graphs in Figure 10 present the premium to college and to secondary education for each year between 1987 and 2002. The college (secondary) premium has been computed as the average wage of males with competed college (secondary) or more and males with less than completed college (secondary) education.

As it was the case for all inequality indicators, returns to college are steeply increasing until 1995 and they decline afterwards. Between 1987 and 1995 they increased by over forty-three per cent; between 1995 and 2002 they decreased by around eight percentage points. The rise of the returns in the first half of the 1990s was of a much bigger magnitude than the fall in the second half of the 1990s. As a result, between 1990 and 2000 returns to college rose by around twenty per cent. In the same period, returns to secondary decreased by about two per cent.[11]

Consistently with returns to high levels of education being an important source of wage inequality, unreported graphs show that the changes in the 90th-10th ratio closely mirror changes in the returns to education: first, the 90th-10th gap increased much more for the college educated than for workers with secondary and less than secondary education. Within-group inequality is also substantially larger for more educated workers when using other inequality measures such as the variance of log wages. Second, returns

---

[8] This trend contrasts with the US experience where several papers have documented a decline in wage inequality in the 1990s which was driven by decreasing inequality in the lower tail of the wage distribution together with a continuining rise in inequality above the median. See, among others, Autor, Katz and Kearney (2005).

[9] For simplicity we do not report the results on the 90th-10th percentile ratios. They are available from the authors upon request.

[10] In order to investigate this question we would need uncensored data on income above the 90th percentile. Tax data would be ideal but they are not available to the general public.

[11] The double change in the wage differential between college and secondary and secondary and less than secondary education in the 1990s resulted into the "convexification" of the wage profile, which has been documented and discussed by Binelli (2008).

to college increased much more at the 90th percentile than at the median and at the 10th percentile of the wage distribution.

The increase in the returns to college in the first half of the 1990s coincided with a period of significant reforms that changed the structure of production and made Mexican economy more open to foreign investment. The reform effort culminated in 1994 when Mexico lowered its trade barriers by entering NAFTA with the US and Canada.

The increase in the skill premium at the time of trade opening contrasts with standard trade theory that would predict a reduction of the college premium when a country like Mexico, that specializes in the production of low-skilled intensive goods, opens to trade. One possible explanation that reconciles theory with empirical evidence is skill-biased technological change (SBTC): exposure to international markets increases the import of skill-intensive goods and induces a change in production towards skilled labour.

A vast empirical literature has found evidence of an increased demand for skilled labour in production associated with trade opening and a series of labor market reforms promoted in Latin America in the 1990s. Most empirical studies have found evidence of a technological change that increased the demand for skilled labour, while a smaller set of studies have estimated a positive impact of trade opening on the skill premium.[12] Riano (2008) finds that the trade-induced skilled-biased technology adoption can only account for about one sixth of the total increase of the skill premium observed in Mexico in the mid 1990s. The trade reforms took place in very turbulent years that were characterized by the peso crisis as well as by a series of economic reforms such as the implementation of a large scale privatization and financial deregulation programs, which could have contributed to the increase in the skill premium. Verhoogen (2008) finds the peso crisis more than the opening to trade being associated with the increased skill premium through an increase in the use of skilled labour and in the production of skill-intensive goods.

Together with the level of education, there are other dimensions of observable heterogeneity that represent important sources of wage dispersion. We focus on the premium to being a male worker and to having labor market experience. In addition, we compute the premium to working in the formal sector. The size of the informal sector and its increase in the mid 1990s suggests that it could be an important factor to explain the evolution of wage inequality in this decade. In section 6 we will discuss the relationship between informality and wage inequality.

We compute the gender premium as the ratio between the average wage of males and females aged between 25 and 60 and the experience premium as the ratio between the average wage of males aged

---

[12]See, among the others, Attanasio, Goldberg and Pavcnik (2002) for Colombia, Courseil and Muendler (2003), Pavcnik, Blom and Goldberg (2002) for Brazil, Bustos (2005) for Argentina, Cragg and Epelbaum (1996) and Hanson and Harrison (1999) for Mexico. Goldberg and Pavcnik (2004) and Winters and McCullach and McKay (2004) provide two complete surveys of the literature on Latin America.

between 45 and 55 and the average wage of males aged between 25 and 35. We define a worker as "informal" if he/she does not pay any social security contribution in either the private or the public sector and we compute the formality premium as the ratio between the median wage of formal and informal male workers aged between 25 and 60.

Graphs C, D and E in Figure 10 present, respectively, the evolution of the gender, experience and formality premium. The gender premium decreased by around nine per cent between 1987 and 1998 and it starts increasing again thereafter. In 2002 is back to the level that it had in 1993. In each year of the sample males earn more than two times what females do.

The experience premium was steeply increasing between 1987 and 1995 and it decreased thereafter. However, the post-1995 decline did not offset the 1987-1995 rise: overall, the premium to experience increased by around twenty per cent between 1987 and 2002. Older workers do earn more but they also experience higher wage dispersion: unreported graphs show that the within-group wage inequality is substantially higher for workers aged fifty or more with respect to younger workers. As shown in graph E, the formality premium is steeply increasing between the end of the 1980s and 1998 and it slightly declines thereafter.

Changes in observable variables account only for some of total dispersion in wages. For each year between 1987 and 2002, we consider all workers aged between 25 and 60 and we run an OLS regression of log hourly real wages on a quadratic polynomial in age as a proxy for labor market experience, gender, and dummy variables for the education level and the formality/informality nature of the main job.

Graph F in Figure 10 presents the variance of the residuals of this regression year by year. The share of the total variance of log wages that is unexplained is steeply increasing up to 1995 and it starts decreasing thereafter. The peso crisis of the mid 1990s resulted into an increased amount of the variance of log wages that is not due to observable characteristics. However, the increase is only temporary: in 2002 the unexplained share of the total variance is back to the value that it had at the beginning of the 1990s.

The variance of the residuals accounts for a large share of the total variation of log wages. A comparison between graph A in figure 9 and graph F in Figure 10 shows that the variance of the residuals accounts for an average of seventy per cent of the total variation of log hourly real wages.

Graph A in Figure 11 reports the variance of log hourly real wages for males and females. At the end of the 1980s the variance for females declines while the one for male increases. From the beginning of the 1990s the two series follow a very similar hump-shaped trend.

As does the variance of log wages, the variance of log yearly hours worked does also follow an hump-shaped profile. Graph B in Figure 11 presents the results for males and females. The two series do follow a very similar trend increasing up to 1995 and slightly declining thereafter. The variance of log hours in

A

Variance l of hourly real wages, ENEU

B

Variance log yearly hours, ENEU

C

Correlation Hours and Wages, ENEU
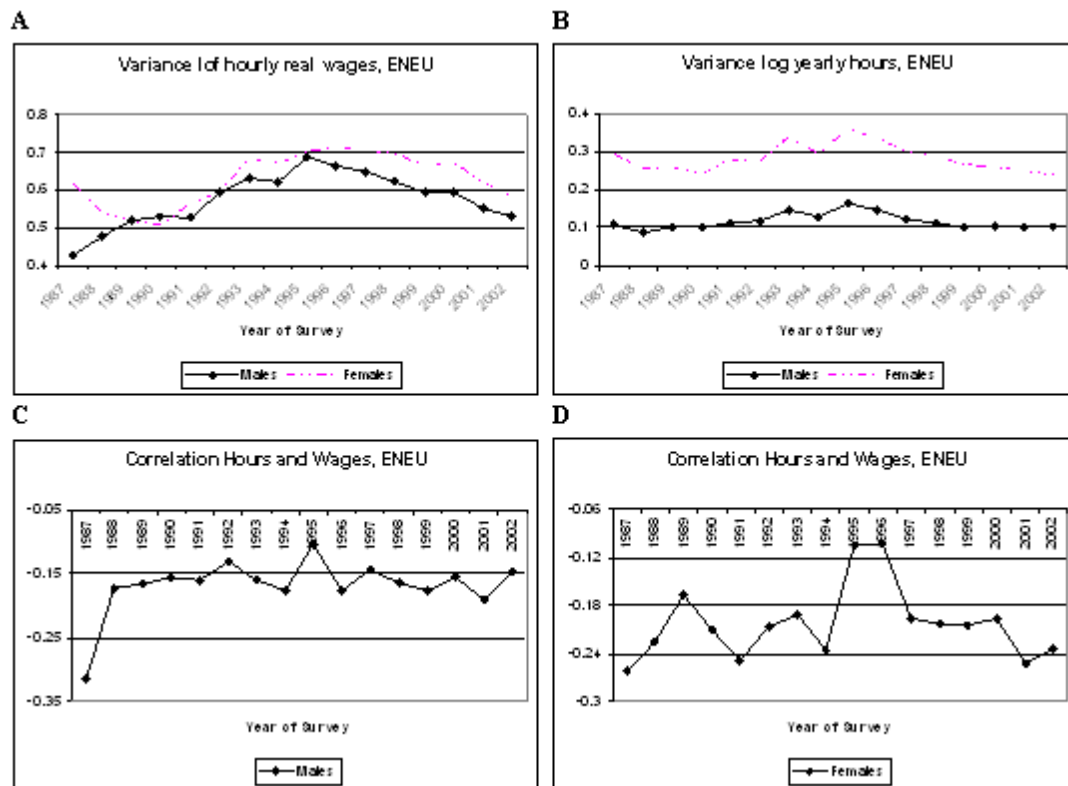
D

Correlation Hours and Wages, ENEU

Figure 11: Log Wages and Log Hours (Source: ENEU)

the female sample is around three times the size of the male variance.

Graph C and D in Figure 11 present the correlation coefficient between hours and wages in the female and in the male sample. In both samples and for all years the correlation is negative and tends to be very noisy, especially for females.

