Mexico in the 1990s: the main cross-sectional facts

Orazio Attanasio* and Chiara Binelli†
PRELIMINARY

This draft: June 2008.

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 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 characterized 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: J23, J24, J31.

*University College London and EDePo-Institute for Fiscal Studies.
†University College London and EDePo-Institute for Fiscal Studies.
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 index for household income in 2005 was 53.1 in Mexico, 45 in the US, 36.8 in the UK and 31.2 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.

First, we present 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.

Then, we highlight some distinctive features that characterized 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.

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

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 Consumer Expenditure Survey (CEX). It is a quarterly household survey that collects individual-level data and it has a rotating panel structure: households are interviewed for five consecutive quarters and in each quarter twenty per cent of the households are replaced by new households that are interviewed for the first time. The survey started in 1981 with progressive increase of the geographic coverage. The sample is selected to be geographically and socio economically representative of the Mexican urban population: by the end of the 1990s it covered approximately sixty two per cent of the national urban population and ninety two per cent of the cities with population greater than one hundred thousand people. By the end of the 1990s, approximately seventy four per cent of the Mexican population lived in urban areas. Recent rounds of the ENEU have national coverage, but since the earlier ones survey urban areas only, we restrict the sample to urban areas in all years.

The main questionnaire is divided in three parts. The first part collects socio demographic information on all household members. The second part contains detailed employment information on individuals at least twelve years old. The third part reports information on the characteristics of the house of residence with additional questions on the characteristics of the building, number and type of rooms and ownership status from the 1994 wave onwards.

Most importantly for our analysis, 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. The survey is representative at the national level and is also representative of rural and urban areas separately. It
is available for 1984, 1989, 1992 and every two years since then. From 1992 the surveys are strictly comparable in terms of sampling frame, sampling methodology, timing and recall periods. 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.\footnote{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 will be followed by additional waves in order to build up a uniquely rich longitudinal database that spans over a period of at least ten years. However, as of today, only the 2002 wave is available to the general public.} 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.

We consider all waves between 1992 and 2002 and 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 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.

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

This section is divided in two parts. In the first part 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. In the second part we study the changes in inequality for each of these variables. We compute different measures of cross-sectional dispersion: the variance of the logarithms, the Gini coefficient, the 90th-10th, the 90th-50th and the 50th-10th percentile ratios. 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 zeros are excluded.

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 wage, 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 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 nuclear households. 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 nuclear households between three and six per cent of the sample. We will investigate the role of changes in household composition to explain the variation in income and consumption inequality.

Details on the sample sizes and summary statistics for the main variables of interest in the ENEU and in the ENIGH samples are reported in Appendix A.

3.1 Cross-sectional means

3.1.1 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 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 of the real hourly wage for males and females. As expected, for each year of the sample, the level of the wages is significantly lower for females. However, the two series follow a very similar trend. They both increase up to the mid

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. 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 of the real hourly wages of male and female 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 ten to twenty per cent for each year between 1992 and 1998 and of around thirty per cent in 2000 and 2002.

The discrepancy in the level of wages in the ENEU and in the ENIGH samples could be due to the differences in the way 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 over the number of self-reported worked hours. On the contrary, the ENEU does not distinguish between different sources of income.

### 3.1.2 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 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 in both samples. As expected, the female curve lies well below the male one: women do work on average four hundred and fifty hours per year less than men.

---

²For a detailed description of Mexican peso crisis, see Arner (1996).
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 experiences a drop of around two percentage points between 1994 and 1995 reaching the lowest value of around eighty-seven per cent in the year 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, 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.

3.1.3 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 (alimony, child support, transfers from relatives, etc.) and income from retirement plans. 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 as the sum of dividends on stocks, interests on bonds and bank accounts net of 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 as the

³For 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 an easier comparability of the results.
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 additional owned property.

The sum of our second measure of labor income and net financial income gives a measure of total pre-government 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. However, the ENIGH does not report information on income from public transfers and social programs. Therefore, our most comprehensive measure of household income does not account for the role of government’s transfers.

Figure 6 and 7 present the mean of the two measures of equivalized labor income in the aggregate and separately for urban and rural areas. In the overall sample as well as in the urban and rural one the series for the two labor income measures 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.

Figure 8 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 very small proportion of total asset income. Therefore, the series for net financial income almost coincides with the one for total asset income. We only report the results for net financial income.

