This version: March, 2018

A GUIDE TO

LABLAC

LABOR DATABASE

FOR LATIN AMERICA AND THE CARIBBEAN

CEDLAS [*]
and
The World Bank [**]

[*] Centro de Estudios Distributivos, Laborales y Sociales at Universidad Nacional de La Plata.
[**] The World Bank's LAC Poverty and Gender Group (LCSPP).

1. INTRODUCTION

LABLAC is a database of labor statistics constructed from microdata of the Latin American (LA) labor surveys, developed by CEDLAS (Universidad Nacional de La Plata) and The World Bank's LAC poverty group (LCSPP). This guide describes the contents of each table in the database, and briefly explains the main methodological decisions taken to construct the statistics.

LABLAC complements SEDLAC, the *Socioeconomic Database for Latin America and The Caribbean*, by presenting statistics for the labor market on a monthly or quarterly basis.

A large-scale dataset like LABLAC implies a daily work of revising and controlling the construction of variables and statistics, identifying mistakes, and searching for more efficient ways of processing the data and presenting the statistics. In this sense, LABLAC is expected to be changing over time. The suggestions of researchers and users are important inputs to improve the database. LABLAC is intended to be a space of interactions among researchers on labor issues in LAC.

On the comparability of statistics

Labor surveys are not uniform across LAC countries. They significantly differ in geographical coverage and questionnaires. Surveys are also different within countries over time. LA governments have been improving their surveys over the past decade, changing coverage and questionnaires. The issue of comparability is, hence, of a great concern. In particular, how comparable (across countries and over time) are the statistics shown in LABLAC?

We make all possible efforts to make statistics comparable across countries and over time by using similar definitions of variables in each country/year, and by applying consistent methods of processing the data. However, perfect comparability is not assured, as the coverage and questionnaires of labor surveys differ among countries, and frequently also within countries over time. Hence, a trade-off arises between accuracy and coverage. If we want to be ambitious in the analysis, we have to pay the cost of losing accuracy and getting into comparability problems. Sometimes these problems are too severe and it is convenient to restrict the analysis. This guide and other documents in our web page provide the user with relevant information to decide on that trade-off. The final decision whether making a comparison or not depends on the preferences and specific needs of each user.

Citation

Information taken from this database should be cited as "Source: LABLAC (CEDLAS and The World Bank)" or "Source: Labor Database for Latin America and the Caribbean – SEDLAC (CEDLAS and The World Bank)". We advise making reference to the date when the database was consulted, as statistics may change.

2. LABLAC Tables

This section has four files: *employment_LABLAC.xls*, *wages&hours_LABLAC.xls*, *labor_benefits_LABLAC.xls* and *surveys_LABLAC.xls*

The file *employment_LABLAC.xls* has the following sheets:

labor force : Labor force participation by gender, age, education, and area.

employment: Share of adults employed by gender, age, education, and area.

unemployment: Unemployment rates by gender, age, education, and area.

extended unemployment: Rates of unemployed and inactive people who wish to work by gender, age, education, and area.

underemployment: Rate of employed people that work less than 15 hours a week by gender, age, education, and area.

involuntary underemployment: Rate of employed people that work less than 15 hours a week and wish to work more hours by gender, age, education, and area.

duration: Duration of unemployment (in months) by gender, age, education, and area.

change: Share of workers willing to change employment and/or increase hours of work by gender, age, education, and area.

structure: Distribution of workers by gender, age, education and area.

stru_region: Distribution of workers by region.

stru_type: Distribution of workers by labor relationship, type of firm and labor category.

stru_sector: Distribution of workers by economic sector.

informal_1: Share of adults workers in informal jobs. Definition 1: Informal=salaried workers in small firms, non-professional self-employed and zero-income workers.

informal_2: Share of salaried adults workers in informal jobs. Definition 2: Absence of social security rights.

child: Share of children aged 10-14 who are employed.

MDG: Share of women in non-agricultural employment.

The file *wages&hours_LABLAC.xls* has the following sheets:

wage_1: Hourly wages in main activity in nominal LCU by gender, age, education, and area.

wage_2: Hourly wages in main activity in nominal LCU by type of job.

wage_3: Hourly wages in main activity in nominal LCU by informality status.

wage_4: Hourly wages in main activity in nominal LCU by sector.

hours_1: Weekly hours of work by gender, age, education, and area.

hours_2: Weekly hours of work by type of job.

hours_3: Weekly hours of work by informality status.

hours_4: Weekly hours of work by sector.

li_1: Monthly labor income in LCU by gender, age, education, and area.

li_2: Monthly labor income in LCU by type of job.

li_3: Monthly labor income in LCU by informality status.

li_4: Monthly labor income in LCU by sector.