4.2 Labor and asset income

The graphs in Figure 12 present the variance of the logs, the Gini coefficient and the 90th-50th and 50th-10th percentile ratios for equivalized household labor income in the overall sample and separately for urban and rural areas.

All measures exhibit a spike in the years of the peso crisis: income inequality increased between 1992 and 1994 and decreased sharply between 1994 and 1996. Afterwards, inequality appears to be slightly increasing when measured with the variance of the logs while it decreases if the Gini coefficient is used.

There are significant differences between trends in income inequality in urban and rural areas. Looking at the Gini coefficient, the decrease in inequality between 1994 and 2002 observed in the aggregate data is driven entirely by the trend in urban areas; in rural areas, on the contrary, income inequality is increasing in these years. In the same way, the increasing trend of the variance of the logs observed in the overall sample between 1996 and 2002 is driven by the trend in rural areas while in urban areas income inequality tends to be constant or slightly decreasing.

As it was the case for hourly wages, the trend in aggregate inequality hides significant differences between changes in inequality in the upper and lower tail of the distribution. The 50th-10th ratio of household income rises between 1996 and 2002 while the 90th-50th declines monotonically from 1994 onwards. Income inequality increases (declines) below (above) the median of the distribution.

Interestingly, the evolution of income inequality for different percentiles of the distribution differs significantly in urban and in rural areas. While in rural areas after 1996 inequality increases both in the upper and in the lower tail of the distribution, in urban areas it decreases at both tails.

Urban and rural areas do differ in many respects. One is the sectorial composition due to a different mix of economic activities; two other factors are female labor force participation and the average education level of the workforce. A fourth very important factor is migration flows from rural to urban areas that affect the composition of the workforce in the two areas.

For each household the ENIGH reports information on the actual place of residence but there is no information on previous places of residence. Therefore, it is not possible to study migration flows. On the contrary, the ENEU reports information on whether the household was previously living in a different place and did migrated to the actual place of residence. As already discussed, the ENEU collects detailed information on wages for urban areas only. Therefore, we can compare the trends in wage inequality that

25

we obtain by using the sample of urban residents that include both households that did and that did not migrate from rural areas and the sample of households that live in urban areas and did not migrate from rural areas. Unreported results show that the trend in wage inequality for each of the two samples is very similar. Therefore, the decline in wage inequality, and most likely the one in income inequality given that income from wages is the main source of income for the vast majority of Mexican households, does not seem to be driven by changes in the composition of workers due to migration flows from rural to urban areas.

Together with a measure of equivalized household labor income, we define a measure of residual income. We compute it as the residuals of an OLS regression of the household log equivalized income on a quadratic polynomial in the age of the household head and dummy variables for the education level of the household head and her spouse, and for family composition (one-person household, couple without children, couple with children, non-couple households). We run an OLS regression year by year on the aggregate data as well as separately for the urban and the rural sample.

Graph A in Figure 13 presents the variance of the log raw, equivalized and residual labor income. The three series follow a very similar trend between 1989 and 1996 with a steep rise in 1994 and a sharp decrease in 1996. The share of the residual (unexplained) component out of the total variance of log income is at an average of eighty-five per cent throughout the sample period.

Figure 12: Inequality Measures, Equivalized Household Labor Income (Source: ENIGH)

A

Variance log equivalized household income

B

Gini coefficient equivalized household income

C

Ratio 90/50 equivalized household income

D

Ratio 50/10 equivalized household income
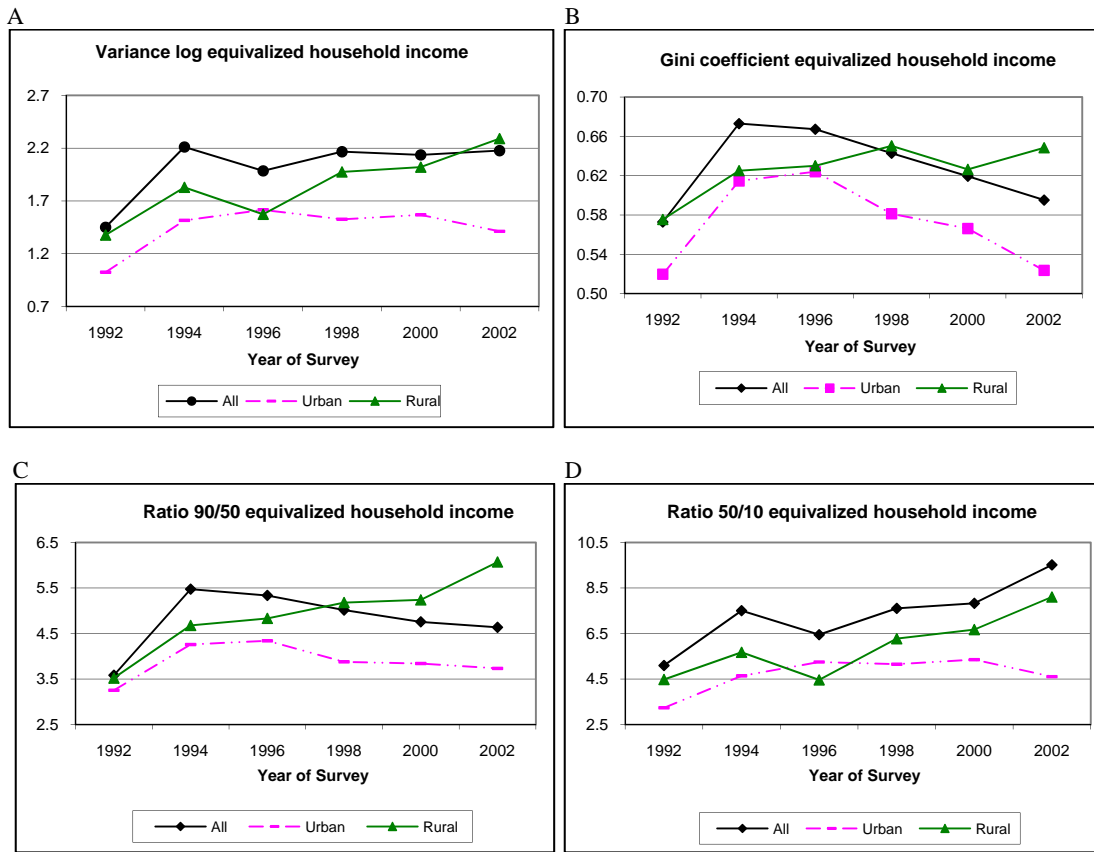
Figure 13: Decomposition Variance Log Equivalized Labor Income (Source: ENIGH)

A

Variance log labor income, All

Legend: Equivalized, Raw, Residual

B

Decomposition var log equivalized labor income, All

Legend: Age, Education, Household composition

C

Decomposition var log equivalized labor income, Urban

Legend: Age, Education, Household composition

D

Decomposition var log equivalized labor income, Rural

Legend: Age, Education, Household composition
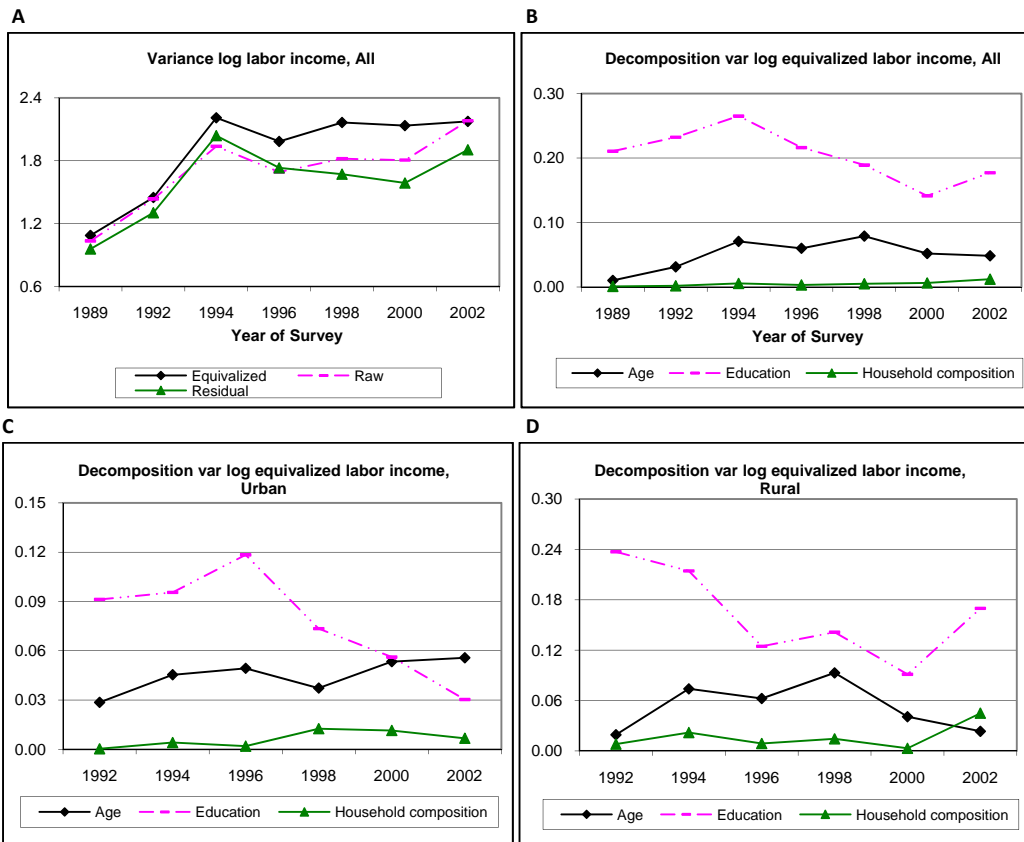
In order to disentangle the contribution of each of the observable variables to the changes in income inequality, for every year between 1989 and 2002 we compute the cross-sectional variance of the fitted values of the polynomial in age, the education and the household composition dummies.