In the aggregate data as well as in both the urban and the rural sample, net financial income decreases monotonically in the period 1992-1996 and it is almost constant until the year 2000. It then increases in the last year of the sample to reach a value of around two hundreds pesos per equivalent unit in both urban and rural areas.

3.1.4 Consumption

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 including and by 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 expenditures.
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.

Figure 9 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 household consumption, the series for the second measure of consumption follows a very similar trend. Therefore, we only report the results on the first measure of consumption.

Consumption decreases between 1992 and 1996 when it starts increasing slightly. The initial decrease and following increase in consumption is concentrated in urban areas while in rural areas consumption is essentially flat throughout the sample period.

3.1.5 Wealth

As a measure of wealth we use total financial wealth within a household, that is the sum of financial wealth hold by all household 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 residential real estate owned 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.

Figure 10 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.

3.2 Dispersion measures

3.2.1 Wages and hours of work

The figures in Panel 1 present the evolution of the variance of the logarithm, the Gini coefficient, the 90th-10th, the 90th-50th and the 50th-10th percentile ratios of the hourly real wages for each ENEU
wave between 1987 and 2002.

As documented by several previous studies on wage inequality in Mexico\textsuperscript{4}, all inequality indicators increase sharply between the end of the 1980s and 1995 and tend to decline afterwards. Between 1987 and 1995 the 90th-10th percentile ratio increased by around twenty-three per cent, the logarithm of the variance by thirty-one per cent and the Gini coefficient by five percentage points. Between 1995 and 2002 the 90th-10th ratio and the logarithm of the variance fell, respectively, by around twenty-two and twenty per cent, while the Gini coefficient decreased of around seven percentage points.

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 figures 4 and 5, while the 50th-10th ratio of hourly wage continues to rise throughout the 1990s, the 90th-50th declines sharply after 1995. The deceleration of aggregate inequality growth in the second half of the 1990s seems to reflect a reversal in upper-tail inequality expansion paired with a continuing rise in inequality below the median of the wage distribution.\textsuperscript{5}

The inequality increase between 1990 and 1995 (and the decrease afterwards) measured with the 90th-10th ratio is less (more) pronounced than the one measured with the variance of the logarithms. This is evidence that most of the change in the variance of the log hourly wage is due to changes in the top ten per cent of the earnings distribution. Given the tenth percentile mostly consists of workers with post secondary education, returns to College are a potential important source of wage dispersion.

The first two figures of panel 2 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.

Returns to College are monotonically and steeply increasing. Between 1990 and 2000 they rose by around twenty per cent. In the same period, returns to Secondary decreased by about two per cent.\textsuperscript{6}

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.

\textsuperscript{4}See, among the others, Hanson and Harrison (1999), Robertson (2004), Airola and Juhn (2005) and Manacorda and Bosch (2008).
\textsuperscript{5}These differential trends do contrast 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 continuing rise in inequality above the median. See, among the others, Autor, Katz and Kearney (2005).
\textsuperscript{6}The double change in the wage differentials between College and Secondary and Secondary and less than Secondary education in the 1990s resulted into the "convexification" of the wage profile. See Binelli (2008).
and the experience premium as the average wage of males aged 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.

Figures 3 to 5 in panel 2 present the evolution of the gender, experience and formality premium. The gender premium decreases 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 is steeply increasing between 1987 and 1995 and it decreases thereafter. However, the post-1995 decline did not offset the 1987-1994 rise: overall, the premium to experience increased by around twenty per cent between 1987 and 2002.

As shown in figure 5, the formality premium is steeply increasing between the end of the 1980s and 1998 and it slightly declines thereafter. The premium differs significantly by job categories. In the data there are four main types of informal workers: self-employed, employers, wage and piece workers. We group employers and self-employed in the self-employment category and wage and piece workers in the employees category. As figure 6 shows, in each year of the sample, the formality premium for the self-employed is at a much higher level than the one for employees.

Changes in observable variables accounts only for some of the total variation in wage dispersion. 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 wage on a quadratic polynomial in age as a proxy for labor market experience, and dummy variables for the education level and the formality/informality nature of the main job.

