

MESEP / MISSION FOR THE SPLICING OF
EMPLOYMENT, POVERTY AND INEQUALITY SERIES

Monetary Poverty in Colombia: New methodology and 2002-2010 figures

MESEP's 2nd phase results

National Planning Department – DNP
National Administrative Department of Statistics – DANE

February 2012

Hernando José Gómez/
Esteban Piedrahita
(DNP Director)
Juan Mauricio Ramírez
(DNP Deputy director)

Jorge Bustamante/
Héctor Maldonado
(DANE Director)
Christian Jaramillo/
Carlos Sepúlveda
(DANE Deputy director)

Committee of Experts

Jorge Iván González – Universidad Nacional de Colombia
Manuel Ramírez – Universidad del Rosario Colombia
Carlos Eduardo Vélez – Independent consultant
Mauricio Santa Maria – Fedesarrollo
João Pedro Azevedo – World Bank
Juan Carlos Feres – ECLAC

Roberto Angulo
(Technical Secretary)

DNP Technical Team

José Fernando Arias
Director of Social Development

Guillermo Rivas
Alejandro Mateus
Sandra Alvarez
Yadira Díaz
Rosa Rodríguez
Francisco Espinosa

DANE Technical Team

Eduardo Efraín Freire Delgado
*Director of Methodology and Statistical
Production*

Diana Carolina Nova
Rocío Bermúdez
Claudia Gordillo
Sandra Guayazán
Carlos Arturo Mora
Andrés Castañeda
Bernardo Atuesta
Diana Esperanza Sánchez
Silvia Esperanza Botello
Daniel Valderrama

ECLAC Technical Team

Juan Carlos Ramírez
Olga Lucía Acosta

Acknowledgements:

Javier Herrera, Xavier Mancero, Darwin Cortés, Jairo Nuñez, Manuel Muñoz, Luis Carlos Gómez and Irma Parra.

Translation team:

Proofreading in Spanish: Victor Andrés Arévalo
Translation: Juliana Mosquera, Ana Lucía Martínez
Proofreading in English: Ximena Díaz

Contents

Introduction

- 1. Background: Objectives and activities of MESEP**
 - 1.1 First Phase of MESEP: Splicing of Employment, Poverty and Inequality Series
 - 1.2 Second Phase of MESEP: Construction of new poverty lines
 - 1.3 Composition and working methodology of MESEP

- 2. Construction of the *income* variable**
 - 2.1 Construction of the aggregate income
 - 2.2 Correction by omission and debugging of the income report for WAP
 - 2.2.1 Missing Values
 - 2.2.2 Extreme Values
 - 2.2.3 Correction of "False Zeros"
 - 2.3 Correction by omission and imputation of false zeros and extreme values: *Hot Deck*
 - 2.4 Imputation of homeownership
 - 2.5 Per capita income of the spending unit
 - 2.6 Income underreporting and the adjustment methodology to National Accounts
 - 2.7 Summary: Differences with the previous methodology

- 3. Assessment of the source of information for the construction of the new poverty lines: The National Survey of Income and Expenditure (ENIG) for the 2006-2007 period**
 - 3.1 Internal consistency
 - 3.2 External consistency
 - 3.3 Implications of the consistency assessment of ENIG for the 2006-2007 period

- 4. Alternative methodologies for the construction of poverty lines**
 - 4.1 Food Energy Intake Method
 - 4.2 Cost of Basic Needs Method
 - 4.3 Alternatives to fix inconsistencies
 - 4.4 Methodology applied by ECLAC

- 5. Methodology applied methodology for the construction of new poverty lines**
 - 5.1 Construction of per capita current expenditure at the spending unit level
 - 5.2 Construction of the Spatial Price Deflator (SPD)
 - 5.3 Household sorting by percentile of deflated current expenditure
 - 5.4 Application of the iterative method for the selection of the reference population
 - 5.5 Construction of basic food baskets
 - 5.6 Adjustment of the basic food basket to meet calorie requirements
 - 5.7 Assessment of the adjusted basic food baskets
 - 5.8 Transfer from IL to PL using the Orshansky coefficient
 - 5.9 Considerations with respect to the construction of the new poverty lines
 - 5.10 Estimation of the incidence of poverty
 - 5.11 Validity verification of the estimates and the selected reference population
 - 5.12 Update of the poverty lines due to annual inflation
 - 5.13 Summary: Main characteristics of the new poverty line