Graph B in Figure 13 presents the contribution of each of the observable variables to the total variance of log equivalized household income. Among the observables, education is by far the variable with the highest explanatory power: on average, over the sample period, it accounts for seventy-six per cent of the explained variance. The contribution of age is at an average value of twenty-two per cent and the share of the explained variance accounted for by household composition is at an average of two per cent.

In the aggregate data the share of the total variance of log equivalized income accounted for by observables decreased from around thirty per cent in 1989 to fourteen per cent in 2002. Observables have a higher explanatory power in the rural sample: on average they account for thirteen per cent of the total variance of log income, while the average for the urban sample is at around nine per cent.

Graphs C and D in Figure 13 present the decomposition of the total variance by observables in urban and rural areas. As it was the case for the aggregate data, in both samples education accounts for the biggest share of the explained variance. The education contribution is at an average of fifty-eight and seventy per cent, respectively, in the urban and in the rural sample. Age explains around thirty-six and twenty-three per cent of the total explained variance, respectively, in urban and rural areas, and the average contribution of household composition is below eight per cent in both samples.

However, there are some differences in the trend of the variance accounted for by each of the observable components: in the urban sample the education contribution is at the highest level in 1996 and it declines afterwards, while in the rural sample it is at its highest level in 1992, it decreases afterwards and it increases again only in the year 2002. The age contribution increases from 1992 to 1996, decreases in 1998 and increases monotonically thereafter in the urban sample, while it increases up to 1998 and it decreases thereafter in the rural one.

The graphs in Figure 14 present the variance of log hourly wage and log earnings of the household head, log labor income, log labor income plus private transfers and log total after-tax household income for the overall sample and separately for urban and rural areas.

Figure 14: Variance Log Wages, Earnings and Income (Source: ENEU and ENIGH)
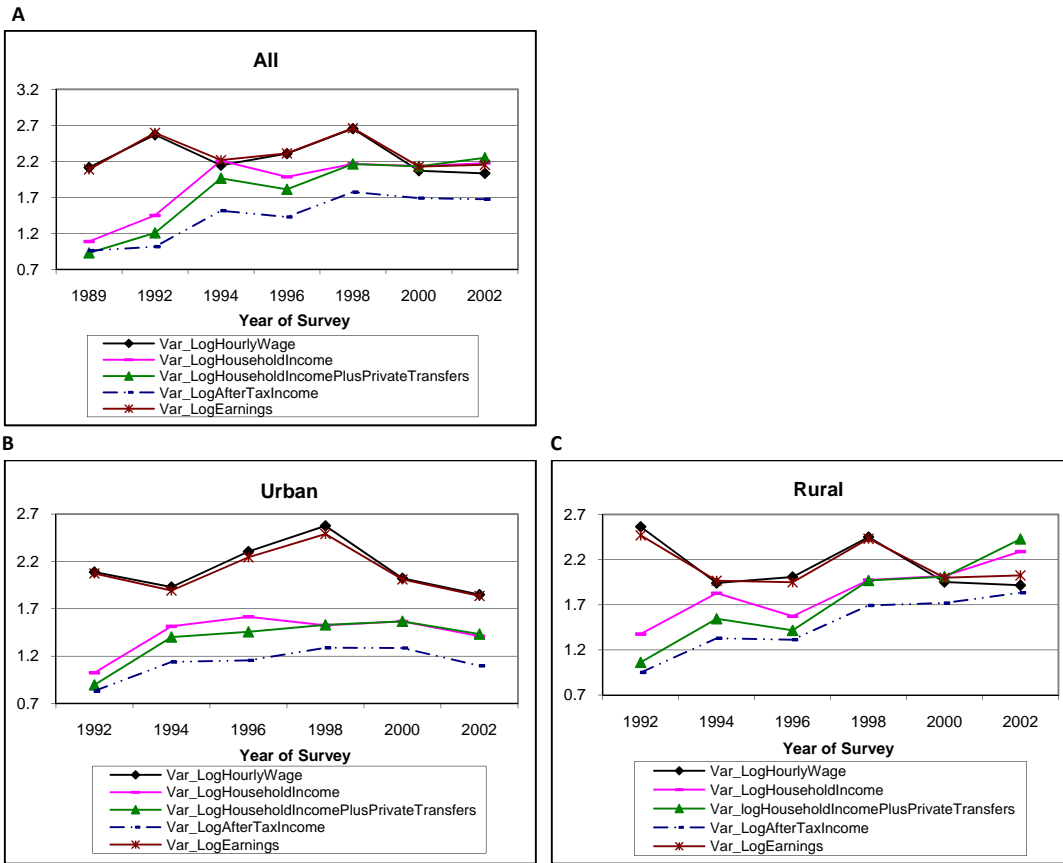
Both in the aggregate data and in the urban and rural sample, the extent of inequality measured with hourly wages or labor earnings is the highest and the extent measured with our most comprehensive definition of income, i.e. the after-tax household income, is the lowest. The extent of inequality measured with equivalized labor income plus private transfers is lower than the extent measured with labor income that does not include income from private transfers and it is higher than the extent of inequality measured with after-tax household income. This ranking suggests that some insurance/risk sharing mechanism among Mexican households could be in place.

4.3 Consumption

The graphs in Figure 15 present the variance of the logs, the Gini coefficient and the percentile ratios for equivalized household consumption. The results obtained by including or excluding education and out of pocket health expenditures look very similar. We present the results that include these two expenditure categories.

In the aggregate data as well as in both the urban and the rural sample consumption inequality increases significantly over time. Between 1992 and 2002 the variance of log consumption increased by over fifty per cent and the Gini coefficient increased by nine percentage points.

In the aggregate data and in the urban sample consumption inequality increases non monotonically with two peaks in 1994 and in 1998. The peso crisis did result into an initial increase and sudden decline in consumption inequality, which is then more than compensated by a steep inequality increase between 1996 and 1998.

On the contrary, rural areas did not experience the sharp increase in consumption inequality in 1994 and the steep decline in the years that followed the peso crisis. Instead, consumption inequality increased up to 1998, decreased in 2000 and it started increasing again in 2002.

Differently from what we documented in the previous section for aggregate labor income, in the second half of the 1990s consumption inequality did rise both at the top and at the bottom of the distribution in urban as well as in rural areas. As shown in graphs C and D in Figure 15, between 1996 and 2002 both the 50th-10th and the 90th-50th ratios do rise in all samples.

Figure 15: Inequality Measures, Equivalized Household Consumption (Source: ENIGH)

A

Variance log equivalized household consumption

B

Gini equivalized household consumption

C

Ratio 90/50 equivalized household consumption

D

Ratio 50/10 equivalized household consumption
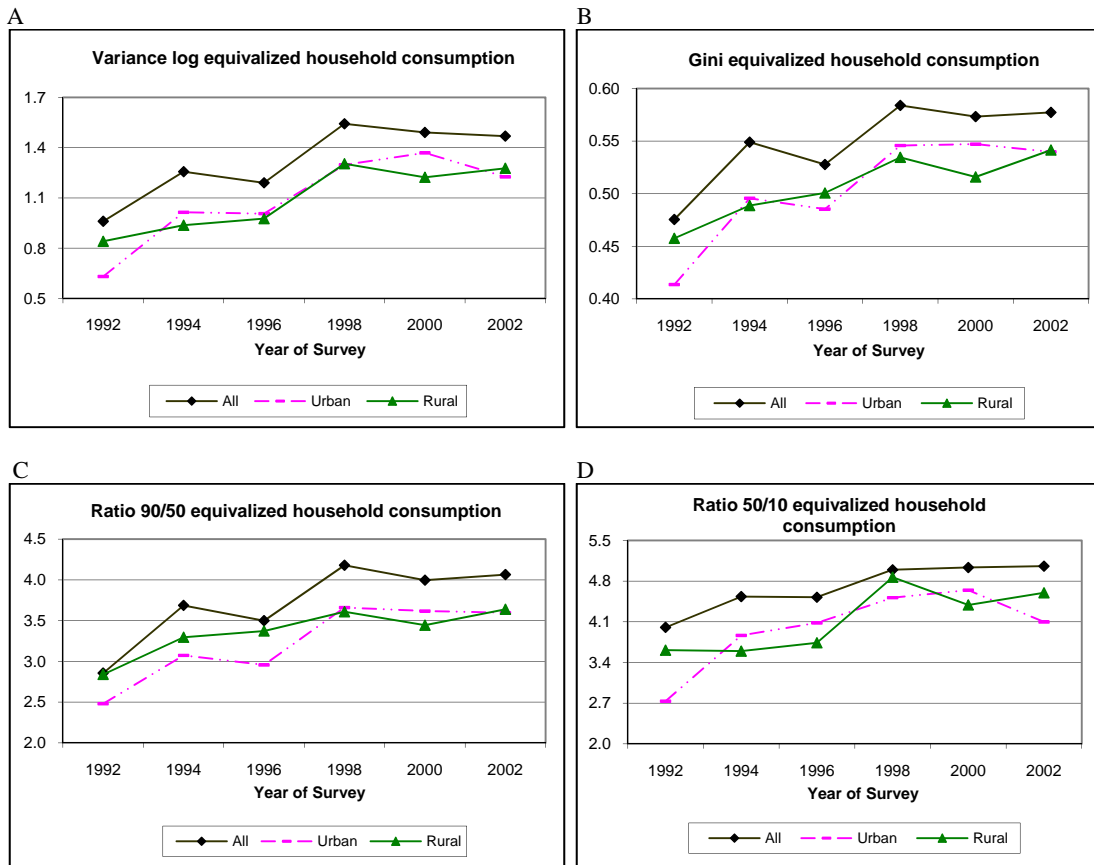
We apply to log equivalized household consumption the same decomposition that we did for log equivalized income. We compute a measure of the variance of log residual consumption by running an OLS regression of log equivalized consumption on the same set of controls that we used to decompose household income. We then assess the contribution of each of the observable variables to the changes in consumption inequality by computing the cross-sectional variance of the fitted values of the polynomial in age and the dummies for the education and the household composition variables. We run the regression for each year between 1992 and 2002 on the aggregate data as well as separately on the urban and rural sample.