Figure 7 in panel 2 presents the variance of the residuals of this regression year by year. The share of the total variance 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. Figure 1 in panel 3 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 trend. As for the variance of the residuals, they reach a peak in 1995 and decline thereafter. A comparison between figure 1 in panel 3 and figure 7 in panel 2 shows that the variance of the residuals accounts for an average of seventy per cent of the total variation of log hourly real wages.

The variance of the log yearly number of worked hours does also show an humpshaped profile. Figure
2 in panel 3 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 the female sample is around three times the size of the male variance.

Figure 3 and 4 in panel 3 present the correlation coefficient between hours and wages in the female and the male sample. In both samples and for all years the correlation is negative and tends to be very noisy, especially for females.

3.2.2 Labor and asset income

The figures in panel 4 present the variance of the logs, the Gini coefficient and the 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 increases between 1992 and 1994 and decreases sharply between 1994 and 1996. Afterwards, inequality appears to be slightly increasing when measured with the variance of the logs or the 90/10 percentile ratio while it decreases if the Gini coefficient is used.

There are significant differences between inequality trends 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, inequality is increasing in these years. In the same way, the slightly increasing trend of the variance of the logs and the 90/10 percentile ratio 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 declines at the top of the distribution and increases below the median.

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.

Together with a measure of equivalized household labor income, we compute a measure of residual income defined as the log equivalized income after having controlled for the effects on income of 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 (single 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.

The first figure in panel 5 presents the variance of the log raw, equivalized and residual labor income.
The three series follow a very similar trend between 1992 and 1996 with a steep rise in 1994 and a sharp decrease in 1996. Between 1996 and 2002, log equivalided and raw income increase and converge to a very similar mean value in 2002. 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.

In order to disentangle the contribution of each of the observable variables to the changes in income inequality, for every year between 1992 and 2002 we compute the cross-sectional variance of the fitted values of the polynomial in age, the education and the household composition dummies.

Figure 2 in panel 5 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 decreases from around seventeen per cent in 1992 to eleven 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.

Figures 3 and 4 in panel 5 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 figures in panel 6 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 pre-government household income for the overall sample and separately for urban and rural areas.

Both in the aggregate data and in urban and rural areas, the extent of inequality measured when we consider total pre-government household income is lower than when we consider labor income plus private transfers and the extent measured with the latter tends to be lower than when we compute inequality
using labor income that does not include private transfers. This suggests that some insurance/risk sharing mechanism between Mexican households could be in place.

3.2.3 Consumption

The figures in panel 7 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 inequality increases non monotonically with two peaks in 1994 and in 1998. The peso crisis seems to result in 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 aftermaths of the peso crisis. Instead, consumption inequality increased up to 1998, decreased in 2000 and it started increasing again in 2002.

Differently from the diverting trends in the upper and lower tail of the wage and income distribution, consumption inequality rises both at the top and at the bottom of the distribution in urban as well as in rural areas. As shown in figure 4 and 5, between 1996 and 2002 both the 50th-10th and the 90th-50th ratios do rise in all samples.

We apply to log equivalized household consumption the same decomposition that we did for log equivalized income. We compute a measure of log residual consumption by running an OLS regression of log equivalized consumption on the same set of controls that we used to decompose 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.

Figures 1 and 2 in panel 8 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. Figures 3 and 4 present the decomposition of the variance of log equivalized consumption for the urban and the rural sample.

The variance of log equivalized and residual consumption increase significantly over time while the
variance of log raw consumption is only slightly increasing from the second half of the 1990s.

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 contribution of household composition is much higher in rural areas, with an average of fifteen per cent in the sample, which is double the size the share in the urban one. 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 figures in panel 9 compare the variance of the logs, the Gini coefficient and the percentile ratios for equivalized household income and consumption in the aggregate data. All inequality indicators increase significantly over time for both consumption and income. The increase measured with the variance of the logs is of around fifty-three per cent for consumption and fifty per cent for income. The Gini coefficient increases by around two percentage points for income and ten percentage points for consumption. The 90th-10th ratio more than doubles for income and increases by around eighty per cent for consumption.

All indicators increase steeply in 1994 and decline in 1996. The immediate impact of the devaluation of the mid 1990s was bigger in magnitude for income than for consumption. However, while all indicators of consumption inequality keep on increasing steadily until 2002, the income gap narrows for some deciles of the distribution. While the ratio between the 90th and the 10th income percentile and the 50th and the 10th increases, the one between the 90th and the 50th percentile significantly decreases.