- 6. Levels of poverty and inequality with the new methodology**
 - 6.1 Poverty for the 2002-2010 period
 - 6.2 Extreme Poverty
 - 6.3 Income inequality (Gini coefficient)

7 **The impact of methodological changes on the measurement of poverty**

8 **Conclusions and Recommendations**

9 **Institutional recommendations**

Bibliography

Glossary

Annexes

Annex A.	Variables included in the <i>Extreme values</i> and <i>False Zeros</i> models
Annex B.	Criteria for the definition of cells for the Hot Deck imputation model, by income source
Annex C.	Multinomial logistic model to predict housing tenure
Annex D.	Review of ECLAC's methodology to construct poverty and extreme poverty baskets
Annex E.	Construction of the <i>per capita expenditure</i> variable in ENIG for the 2006-2007 period
Annex F.	Food items for constructing the Spatial Price Deflator (SPD)
Annex G.	Composition of the Urban and Rural Basic Baskets
Annex H.	Calorie requirements
Annex I.	Values of extreme poverty and poverty lines by domain
Annex J.	Poverty and Extreme Poverty Incidence rates for the 13 metropolitan areas
Annex K.	Poverty and Extreme Poverty Rates by Department 2002-2010.

Diagrams

Diagram 1.	Composition of MESEP's 2nd phase
Diagram 2.	Composition of the household survey population
Diagram 3.	Correction by omission and debugging of the income report: Four steps for obtaining the complete income matrix
Diagram 4.	Outliers detection process using quantile regression
Diagram 5.	Detection process of "false zeros" in salaries and profits using nonparametric discriminant analysis
Diagram 6.	Imputation process by <i>Hot Deck</i>
Diagram 7.	On-field application scheme of the housing module for ECH for the 2002-2005 period
Diagram 8.	Alternatives for the election of the Orshansky Coefficient

Graphs

Graph 1.	Income matrix by associated population (WAP)
Graph 2.	Calorie income function
Graph 3.	Value of the poverty line: Comparing Colombia to other Latin American countries
Graph 4.	Incidence of Poverty: National, Urban and Rural. 2002-2010
Graph 5.	National Poverty 2002-2010: New and previous methodology
Graph 6.	Urban and Rural poverty, 2002-2010: New and previous methodology
Graph 7.	Incidence of extreme poverty: National, Urban and Rural. 2002-2010
Graph 8.	National extreme poverty for the 2002- 2010 period: New and previous methodology
Graph 9.	Urban and Rural extreme poverty, 2002-2010: New and previous methodology
Graph 10.	Gini Coefficient: National, Urban and Rural. 2002-2010
Graph 11.	Income inequality (Gini coefficient) 2002-2010: New and previous methodology
Graph 12.	Urban and rural income inequality, 2002-2010: New and previous methodology
Graph 13.	Decomposition of poverty measurement: Differences between the two methodologies
Graph 14.	Decomposition of extreme poverty measurement: Differences between the two methodologies

Tables

Table 1.	MESEP's working methodology
Table 2.	<i>A priori</i> and <i>a posteriori</i> distribution of "zero" and "non-zero" income
Table 3.	Spatial Price Deflator (SPD)

Table 4.	Average Calorie Requirement by Domain
Table 5.	Extreme poverty lines by Domain (by month) in Colombian Pesos of March 2007
Table 6.	Endogenous Orshansky Coefficients and monthly per capita poverty lines by Domain
Table 7.	Endogenous Orshansky coefficients for Latin American countries, ECLAC
Table 8.	Exogenous Orshansky coefficients and poverty lines by month (in Colombian pesos of March 2007)
Table 9.	Orshansky Coefficients by country
Table 10.	Geographic Domains of GEIH
Table 11.	Price index weights for <i>Remaining townships and Rural</i> domains
Table 12.	Poverty 2010: National, Urban, Rural, 13 metropolitan areas (MA) and Remaining urban areas
Table 13.	Changes in poverty level 2002-2010: Comparison between new and previous methodology
Table 14.	Changes of the extreme poverty level for the 2002-2010 period: Comparison of new and previous methodology
Table 15.	Decomposition of extreme poverty and poverty measurements: Differences between the two methodologies