Graphs A and B in Figure 16 present the variance of log raw, equivalized and residual consumption and the contribution of each of the observable variables to the total variance of log equivalized consumption in the aggregate data between 1989 and 2002. Graphs C and D present the decomposition of the variance of log equivalized consumption for the urban and the rural sample between 1992 and 2002.

Figure 16: Decomposition Variance Log Equivalized Consumption (Source: ENIGH)
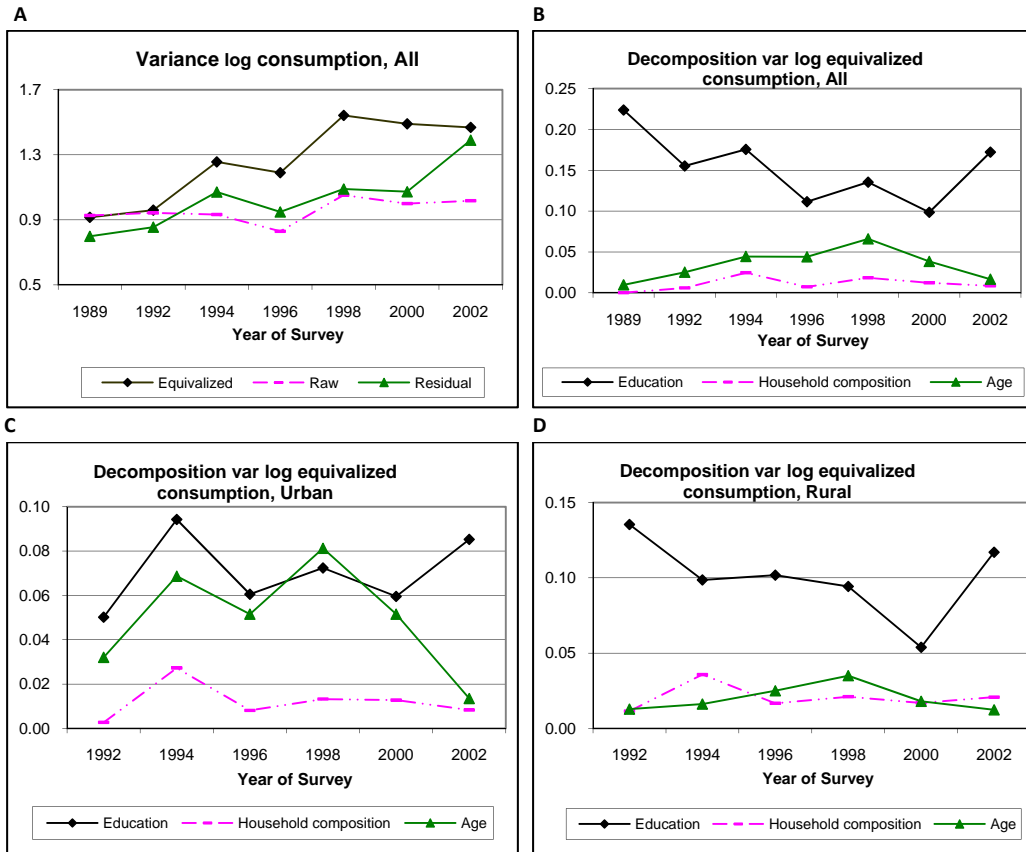
The variance of log consumption increases significantly over time with the biggest increase for the variance of log equivalized consumption. By looking at the decomposition by observable characteristics, as it was the case for log equivalized income, education accounts for the largest share of the explained variance. Over the sample period, it accounts on average for fifty-five and seventy per cent of the explained variance of consumption, respectively, in the urban and in the rural sample. The contribution of age is also important but mainly in the urban sample. For each year of the sample, the share of the explained variance accounted for by household composition is at an average of fifteen per cent in rural areas, while only at an average of seven per cent in the urban ones. Overall, between 1992 and 2002, the share of the total variance of log equivalized consumption accounted for by observables decreased from around thirteen to nine per cent in urban areas, and from around nineteen to twelve per cent in rural areas.

The graphs in Figure 17 compare the variance of the logs, the Gini coefficient and the 90th-50th and 50th-10th percentile ratios for equivalized household income and consumption in the aggregate data. As a sensitivity test, graph B does also report the results for the Gini obtained when the sample includes all the zeros. As it was the case for the Gini coefficient for hourly real wages, the Ginis are higher when the zeros are included. However, the trend remains the same with a steep increase in income inequality in the first half of the 1990s and a decline afterwards.[13]

All inequality indicators increase over time for both consumption and income but to a different extent. In particular, with the exception of the 50th-10th percentile ratio, between 1989 and 2002 consumption inequality did grow faster than income inequality. The evidence is even more stark if we look at Graph C in Figure 17: between 1994 and 2002 the 90th-50th percentile ratio decreased for income and increased for consumption. This is an interesting result which contrasts with the evidence found in other countries such as the US and Italy.[14]

Differences in income and consumption inequality reflect the way households react to temporary and permanent shocks to income. Consumption smoothing implies that permanent shocks that hit income over a long time span result into changes in consumption patterns, while temporary shocks that last for a short time or will be compensated by shocks of the opposite sign are largely absorbed by changes in savings decisions. Therefore, one possible explanation for a faster growth in consumption than in income is that the increase in income inequality that characterized Mexico in the 1990s was primarily of a permanent nature.

All inequality indicators for income and for consumption in Figure 17 show the impact of the peso crisis: they increase steeply in 1994 and decline in 1996. However, the size and the persistence of the impact was different for income and for consumption. First, inequality increased in 1994 much more for

---

[13]We have also computed the 90th-50th and the 50th-10th percentile ratios with the sample that includes the zeros. The results obtained are virtually the same as the ones obtained when the zeros are excluded.

[14]See Jappelli and Pistaferri (2008).

income than for consumption; second, between 1996 and 2002, income inequality increased only at the bottom of the distribution, while consumption inequality did increase both at the top and at the bottom of the distribution. The divergent trend of changes in income inequality at the top and at the bottom of the distribution is consistent with the trends that characterized the changes in wage inequality in the second half of the 1990s.

Figure 17: Inequality Measures Income and Consumption (Source: ENIGH)
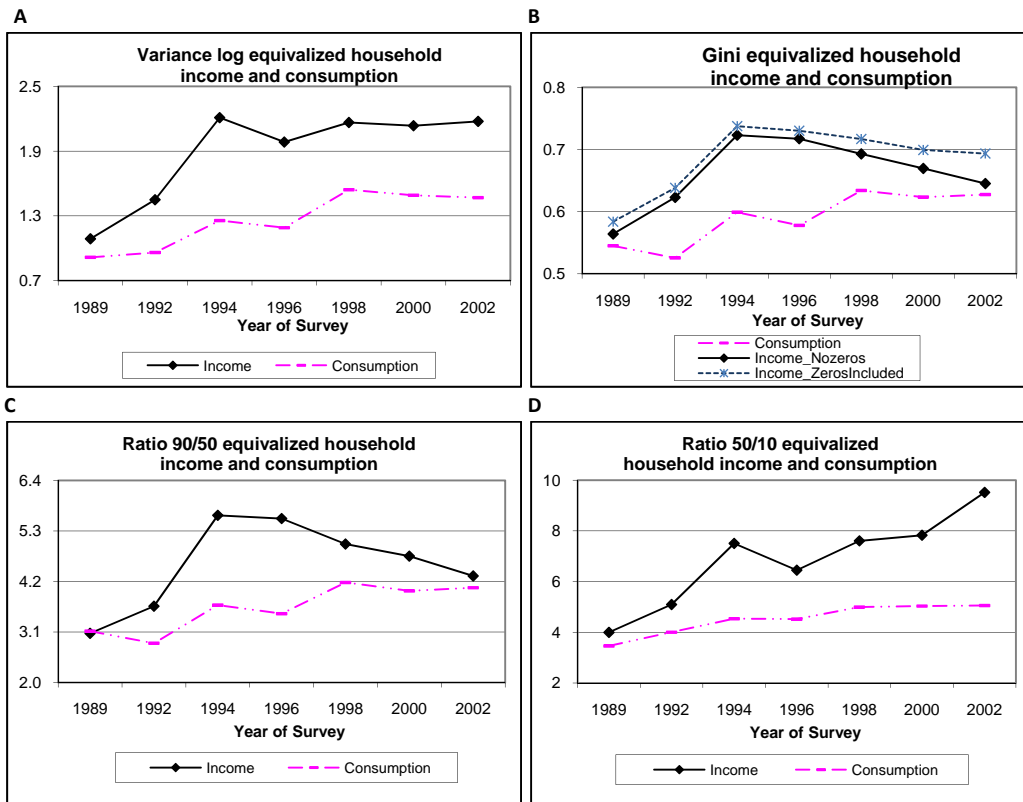
## 4.4 Wage, earnings and consumption inequality over the life cycle

Following Heathcote, Storesletten and Violante (2005) (HSV - hereafter), we exploit the cross-sectional variation between successive cohorts to estimate the evolution of wages, earnings and consumption inequality over the life-cycle. We do the analysis on the aggregate data and we use the variance of the log of the variable of interest as a measure of inequality.