### 3.2.4 Wealth and National Accounts data

The first two figures in panel 12 present the Gini coefficient for net financial wealth and net total wealth in the aggregate data and separately for urban and rural areas. The trends of net financial and total wealth do almost coincide with very small differences in the level of wealth for some years.

As for wages, income and consumption, wealth data show the impact of the peso crisis. All series drop sharply in 1994 and increase in 1996. Household wealth in the aggregate and in urban areas decline steadily between 1996 and 2000 while it increases between 1996 and 1998 in rural areas.

Figures 3 and 4 in panel 12 show the wealth income ratio computed by dividing net financial wealth and net total wealth over total disposable income across all households in the sample. The two series exhibit a pronounced spike in 1996 but remain below one in each year of the sample.

The figures in panel 13 compare three main variables computed from the employment and expenditure surveys and from the Mexican National Accounts data. Figure 1 plots the employment to population
ratio, figure 2 presents the mean monthly earnings\textsuperscript{7} and figure 3 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. Then, we consider all individuals and households in the sample without applying any age restriction.

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, 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 figure 1, 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 1988 and 1993 it is on average four per cent lower than the one computed from the National Accounts data. From 1994 and 2002, it is on average ten per cent higher.

As shown in figure 2, monthly earnings do also follow a very similar trend in the two data sources. As 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 figure 3, 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 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 much more 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.

\textsuperscript{7}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 Mexican Employment Survey.
3.2.5 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-portions 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-portions of the age-cohort regression gives the life-cycle profile having controlled for cohort effects.

The figures in panel 10 present the set of results obtained when we assume away cohort effects and we control for year effects. The figures in panel 11 present the results obtained when we abstract from time effects and we control for cohort effects. We use the ENEU for the analysis on hourly wages and the ENIGH for the analysis on earnings and consumption. All graphs are normalized to match the unconditional average for the corresponding cross-sectional moment over the sample period.

The age profiles of wages and earnings do differ significantly when we control for cohort or for time effects. The profile for 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 always decreasing in both cases but much more when cohort effects are accounted for.

As in HSV, we find that the changes in inequality over the life cycle are generally substantially larger when one controls for cohort effects and that the age profile of inequality for household consumption is less affected by the assumption made about the time/cohort effects.

4 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 figures in panel 14 present the proportion of informal workers among the 25-60 age population,
overall and by some relevant observable characteristics. As documented by previous studies\(^8\), the share of
the informal workers in the 1990s accounts for around half of the urban workforce in Mexico: between
1987 and 2002 it increases only slightly from around fifty to fifty-two per cent. However, in 1995 it jumps
up by four percentage points to reach the value of fifty-six per cent in one year. In 1998 it is back to the
pre-1995 level and it fluctuates around a value in 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 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 as all individuals aged between 25 and 35, the middle
group as those aged between 36 and 50 and the old group as those aged 51 to 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 levels according to the highest completed schooling cycle. As documented
by previous studies\(^9\), the vast majority of informal workers have very low education levels. Figure 4
shows that over sixty per cent of informal workers in 1987 and still more than forty per cent in 2002 have
less than completed primary education. The proportion of workers with completed primary, secondary
or post-secondary education increases 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 increases 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 figure 5, 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 figures in panel 15 present the variance of the logs, the Gini coefficient, the 90th-10th, 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.

\(^8\)See, among the others, Maloney (2004) and Bosch and Maloney (2006).
\(^9\)See, among the others, Maloney (1999).
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 value of the Gini coefficient is of about ten percentage points higher for informal workers.

Inequality is higher among informal workers both at the top and at the bottom of the wage distribution. The 90th-10th, 90th-50th and 50th-10th percentile ratios are on average, respectively, forty-six, twenty-six and sixteen per cent higher for informal than for formal workers.

5 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.\textsuperscript{10}

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 much less investigated. 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.

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 et al. (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.

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\textsuperscript{10} 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 Schenkelman (2007) are two recent contributions that study the factors driving the choice of firms to become informal.
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 successive years shows that the changes in the share of the informal sector have no significant impact on the changes in the Gini.