Acronyms

Acronym	Definition in Spanish	Definition in English
ANA	Ajuste del Ingreso a Cuentas Nacionales	Adjustment of Income to National Accounts
CBN	Costo de las Necesidades Básicas	Cost of Basic Needs
CDI	Ingreso Corriente Disponible	Current Disposable Income
CPI	Índice de Precios al Consumidor	Consumer Price Index
DANE	Departamento Administrativo Nacional de Estadística	National Administrative Department of Statistics
DNP	Departamento Nacional de Planeación	National Planning Department
EAP	Población Económicamente Activa	Economically Active Population
ECH	Encuesta Continua de Hogares	Continuous Household Survey
ECLAC	La Comisión Económica para América Latina y el Caribe	The Economic Commission for Latin America and the Caribbean
ECV	Encuesta de Calidad de Vida	Quality of Life Survey
EIP	Población Económicamente Inactiva	Economically Inactive Population
ENIG	Encuesta Nacional de Ingresos y Gastos	National Income and Expenditure Survey
FAO	La Organización de las Naciones Unidas para la Alimentación y la Agricultura	The Food and Agriculture Organization of the United Nations
GEIH	La Gran Encuesta Integrada de Hogares	The Great Integrated Household Survey
ICA (deduction)	Impuesto de Industria y Comercio, Avisos y Tableros	Industry and Commerce Tax and Supplementary Notice and Boards
IES	Encuesta de Ingresos y Gastos	Income and Expenditure Survey
IiK	Ingreso en Especie	Income in Kind
IL	Línea de Pobreza Extrema	Extreme Poverty line
ILO	Organización Internacional del Trabajo	International Labour Organization
IOS	Ingreso de otras fuentes	Income from other sources
ISA	Ingreso de la Segunda Actividad	Income of the Second Activity
ISU	Ingreso Corriente Disponible de la Unidad de Gasto	Current Disposable Income of the Spending Unit
LCI	Índice de Condiciones de Vida	Living Conditions Index
MA	Áreas Metropolitanas	Metropolitan Areas
MERPD	Misión para el Diseño de una Estrategia para la Reducción de la Pobreza y la Desigualdad	Mission for the design of a Strategy for Reducing Poverty and Inequality
MESEP	Misión para el Empalme de las Series de Empleo, Pobreza y Desigualdad	Mission for the Splicing of Employment, Poverty and Inequality Series
MIFA	Ingreso Monetario de la Primera Actividad	Monetary Income of the First Activity
MIUI	Ingreso Monetario Desempleados e Inactivos	Money Income of the Unemployed and inactive
NPDA	Análisis Discriminante No Paramétrico	Non-Parametric Discriminant Analysis
OC	Coeficiente de Orshansky	Orshansky coefficient
PCISU	Ingreso Per Cápita de la Unidad de Gasto	Per Capita Income of the Spending Unit
PL	Línea de Pobreza	Poverty Line
PPP	Paridad de Poder Adquisitivo	Purchasing Power Parity
SGSSS	Sistema General de Seguridad Social en Salud	General System of Social Security in Health
SNA	Sistema de Cuentas Nacionales	System of National Accounts
SOAT	Seguro Obligatorio de Accidentes de Tránsito	Mandatory Traffic Accident Insurance

SPD	Deflactor Espacial de Precios	Spatial Price Deflator
TI	Ingreso Total	Total Income
TIFW	Ingreso Total Trabajadores Familiares sin Remuneración	Total Income Unpaid Family Workers
TIIW	Ingreso Total Independientes	Total Income Independent workers
TIS	Ingreso Total Asalariados	Total Income Salaried Employees
TIUI	Ingreso Total Desempleados e Inactivos	Total Income Unemployed and Inactive
VAT	Impuesto al Valor Agregado	Value Added Tax
WAP	Población en Edad de Trabajar	Working Age Population
WHO	Organización Mundial de la Salud	World Health Organization

Introduction

This document presents the new methodology for measuring monetary poverty in Colombia, which adopts changes both in the poverty line as well as in the construction of the aggregate household income.