We deal with the standard problem of lack of separate identification of time, cohort and age effects by doing the analysis in two different ways. First, we assume away cohort effects and we regress the cross-sectional moment of interest for age group a in year t on a full set of age group and year dummies. The age profile from the predicted age-portion of the age-time regression gives the life-cycle profile having controlled for time effects.

Second, we assume away time effects and we regress the cross-sectional moment of interest for age group a in cohort k on a full set of age group and cohort dummies. The age profile from the predicted age-portion of the age-cohort regression gives the life-cycle profile having controlled for cohort effects.

For each cross-sectional moment of interest, the graphs in Figure 18 present, respectively, the results obtained when we assume away cohort effects and we control for year effects (panels A, C, E and G) and, viceversa, when we abstract from time effects and we control for cohort effects (panels B, D, F and H). We use the ENEU for the analysis of hourly wages and the ENIGH for the analysis of earnings and consumption. All graphs are normalized to match the unconditional average for the corresponding cross-sectional moment over the sample period.

Figure 18: Variance Log Wages, Earnings and Consumption Over the Life Cycle (Source: ENEU and ENIGH)
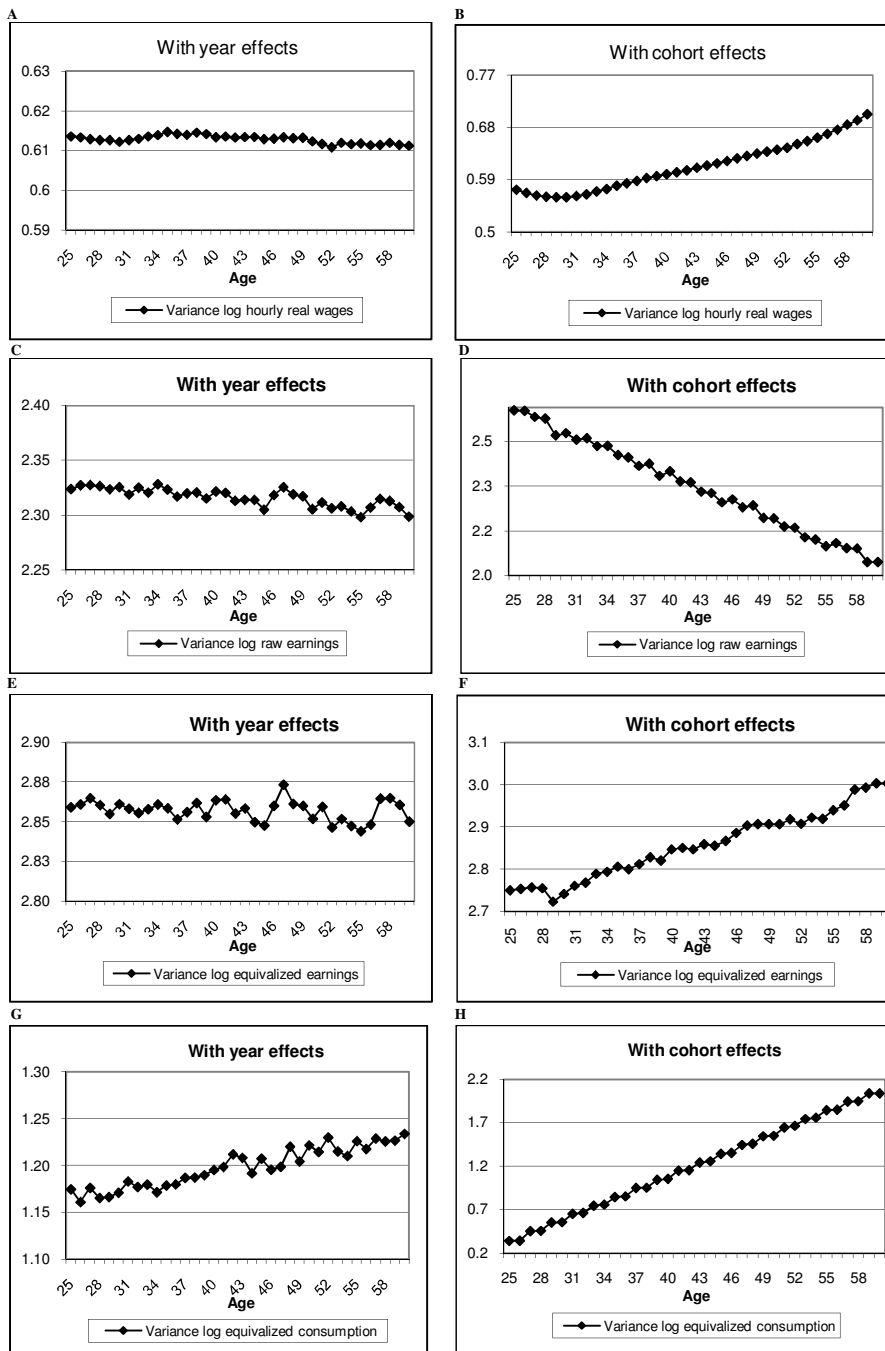
The age profiles of wages and earnings do differ significantly when we control for cohort or for time effects. The profile of log wages and log equivalized earnings exhibits a slightly (steeply) tendency to decrease (increase) when year (cohort) effects are accounted for. The profile of log raw earnings is decreasing in both Figure C and D but much more when cohort effects are accounted for. The variance of log equivalized consumption follows an increasing trend both in Figure G and H but of a much bigger magnitude when cohort effects are controlled for. All in all, as in HSV, we find that the changes in inequality over the life cycle are generally substantially larger when one controls for cohort effects.

We can measure the relative importance of cohort and time effects to explain the dynamics of wage inequality in Mexico between 1987 and 2002. We divide the sample in three main periods: 1987-1991, 1992-1996 and 1997-2002. Following HSV, in each period we compute three different measures of the changes in wage inequality: within-cohort changes given by the average (across birth cohorts) variance of log wages by year of birth, within-age changes given by the average (across age groups) variance of log wages by age and between-age changes given by the average (across age groups) variance of the within-period change in log wages between age groups. Table 1 reports the three different measures of the change in inequality for the variance of log hourly real wages.

As formally shown by HSV, the within-cohort (age) change can be expressed as the sum of the average age (cohort) and the average time effect and the between-age change can be expressed as the difference between the average age and the average cohort effect. Therefore, the within-cohort and the within-age change do share the average time effect and they do differ for the presence of a second term that is the average age effect for the within-cohort change and the average cohort effect for the within-age change.

Given this decomposition, we can evaluate the relative importance of time versus cohort effects with the following statistical test that provides a necessary and sufficient condition for the presence of significant time effects. If time effects are important and cohort effects are negligible, in each period the within-cohort and within-age group measures should have comparable magnitudes and the age and the cohort effect should be statistically significantly different (or, equivalently, the between-age change should be significantly different from zero).

As it is clear from the results in Table 1, in each of the three periods the magnitude of the within-age and within-cohort changes is very similar and the between-age change is statistically different from zero.

In addition, if time effects were not important, the within-cohort change should be similar across time periods and, conversely, if cohort effects were small, the within-period between-age change should be constant across periods.[15] We find that the within-cohort changes are statistically different between periods and we can not reject the hypothesis that the between-age changes are constant over time.

---

[15] These two additional hypothesis are only necessary conditions. They became necessary and sufficient conditions under the additional assumption that time and cohort effects are linear.

	1987-1991	1992-1996	1997-2002
Within cohort	0.5178	0.6541	0.6137
	(0.0270)	(0.0413)	(0.0488)
Within age group	0.5248	0.6814	0.6185
	(0.1030)	(0.1266)	(0.1340)
Between age group	0.5249	0.6771	0.6193
	(0.0983)	(0.1283)	(0.1382)

Table 1: Changes in the Variance of Log Hourly Real Wages by Age and Cohort, standard errors in parenthesis

Therefore, as also found by HSV for the US, we conclude that the evidence is consistent with the presence of important time effects and the absence of significant cohort effects. This conclusion suggests that exogenous shocks such as a skill-biased technological change could be important to explain the dynamics of wage inequality observed in Mexico in the 1990s, while changes in the composition on observable characteristics such as birth cohorts would be irrelevant. This is very consistent with the results found by Binelli (2008): she found that changes in the composition on observables can not explain the trend in wage inequality observed in Mexico in the 1990s, which was driven by changes in the supply of education in reaction to an increased demand for skilled labor in production.

5   Informality

In this section we present a description of the changes in the share of informal workers and their characteristics. We use the ENEU and we consider all individuals aged between 25 and 60. First, we characterize the changes in informality by gender, age, education and type of occupation. Second, we compare different measures of wage inequality for formal and informal workers.

The graphs in Figure 19 present the proportion of informal workers among the 25-60 age population, overall and by relevant observable characteristics as well as the formality premium by job category. As documented by previous studies,[16] the share of informal workers in the 1990s accounts for around half of the urban work force in Mexico: between 1987 and 2002 it increased only slightly from around fifty to fifty-two per cent. However, in 1995 it jumped up by four percentage points to reach the value of fifty six per cent in one year. In 1998 it was back to the pre-1995 level and it fluctuated around a value of between fifty and fifty two per cent until 2002.

With respect to observable characteristics, until the mid 1990s around seventy per cent of informal workers were male. By the year 2002 the proportion of female working informal has increased but it is still below forty per cent.