Ginis: Gini coefficients for the distribution of wages and earnings.

correlation: Linear correlation coefficients between hours of work and hourly wages.

wagegaps: Ratio of hourly wages by educational groups. Prime-age males.

Mincer: Coefficients of educational dummies in Mincer equations. Males and females.

Mincer2: Dispersion in unobservables and conditional gender wage gaps.

The file *labor_benefits_LABLAC.xls* has the following sheets:

contract: Share of salaried workers with labor contracts by age, gender, education and area.

pensions: Share of salaried workers with right to pensions when retired by age, gender, education and area.

health: Share of salaried workers with right to health insurance linked to the job by age, gender, education and area.

others: Share of salaried workers with access to 13th month and holidays. Share of salaried workers unionized.

The file *surveys_LABLAC.xls* contains information about the sample size and contents of all the surveys used to calculate statistics LABLAC.

3. Methodological issues

Type of firm

Workers are classified into three groups according to whether they work in small firms, large firms or the public sector. The latter includes jobs in stated-owned firms,

public schools, hospitals and other services, and public administration.

Informality

There are at least two different concepts that are referred by the term *labor informality*. The "productive" definition pictures informal workers as those in low-productivity, unskilled, marginal jobs, while the "legalistic" or "social protection" definition stresses the lack of labor protection and social security benefits. The productive definition is concerned with the type of job (*e.g*. salaried vs. self-employed, large vs. small firms), while the legalistic definition is concern with the compliance of the labor relationship with some rules, mainly labor protection.

The empirical implementation of the productive notion of informality has been linked to

(i) the type of job (salaried, self-employment), (ii) the type of economic unit (small, large, public sector), (iii) and the worker's skills. Following this practice, we divide the working population into 7 groups:

1. Entrepreneurs (*patrones*)
2. Salaried workers in large private firms
3. Salaried workers in the public sector
4. Salaried workers in small private firms
5. Skilled self-employed
6. Unskilled self-employed
7. Zero-income workers

To implement this classification we include as *unskilled* all individuals without a tertiary or superior education degree, and we define as *small* all firms with 5 or fewer employees. Given that an individual could have more than one job, we apply the classification only to his/her main occupation. We implement the following definition of labor informality:

Definition 1 (productive definition): *An individual is considered an informal worker if (s)he belongs to any of the following categories: (i) unskilled self-employed, (ii) salaried worker in a small private firm, (iii) zero-income worker.*

A second strand of the literature has stressed the "legalistic" or "social protection" notion of informality. Informal firms are those not complying with the norms in terms of labor contracts, labor taxes, and labor regulations, and then their workers have no rights to labor protection or social benefits linked to employment.

The right to receive a pension when retired is the social security benefit most asked in LAC household surveys. However, not all countries have questions on this item, and in those that have, questions are different (see table below). Moreover, in most countries the questions apply only to salaried workers, leaving all the self-employed as missing.

We implement the following legalistic/social-protection definition of informality:

Definition 2 (legalistic or social protection definition): *A salaried worker is informal if he (she) does not have the right to a pension linked to employment when retired.*

Country	A worker is formal if she ..
Argentina	The job has retirement withholding
Brazil	Contributes to the Social Security system
Chile	Is affiliated with any social security system
Colombia	It is currently trading at a pension fund
Ecuador	Has Social Security
Mexico	Has the right to a pension when retired
Nicaragua	Is inscribed with the Obligatory Social Security
Paraguay	Contribute to a retirement fund for this work
Peru	Is affiliated with any social security system
Uruguay	Has contributions to the Retirement Fund

Mincer equations

We run regressions of the logarithm of the hourly wage in the main occupation for adults aged 25 to 55. As regressors we include educational dummies, age, age squared, an urban dummy and regional dummies. We estimate the model separately for men and women by Heckman maximum likelihood methods. The selection equation includes in addition school enrollment and the number of children. In the tables we report the marginal "returns" to completing each educational level.

The Mincer equation is also informative on two interesting factors: the role of unobservable variables and the gender wage gap. The error term in the Mincer regression is usually interpreted as capturing the effect on hourly wages of factors that are unobservable in household surveys, like natural ability, contacts and work ethics. An increase in the dispersion of this error term may reflect an increase in the returns to these unobservable factors in terms of hourly wages. We show the standard deviation of the error term of each Mincer equation.

The coefficients in the Mincer regressions are usually different for men and women, indicating that they are paid differently even when having the same observable characteristics (education, age, location). To further investigate this point we simulate the counterfactual wage that men would earn if they were paid like women. The last column in sheet *Mincer2* reports the ratio between the average of this simulated wage and the actual average wage for men. Cases in which this ratio is less than one reflect the fact that women earn less than men even when controlling for observable characteristics.