The methodology provides an updated and more accurate measurement, in technical and empirical terms, of poverty in Colombia with respect to three aspects. First, the construction of the poverty line uses the latest statistical basis of the Colombians' consumer habits, the 2006-2007 National Household Income and Expenditure Survey (ENIG for its acronym in Spanish); (the previous methodology used the ENIG for the 1994-1995 period). Second, it incorporates recent methodological developments, widely accepted by international experts. Third, it uses a more accurate measurement of the aggregate income, excluding the adjustment to National Accounts that introduces an uncontrollable bias in the distribution of income and the composition of poverty.

On the other hand, the new methodology for measuring poverty relates to a greater extent to current methodologies of other Latin American countries, facilitating its comparability in the regional context.

In addition, one of the most important sub-products of MESEP was the institutionalization of methodologies and calculation algorithms. The computer algorithm used to generate estimates is now shared by: the National Administrative Department of Statistics (DANE for its acronym in Spanish), the National Planning Department (DNP for its acronym in Spanish), the World Bank and the Economic Commission for Latin America and the Caribbean (ECLAC). In addition, a mechanism for the dissemination of this information to the general public will be designed.

Measurements of poverty based on the new methodology show trends that are similar to those of the previous methodology (Mission for the Splicing of Employment, Poverty and Inequality Series – MESEP Phase 1). However, the rates of the new series differ from the previous ones. Poverty and extreme poverty achieve lower levels (the latter with a higher rural component), whereas income inequality measured by the Gini coefficient is slightly lower, but very similar.

In order to guarantee the transparency of the measurements of poverty, the series using both methodologies will be published for a period of at least two years. With the same purpose, DANE will make available to users the household surveys used by MESEP for its measurements.

In the first section, this paper presents the new methodology for measuring monetary poverty as well as extreme poverty, poverty and inequality figures for the 2002-2010 period. The second section outlines the objectives and activities of MESEP. The third section describes the procedure for constructing the aggregate household income. The fourth section presents an analysis of the survey used to build the new lines. The fifth section briefly reviews some methodologies commonly used to construct poverty lines. The sixth section presents the methodology adopted to define new poverty lines in Colombia. Sections seven and eight present updated 2010 figures and the impact of methodological changes respectively. Section nine provides some conclusions and recommendations of MESEP's second phase.

1. Background: Objectives and activities of MESEP

As of the second semester of 2006 a problem arose regarding the comparability of employment, poverty and inequality series as a result of the methodological change concerning the transition from the Continuous Household Survey (ECH for its acronym in Spanish) to the Integrated Household Survey (GEIH for its acronym in Spanish)¹. In order to address these difficulties, which compromised the credibility of key indicators of the labor market and living conditions in Colombia, DANE and DNP began to work from a technical perspective, giving rise to the Mission for the Splicing of Employment, Poverty and Inequality Series (MESEP, January 2009). MESEP worked in two phases that are summarized below.

1.1 First Phase of MESEP: Splicing of Employment, Poverty and Inequality Series

In its first phase (January to September 2009), the MESEP carried out the splicing of employment, poverty and inequality figures in order to obtain comparable series for the 2002-2008 period. This splice was in response to the methodological change from ECH to GEIH, which affected the inter-temporal comparability of the indicators as of the second semester of 2006.

The following results were achieved by means of the technical work undertaken by DANE and DNP in November 2007, and which continued throughout the first phase of the MESEP:

- *Identification* of the factors that disturbed the comparability of the employment, poverty and inequality series.
- The *adoption*, by DANE, of appropriate corrective measures for data collection operations with respect to the 2008 GEIH survey, which resulted in the improvement of the statistical quality of the surveys: ECH for the 2002-2006 period and GEIH for the 2006-2007 period.
- The *production* of the spliced Employment, Poverty and Inequality series for the 2002-2008 period. In 2010, MESEP presented the 2009 updated series along with the methodology prevailing at the time. While these estimates included some changes to estimate the incidence of poverty, the values of the poverty lines based on data from the 1994-1995 Income and Expenditure Survey (IES) were essentially preserved.

1.2 Second Phase of MESEP: Construction of new poverty lines

¹ For a detailed description of the implemented changes, refer to MESEP (2009).