We define three age groups: the young group of all individuals aged between 25 and 35, the middle

---

[16]See, among the others, Maloney (2004) and Bosch and Maloney (2006).

group of those aged between 36 and 50 and the old group of those aged between 51 and 60. The middle group accounts for the highest share of informal workers increasing from forty seven per cent in 1987 to above fifty per cent in 2002. The proportion of old informal workers is stable at around twenty per cent throughout the sample period while the share of the young group slightly declines from thirty five to thirty per cent.

We define four education groups according to the highest completed schooling level - less than primary, primary, secondary and post secondary. As documented by previous studies,[17] the vast majority of informal workers have very low education levels. Graph D shows that over sixty per cent of informal workers in 1987 and still more than forty per cent in 2002 had less than completed primary education. The proportion of workers with completed primary, secondary or post-secondary education increased over time but remains at low levels. The proportion of those with primary increased from around fourteen to twenty-three per cent, while the share of workers with completed secondary and post-secondary education increased from around ten to sixteen per cent.

We distinguish between four types of informal jobs: self-employed (individuals working on their own with no employees at their dependency), employers (individuals working on their own with some employees at their dependency), wage and piece workers. As shown in graph E in Figure 19, in each year of the sample self-employment represents the largest share of the informal sector accounting for over forty per cent of the informal labor force. Informal salaried workers represent between thirty and thirty five per cent of informal workers. By adding the "employer" category, these three categories together account for almost ninety per cent of all informal workers. The proportion of piece workers is stable at around ten per cent throughout the sample period.

The job category is an important determinant of the formality wage premium. We group employers and self-employed in the self-employment category and wage and piece workers in the employees category. As graph F in Figure 19 shows, in each year of the sample, the formality premium for the self-employed is at a much higher level than the one for the employees. The ratio of the mean wage for formal and informal employees is significantly higher than one only from the year 1992 onwards.

The formality wage premium reflects differences in observable characteristics between formal and informal workers and in the returns to those characteristics. In order to disentangle the component of the premium that is due to differences in the distribution of individual characteristics and in the returns to these characteristics, we can compute an Oaxaca decomposition of the wage differential between formal and informal workers. We consider all workers aged between 25 and 60 and we run two separate OLS regressions for formal and for informal workers of the log hourly real wage on education dummies, a quadratic polynomial in age, gender, dummies for the four job categories (self-employed, employers, wage

---

[17]See, among the others, Maloney (1999).

and piece workers) and a constant. The results show that around eighty per cent of the wage differential between formal and informal workers is due to differences in the returns to the observable characteristics that we included in the regressions, while only twenty per cent is due to a different distribution of these observable characteristics between the two groups.

Together with the wage differential between formal and informal workers, it is interesting to compare inequality within each of these two groups. The graphs in Figure 20 present the variance of the logs, the Gini coefficient, the 90th-50th and the 50th-10th percentile ratios of the hourly real wage for formal and informal workers. All measures follow a similar trend over time in the two samples. However, for all indicators and in each year of the sample, inequality is at a much higher level among informal workers. The variance of the log wages for informal workers is more than double the size of the one for formal workers at the end of the 1980s and it is on average sixty per cent higher throughout the 1990s. The Gini coefficient is on average ten percentage points higher for informal workers.

Figure 19: Share Informal Workers by Observable Characteristics and Formality Premium by Job Type (Source: ENEU)
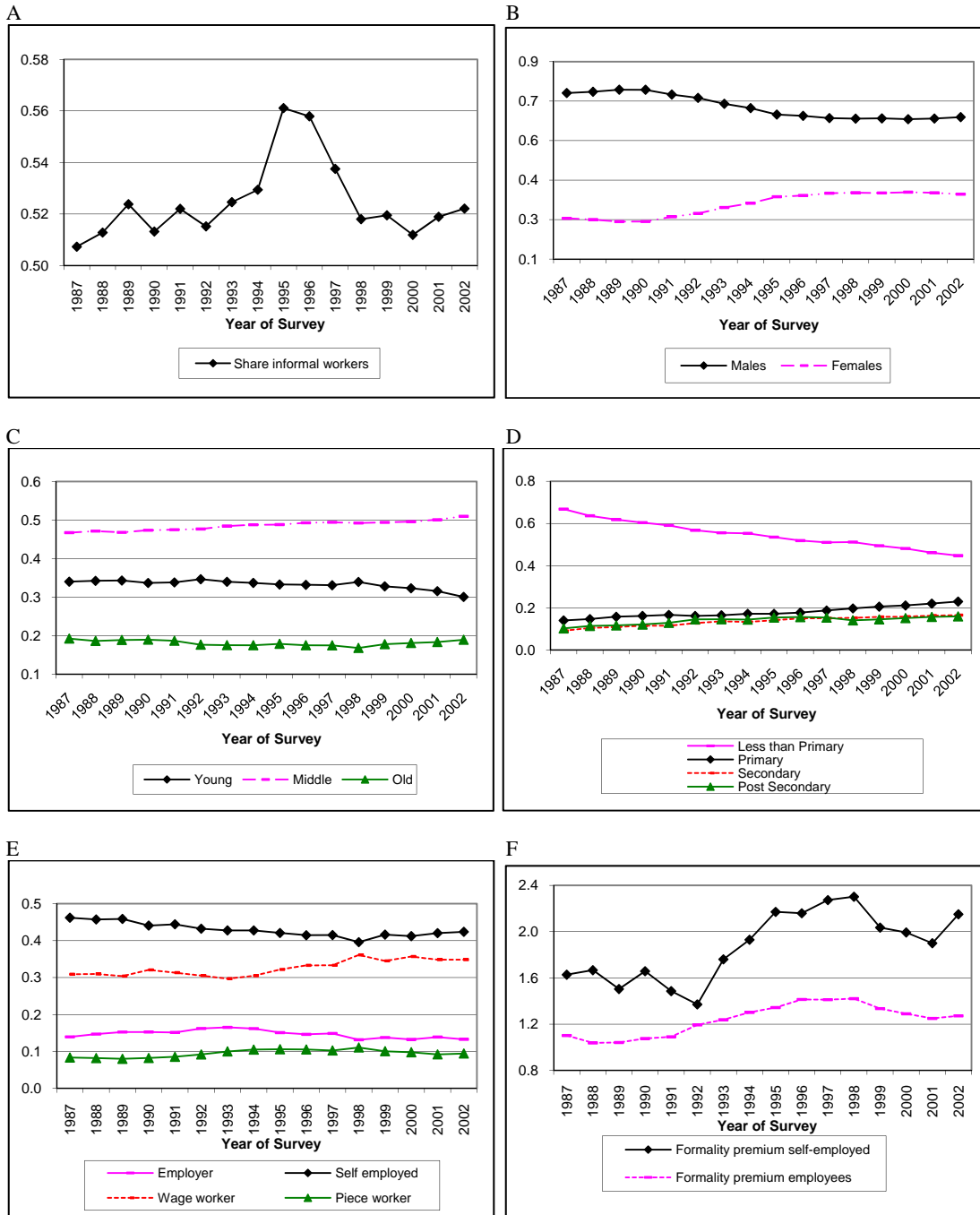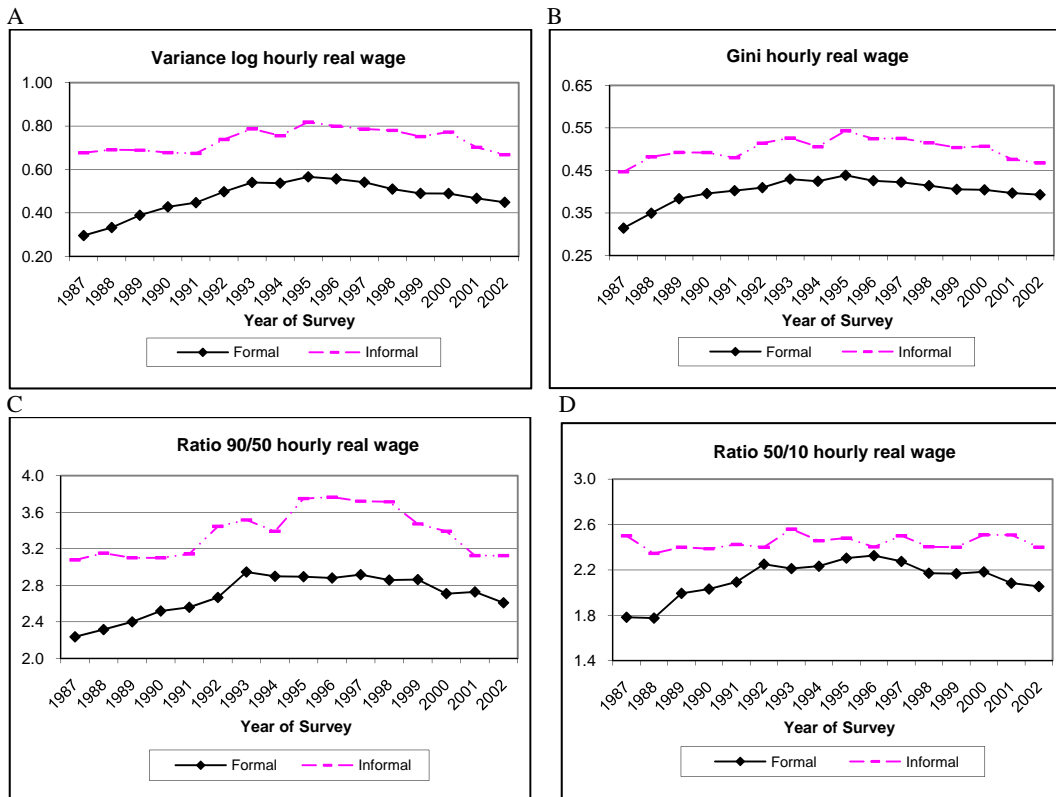

44

Figure 20: Wage Inequality for Formal and Informal Workers (Source: ENEU)

A

Variance log hourly real wage

B

Gini hourly real wage

C

Ratio 90/50 hourly real wage

D

Ratio 50/10 hourly real wage

Inequality is higher among informal workers bot at the top and at the bottom of the wage distribution. The 90th-50th and 50th-10th percentile ratios are on average, respectively, twenty-six and sixteen per cent higher for informal than for formal workers. With respect to the inequality statistics for hourly real wages computed in the aggregate sample that we presented in the graphs in Figure 9, it is interesting to note that the increase in the 50th-10th percentile ratio of the hourly real wage does only characterizes formal workers. Among informal workers inequality has remained pretty stable below the median of the wage distribution throughout the sample period.