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 figures in panel 16 present the share of the informal workers and two measures of wage inequality, the variance of the logs and the Gini coefficient of hourly real wage, for each ENEU wave between 1987 and 2002. The increase in informality between 1993 and 1995 does match almost perfectly 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 peso crisis. The correlation coefficients between the proportion of informal workers and each of the two inequality measures give a first indication of the relationship between informality and wage inequality. For the period 1987-1993 the correlation is 0.72 for the Gini coefficient and 0.63 for the variance of the logs; between 1994 and 1998, the coefficients increase to, respectively, 0.89 for Gini and 0.91 for the variance of the logs. Finally, in the last years of the sample, the correlation coefficients become negative: -0.64 for Gini and -0.70 for the variance of the logs.

We pool the data between 1987 and 2002 and we run a regression of the Gini coefficient of hourly wage 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.

Table 1 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 wage (Model 2). We only report the education and age shares’ coefficient 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 coefficient of the share of informal workers has a value of around 0.83 in the Gini regression and a value 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 the changes in the size of the informal sector and the changes 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 successive years on the change in the share of informal workers, the education and age controls.

Table 2 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 1, we only report the education and age shares’ coefficients that are significant in any of the two regressions.

We find that the 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.

6 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 to develop the analysis separately for urban and rural areas.

We document the trends in the level and dispersion of real hourly wages, labor and asset income, consumption and wealth. As measures of dispersion, we use the Gini coefficient, the variance of the logs and the 90th-10th percentile ratio. 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 component.

We find that both wage, consumption and income inequality increase significantly over time. The increase measured with the variance of the logs is of around fourteen per cent for hourly real wages, fifty-three per cent for household consumption and seventy-five per cent for household income.

Between 1992 and 2002, the share of the residual (unexplained) component of the variance of household consumption and income is at an average value of, respectively, eighty and eighty-five per cent. The share of the total variance of log equivalized household income (consumption) accounted for by observables decreases from around eighteen (nineteen) in 1992 to around eleven (thirteen) per cent in 2002.

Among the observables, education accounts for the biggest share of the explained variance of consumption, income and wages. Together with education, the informal nature of the job appears as an important determinant of wage dispersion. 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 in the 1990s. 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 of a measure of wage inequality on the share of informal workers. After controlling for changes in the skill and age composition, 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.

These results could vary under different specifications if other possibly relevant variables were to be included among the controls. The list of important omitted variables includes unemployment and inflation rates, poverty rates and proxies for economic and financial liberalization that characterized 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.
References