6   Informality and wage inequality

The growth of the informal sector in many developing and transition economies has generated an extensive empirical and theoretical literature that studies the determinants of the choice to go informal and the characteristics of informal firms and workers.[18]

More recently, some contributions have studied the relationship between the changes in inequality and in the size of the informal economy by looking at the impact of inequality on informality. Chong and Gradstein (2007a) propose a theoretical model in which an increase in income inequality causes a bigger informal sector by lowering the relative benefits from becoming formal with the effect being stronger the weaker are the institutions and the protection of property rights in the formal sector. They test the theoretical predictions of the model using cross-country regressions over the period 1990-2000 and find empirical support for a positive impact of inequality on the size of the informal sector.

The reverse relationship of informality on inequality has been less investigated. On the theoretical side, the impact of informality on income inequality has been studied as a side effect of changes in tax revenues and in the availability of public goods. Johnson, Kaufmann and Zoido-Lobaton (1998) develop a model of the public sector where the share of the informal sector is negatively related to tax revenues. The decrease in tax revenues reduces the availability of public goods, which further reduces the willingness of the private sector to pay taxes. Loayza (1996) does also model the supply of congestible public goods as a negative function of the share of the informal sector and finds supporting empirical evidence among Latin American economies.

An increase in informality could also lead to an increase in overall wage inequality because of the labor market regulations and wage determination mechanism that characterize informal labor markets. While in the formal sector wages tend to be kept artificially stable by collective bargaining agreements and employment protection legislation, in the informal sector wages are unregulated and therefore flexible by design. An increase in the proportion of informal workers would therefore result into higher wage

---

[18]Maloney (1999) is one of the first contribution that studies the determinants of the choice of individual workers to go informal. Dessy and Pallage (2003) and DePaula and Schenkeiman (2007) are two recent contributions that study the factors driving the choice of firms to become informal.
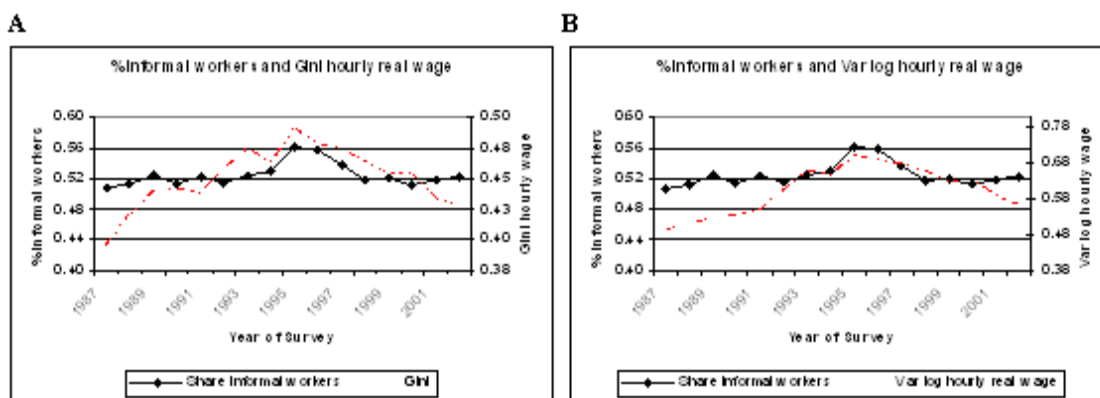
Figure 21: Share Informal Workers and Measures of Wage Inequality (Source: ENEU)

dispersion.[19]

On the empirical side, to the best of our knowledge, Rosser, Rosser and Ahmed (2000 and 2003) are the only two empirical papers that study the direct impact of informality on inequality. Rosser, Rosser and Ahmed (2003) focus on a group of eighteen transition economies in the 1990s and run a cross-country regression of the level and the changes in the Gini index for household income on a proxy for the size of the informal sector, the inflation rate, the GDP, the unemployment rate and an index of democratic rights and economic freedom. They find that the share of the informal sector has a positive and significant impact on the level of the Gini index. On the contrary, the regression run in first differences between two consecutive years shows that the changes in the share of the informal sector have no significant impact on the changes in the Gini index.

The results in the previous section show that for all indicators and both at the upper and at the lower tail of the distribution wage inequality is much higher among informal workers. This evidence suggests that changes in informality could be an important factor to explain the evolution of wage inequality. In order to investigate the relationship between informality and wage inequality we start by comparing the changes in the size of the informal sector and in some inequality indicators. The two graphs in Figure 21 present the share of informal workers and two measures of wage inequality, the variance of the logs and the Gini coefficient of hourly real wages, for each year of the ENEU between 1987 and 2002. The increase in informality between 1993 and 1995 does match very closely the size of the rise in wage inequality observed in Mexico in these years. The informal sector seems to be serving as a safety net: it takes on added significance in the hard times of the immediate wake of the peso crisis.

The changes in the share of informal workers not only track closely the changes in the inequality indicators in the mid 1990s, but also seem to follow a similar trend before and after the years of the

---

[19] We thank an anonymous referee for having suggested this channel.

peso crisis. The correlation coefficient between the share of informal workers and the Gini coefficient (the variance of the logs) is 0.72 (0.63) for the period 1987-1993, 0.89 (0.91) between 1994 and 1998, and -0.64 (-0.70) in the last years of the sample.

Therefore, changes in informality and in wage inequality are highly correlated. Given the two-way relationship between informality and wage inequality that we discussed earlier on, the direction of causation could go both ways. Therefore, without a valid instrument to act as an exogenous source of variation, all regression analysis does remain of a pure descriptive nature.

Being aware of this identification problem, the changes in the correlation coefficient provide some hints on the direction of causation. The increase in the correlation between changes in informality and in wage inequality in the years of the peso crisis suggests that the most plausible direction would go from informality to inequality. The peso crisis resulted into significant losses of jobs and a steep rise in unemployment. Evidence from the ENEU shows that the unemployment rate increased by almost three percentage points between 1993 and 1995. To many unemployed people the informal sector remained as the only option available; as a consequence, the size of the informal sector increased. Given that inequality is higher among informal workers, this resulted into an increase in overall wage inequality.

This motivates a regression analysis where the dependent variable is an indicator of wage inequality and the share of informal workers is included among the explanatory variables. We pool the data between 1987 and 2002 and we run a regression of the Gini coefficient of hourly real wages on the share of informal workers, the proportion of each of the four education levels (to control for changes in the skill composition), the proportion of the population belonging to each of the three age groups that we defined in the previous section (to control for changes in the age composition) and year. We run the same regression for the variance of the logs. Even if the regression coefficients have no causal interpretation, they provide a robustness' test of the relationship between informality and wage inequality to changes in the skill and age composition of the population over time.

Table 2 presents the results for the regression where the dependent variable is, respectively, the Gini coefficient (Model 1) and the variance of the logs of hourly real wages (Model 2). We only report the education and age shares' coefficients that are significant in any of the two regressions. We find that the share of informal workers has a strong positive correlation with both the Gini coefficient and the variance of the logs. The share of informal workers has an estimated coefficient of around 0.83 in the Gini regression and of around 0.87 when the dependent variable is the variance of the logs. In both regressions it is statistically significant at the five per cent level.

Finally, we test for the presence of a relationship between changes in the size of the informal sector and in wage inequality. We do so by running the regressions described above in first differences: we regress the change in each of the two inequality measures between two consecutive years on the change

Parameter	Model 1	Model 2
Share Informal Workers	0.8291	0.8707
	(0.2151)	(0.1747)
Share Primary	−1.4198	−1.2695
	(0.3525)	(0.5062)
Share Secondary	0.4665	−0.7509
	(0.5527)	(0.1629)
Share Young	1.4184	1.7797
	(0.2124)	(0.1369)
Year	−0.0121	−0.0028
	(0.0045)	(0.0110)
Constant	24.0096	6.2564
	(9.1755)	(21.6066)

Table 2: Wage Inequality and Informality, Level Regression, standard errors in parenthesis

in the share of informal workers, the education and age variables.

Parameter	Model 3	Model 4
Share Informal Workers	0.8355	0.8259
	(0.2063)	(0.1527)
Share Primary	−1.0976	−0.6769
	(0.4866)	(0.5597)
Share Secondary	0.3616	−0.5302
	(1.0222)	(0.4652)
Share Young	1.4368	1.4363
	(0.6051)	(0.4512)
Constant	−0.0074	0.0054
	(0.0087)	(0.0214)

Table 3: Wage Inequality and Informality, Growth Regression, standard errors in parenthesis

Table 3 presents the results for the regression where the dependent variable is, respectively, the first difference of the Gini coefficient (Model 3) and of the variance of the logs of hourly real wage (Model 4). As we did in Table 2, we only report the education and age shares' coefficients that are significant in any of the two regressions. We find that changes in the share of informal workers are positively and strongly associated with changes in wage inequality. In both regressions the coefficient of the change in the share of informal workers has a value of around 0.83 and it is statistically significant at the five per cent level.

The regressions that we have run only contain a limited number of controls. The list of important omitted variables includes unemployment and inflation rates, poverty rates and proxies for the economic reforms that characterized the Mexican economy in the 1990s. A careful choice of the relevant variables to include in the empirical specification could only be driven by a sound theoretical model that will have clear predictions to be tested. This task is left for future research.

7 Conclusion

This paper provides a complete characterization of the main changes in the distribution of earnings, income, consumption and wealth across Mexican households in the decade of the 1990s. We combine employment and wage information from the Mexican Urban Employment Survey (ENEU) with income, wealth and consumption data from the Mexican Income and Expenditure Survey (ENIGH). The ENIGH has a national coverage, which allows us to develop the analysis separately for urban and rural areas.

First, we document the trends in the level and dispersion of individual real hourly wages and household labor and asset income, consumption and wealth. As measures of dispersion, we use the Gini coefficient, the variance of the logs and the 90th-50th and 50th-10th percentile ratios. We choose the variance of the logs as our preferred measure of cross-sectional inequality and we decompose the changes over time in explained and residual components.

We find that both wage, consumption and income inequality increased significantly in the 1990s. For both wages and income the overall trend results from an initial rise in inequality in the first half of the 1990s and a subsequent drop or stable trend in the second half of the decade. Also, while the increase in wage and income inequality before the mid 1990s is generalized across the distribution, after 1995 inequality at the top and bottom of the distribution diverged with a steep decrease in inequality at the top. Only part of these dynamics is explained by the premium to observable characteristics that are important for wage determination, such as the premium to college education. The share of the variance of log wages and household income that is due to the returns to unobservable characteristics is high and increasing over time. Differently from income, dispersion in log household consumption did increase throughout the decade, especially in the top half of the distribution.

Then, we estimate the changes in wage, earnings and consumption inequality over the life-cycle. We do find strong evidence of time effects as opposed to cohort effects. This suggests that exogenous shocks such as a skill-biased technological change could be important to explain the dynamics of wage inequality observed in Mexico in the 1990s, while changes in the composition on observable characteristics such as birth cohorts would be irrelevant. This is very consistent with the results found by Binelli (2008): she found that changes in the composition on observables can not explain the trend in wage inequality observed in Mexico in the 1990s, which was driven by changes in the supply of education in reaction to an increased demand for skilled labor in production.

Finally, we focus on one important feature that characterized Mexican economy in the 1990s: the increase in the size of the informal sector and its relationship with the changes in wage inequality. We show that wage inequality is much higher among informal workers and that the changes in the size of the informal sector do track closely the changes in wage inequality, particularly in the mid 1990s, at the

height of the peso crisis.

These two pieces of evidence taken together suggest that changes in informality could be an important factor to explain the evolution of wage inequality in this decade. This prediction is confirmed by the results of a simple regression analysis of a measure of wage inequality on the share of informal workers. After controlling for changes in the skill and age composition, we find that the share of informal workers has a strong positive correlation with both the Gini coefficient and the variance of the logs of hourly real wage. We also find the correlation to be strong in first differences: changes in the share of informal workers are positively and strongly associated with changes in wage inequality.

References

[1] Airola J., Juhn C., (2005) "Wage Inequality in Post-Reform Mexico", IZA Discussion Paper, No. 1525.

[2] Attanasio O., Goldberg PK., Pavcnik N., (2002) "Trade Reforms and Wage Inequality in Colombia", EdePo discussion paper EWP03/01.

[3] Attanasio O., Székely M. (2004), "Wage shocks and consumption variability in Mexico during the 1990s", *Journal of Development Economics*, Vol. 73, No. 1, pp. 1-25.

[4] Attanasio O., Binelli C., (2003) "Redistribution Policy: Theoretical Motivation and Empirical Evidence", paper prepared for the AfD/EUDN Conference, Paris, November 2003.

[5] Autor D. H., Katz L. F., Kearney M. S., (2005) "Trends in U.S. Wage Inequality: Re-Assessing the Revisionists", NBER Working Paper 11627.

[6] Banco de Mexico, *Informe Sobre la Politica Monetaria*, September 1996.

[7] Binelli C., (2009) "Returns to Education and Increasing Wage Inequality", Oxford University, Mimeo.

[8] Bosch M., Maloney W. F., (2006) "Gross Worker Flows in the Presence of Informal Labor Markets: The Mexican Experience 1987-2002", The World Bank, Washington, D.C.

[9] Bosch M., Manacorda M., (2008) "Minimum Wages and Earnings Inequality in Urban Mexico. Revisiting the Evidence", LSE, Mimeo.

[10] Bustos P. (2005), "The Impact of Trade on Technology and Skill Upgrading: Evidence from Argentina", Mimeo.

[11] Chong A., Gradstein M., (2007a) "Inequality and informality", *Journal of Public Economics*, 91, pp. 159-179.

[12] Chong A., Gradstein M., (2007b) "Inequality and institutions", *Review of Economics and Statistics*, 89(3), pp. 454-465.

[13] Corseuil C. H., Muendler M., (2003) "Wage Gaps, Capital and Skills", Mimeo.

[14] Cragg M. I., Epelbaum M., (1996) "Why has wage dispersion grown in Mexico? Is it the evidence of reforms or the growing demand for skills?", *Journal of Development Economics*, Vol.51, 99-116.

[15] De Paula A., Scheinkman J. A., (2007) "The Informal Sector", Mimeo.

[16] Dessy S., Pallage S., (2003) "Taxes, inequality and the size of the informal sector", *Journal of Development Economics*, 70, pp. 225-233.

[17] Feenstra R. C., Hanson G., (1997) "Foreign Direct Investment and Relative Wages: Evidence from Mexico's Maquiladoras", *Journal of International Economics*, 42(3-4), 371-393.

[18] Goldberg PK., Pavcnik N., (2004) "Trade, Inequality, and Poverty: What Do we Know? Evidence from recent trade liberalization episodes in developing countries", *Brookings Trade Forum* 2004, Brookings Institute.

[19] Heathcote J., Storesletten K., Violante G., (2005) "Two Views of Inequality Over the Life-Cycle", *Journal of the European Economic Association* (Papers and Proceedings), Vol. 3 (2-3), pp. 543-552.

[20] Hanson G., Harrison A., (1999) "Trade Liberalization and Wage Inequality in Mexico", *Industrial and Labor Relations Review*, 52(2): 271-288.

[21] Jappelli T., Pistaferri L., (2008) "Financial Integration and Consumption Smoothing", CSEF Working Paper No. 200.

[22] Johnson S., Kaufmann D., Zoido-Lobaton P., (1998) "Regulatory Discretion and the Unofficial Economy", *American Economic Review Papers and Proceedings*, 88, 2:387-392.

[23] Loayza N. V., (1996) "The Economics of the Informal Sector: A Simple Model and Some Empirical Evidence from Latin America", Carnegie-Rochester Conference Series on Public Policy, 45: 129-162.

[24] Maloney W. F., (2004) "Informality Revisited", *World Development*, Vol. 32, No. 7, pp. 1159-1178.

[25] Maloney W. F., (1999) "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico", The World Bank Economic Review, Vol. 13, No. 2, pp.275-302.

[26] Pavcnik N., Blom A., Goldberg P., (2002) "Trade Liberalization and Labor Market Adjustment in Brazil", Mimeo.

[27] Riano A., (2008) "Trade, Technology Adoption and the Rise of the Skill Premium in Mexico", Mimeo.

[28] Roberston R., (2004) "Relative Prices and Wage Inequality: Evidence from Mexico", *Journal of International Economics*, 64(2), 387-409.

[29] Rosser J. B., Rosser M. V., Ahmed E., (2000) "Income Inequality and the Informal Economy in Transition Economies", *Journal of Comparative Economics*, 28, pp. 156-171.

[30] Rosser J. B., Rosser M. V., Ahmed E., (2003) "Multiple unofficial economy equilibria and income distribution dynamics in systemic transition", *Journal of Post Keynesian Economics*, Vol. 25, No. 3, pp. 425-447.

[31] Rubalcava L., Teruel G., (2004) "Un Análisis sobre las Economías de Escala en los Hogares Mexicanos", Mimeo.

[32] Rubalcava L., Santana A., Teruel G., (2005) "Escalas de Equivalencia para Mexico" in M. Székely, Números que mueven al mundo: La Medición de la Pobreza en México CIDE, ANUIS, Porrúa, 577-610.

[33] United Nations Development Programme, *Human Development Report 2007/2008.*

[34] Verhoogen E., (2008) "Trade, Quality Upgrading and Wage Inequality in the Mexican Manufacturing Sector.", *Quarterly Journal of Economics,* Vol. 123, No. 2.

[35] Winters A., McCullach N., McKay A., (2004) "Trade Liberalization and Poverty: the Evidence So Far", *Journal of Economic Literature*, 62, 72-115.

[36] Whitt J. A., (1996) "The Mexican Peso Crisis", *Economic Review Federal Reserve Bank of Atlanta*, January/February.

Appendix A - Descriptive Statistics

Variable	ENEU		ENIGH	
	Mean	SD	Mean	SD
	N = 3831264		N = 982647	
Yearly Hours worked				
Males	2172.167	984.3918		
Females	854.212	1094.412		
Employment rate				
Full-time	0.648	0.025		
Part-time	0.127	0.013		
Real hourly wage				
Males	28.135	49.803		
Females	22.792	37.974		
Equivalized labor income				
Urban			1038.933	2290.669
Rural			336.862	749.773
Equivalized asset income				
Urban			247.353	1640.912
Rural			152.685	752.645
Equivalized consumption				
Urban			738.252	1002.854
Rural			409.632	688.648
Equivalized financial wealth				
Urban			369.375	7500.763
Rural			167.421	1512.93

Table 4: Summary Statistics ENEU and ENIGH