Figure 4

Full time employment rate, ENEU

Year of Survey

Females

Males

Figure 5

Part time employment rate, ENEU

Year of Survey

Females

Males
Figure 9

Equivalized consumption, ENIGH

Figure 10

Equivalized net financial wealth, ENIGH
Figure 5

50/10 percentile equivalized consumption, ENIGH

Year of Survey

Panel 8

Figure 1

Variance log consumption, All, ENIGH

Year of Survey

Figure 2

Decomposition log equivalized consumption, All, ENIGH

Year of Survey

Figure 3

Decomposition log equivalized consumption, Urban, ENIGH

Year of Survey

Figure 4

Decomposition log equivalized consumption, Rural, ENIGH

Year of Survey
Panel 10
Figure 1

**Variance log hourly real wages life cycle, ENEU**

- 0.59
- 0.60
- 0.61
- 0.62
- 0.63

Age
With year effects

Figure 2

**Variance log raw earnings life cycle, ENIGH**

- 2.25
- 2.30
- 2.35
- 2.40

Age
With year effects

Figure 3

**Variance log equivalized earnings life cycle, ENIGH**

- 2.80
- 2.83
- 2.85
- 2.88
- 2.90

Age
With year effects

Figure 4

**Variance log equivalized consumption life cycle, ENIGH**

- 1.10
- 1.15
- 1.20
- 1.25
- 1.30

Age
With year effects

Panel 11
Figure 1

**Variance log hourly real wages life cycle, ENEU**

- 0.50
- 0.59
- 0.68
- 0.77

Age
With cohort effects

Figure 2

**Variance log raw earnings life cycle, ENIGH**

- 2.0
- 2.2
- 2.3
- 2.5

Age
With cohort effects
### Table 1

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share informal workers</td>
<td>0.8291**</td>
<td>0.8707**</td>
</tr>
<tr>
<td></td>
<td>(0.2151)</td>
<td>(0.1747)</td>
</tr>
<tr>
<td>Share Primary</td>
<td>-1.4198**</td>
<td>-1.2695**</td>
</tr>
<tr>
<td></td>
<td>(0.3525)</td>
<td>(0.5062)</td>
</tr>
<tr>
<td>Share Secondary</td>
<td>0.4665</td>
<td>-0.7509**</td>
</tr>
<tr>
<td></td>
<td>(0.5527)</td>
<td>(0.1629)</td>
</tr>
<tr>
<td>Share Young</td>
<td>1.4184**</td>
<td>1.7797**</td>
</tr>
<tr>
<td></td>
<td>(0.2124)</td>
<td>(0.1569)</td>
</tr>
<tr>
<td>Year</td>
<td>-0.0121**</td>
<td>-0.0028</td>
</tr>
<tr>
<td></td>
<td>(0.0045)</td>
<td>(0.0110)</td>
</tr>
<tr>
<td>Constant</td>
<td>24.0096**</td>
<td>6.2564</td>
</tr>
<tr>
<td></td>
<td>(9.1755)</td>
<td>(21.6066)</td>
</tr>
<tr>
<td># obs = 16</td>
<td># obs = 16</td>
<td></td>
</tr>
</tbody>
</table>

*Standard errors in parentheses. The model includes dummy variables for State and year of birth. Dependent variable is the Gini coefficient for hourly real wage for Model 1 and the variance of the log hourly real wage for Model 2. ** indicates significance at the five per cent level.

### Table 2

<table>
<thead>
<tr>
<th></th>
<th>Model 3</th>
<th>Model 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share informal workers</td>
<td>0.8355**</td>
<td>0.8259**</td>
</tr>
<tr>
<td></td>
<td>(0.2063)</td>
<td>(0.1527)</td>
</tr>
<tr>
<td>Share Primary</td>
<td>-1.0976**</td>
<td>-0.6769</td>
</tr>
<tr>
<td></td>
<td>(0.4866)</td>
<td>(0.5597)</td>
</tr>
<tr>
<td>Share Secondary</td>
<td>0.3616</td>
<td>-0.5302</td>
</tr>
<tr>
<td></td>
<td>(1.0222)</td>
<td>(0.4652)</td>
</tr>
<tr>
<td>Share Young</td>
<td>1.4368**</td>
<td>1.4363**</td>
</tr>
<tr>
<td></td>
<td>(0.6051)</td>
<td>(0.4512)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0074</td>
<td>0.0054</td>
</tr>
<tr>
<td></td>
<td>(0.0087)</td>
<td>(0.0214)</td>
</tr>
<tr>
<td># obs = 15</td>
<td># obs = 15</td>
<td></td>
</tr>
</tbody>
</table>

*Standard errors in parentheses. The model includes dummy variables for State and year of birth. Dependent variable is the first difference of the Gini coefficient of hourly real wage for Model 3 and of the variance of the log hourly real wage for Model 4. ** indicates significance at the five per cent level.
APPENDIX A

<table>
<thead>
<tr>
<th>Variable</th>
<th>ENEU Mean</th>
<th>ENEU SD</th>
<th>ENIGH Mean</th>
<th>ENIGH SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yearly Hours worked</td>
<td>N = 3807300</td>
<td></td>
<td>N = 916956</td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>2114.57</td>
<td>1011.899</td>
<td>1038.933</td>
<td>2290.669</td>
</tr>
<tr>
<td>Females</td>
<td>815.187</td>
<td>1078.917</td>
<td>336.862</td>
<td>749.773</td>
</tr>
<tr>
<td>Employment rate</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time</td>
<td>0.633</td>
<td>0.023</td>
<td>247.353</td>
<td>1640.912</td>
</tr>
<tr>
<td>Part-time</td>
<td>0.143</td>
<td>0.017</td>
<td>152.685</td>
<td>752.645</td>
</tr>
<tr>
<td>Real hourly wage</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>28.274</td>
<td>48.071</td>
<td>735.201</td>
<td>1000.022</td>
</tr>
<tr>
<td>Females</td>
<td>22.657</td>
<td>34.825</td>
<td>411.486</td>
<td>696.472</td>
</tr>
<tr>
<td>Equivalized labor income</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>1038.933</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>336.862</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equivalized asset income</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>247.353</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>152.685</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equivalized consumption</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>735.201</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>411.486</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equivalized financial wealth</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>369.376</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>167.421</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>