During its second phase (September 2009 to August 2011) MESEP focused on the design of the new methodology for measuring monetary poverty. This work enabled:

- *Constructing* the new poverty and extreme poverty lines, using the ENIG for the 2006-2007 period.
- *Reviewing and defining* other elements that compose the methodology for measuring monetary poverty in Colombia (imputation of income, updating the values of the lines, etc.).
- *Estimating* official poverty, extreme poverty and Gini figures for the 2002-2010 period.
- *Proposing* an institutional arrangement for estimating and officializing poverty figures (data as of 2011 and onwards).

1.3 Composition and working methodology of MESEP

Three groups formed MESEP: a decision-making group, an experts group and a technical group. Both the DNP Director at the time, Hernando José Gómez (preceded by Esteban Piedrahita) and the DANE Director Jorge Bustamante (preceded by Héctor Maldonado) established the decision-making group, leader of MESEP. The members of the experts group were: Juan Mauricio Ramírez (DNP Deputy Director); Christian Jaramillo, current DANE Deputy director (preceded by Carlos Sepúlveda); Jorge Iván González (Universidad Nacional de Colombia); Manuel Ramírez (Universidad del Rosario, Colombia); Carlos Eduardo Vélez (independent consultant), Mauricio Santa María² (FEDESARROLLO³), João Pedro Azevedo (World Bank representative) and Juan Carlos Feres (ECLAC representative). The technical group comprised professionals from both DANE and DNP.

² Mauricio Santa María resigned when he was appointed Minister of Social Welfare in 2010. He participated from September 2009 to August 2010

³ Foundation for Higher Education and Development

Diagram 1. Composition of MESEP’s 2nd phase



Source: MESEP’s technical team

The two phases of MESEP were developed through work sessions in which the results of the exercises proposed by the experts group were presented. DNP and DANE teams simultaneously reviewed the exercises in order to ensure consistency in the results and develop technical skills in both entities. The experts group was responsible for validating the results and making the decisions that led to the final version of the methodology for measuring monetary poverty.

During the second phase of the MESEP, 25 working sessions were conducted and 35 technical reports prepared jointly by the technical teams of both DNP and DANE were presented.

The agreements and the final version of the methodology for measuring poverty in Colombia are described in the following sections.

Table 16. MESEP's working methodology

2nd phase period	September 2009 – August 2011
Working sessions	25
Presentations	35
Methodology	<ol style="list-style-type: none">1. Working sessions of the experts group to discuss reports and analyses prepared by the DANE-DNP technical group and international experts.2. Reaching methodological agreements based on the technical consistency of statistical processes and in the coherence of empirical results.

Source: MESEP's technical team

2. Construction of the *income* variable

Two elements are required for the measurement of monetary poverty: the value of the poverty lines and the disposable per capita income of the spending units⁴. The latter is obtained from a set of questions that are included in the forms of household surveys. In order for this measurement to represent a plausible structure of the flow of household income, a set of items that provides a consistent approach with the disposable income that these units use to meet their expenses must be considered in its definition.

ECH, and its further version, GEIH are multipurpose tools that have had the measurement of labor market indicators as their central purpose. They also enable obtaining information to calculate indicators of income distribution, unsatisfied basic needs, living conditions index and other aspects of these conditions in households.

The measurement of income in household surveys has been marked by an emphasis on labor income, perhaps following the reasonable assumption that most of the household income in developing countries comes from labor income (Grosh and Glewwe, s. f.: 219). Other components of the total income of the household spending unit are collected with a lower level of detail. Due to the above reasons, household surveys imperfectly capture the wealth and assets of households. It should be noted that in recent years the battery of questions focusing on income have shown significant improvements in all items, whether they are labor or non-labor income.

The first step in the construction of the *income* variable consists of estimating the aggregates for each of its sources. Second, an imputation model is applied to correct biases⁵ on information reported by the population in household surveys. Third, a consolidation of the total income of the working-age population (hereinafter WAP) is obtained, which is then aggregated at the spending unit level. Fourth, depending on the status of housing tenure, an imputation for its ownership is made. Finally, the value of the per capita income of the spending unit is calculated, from which the incidence of poverty and extreme poverty is calculated.

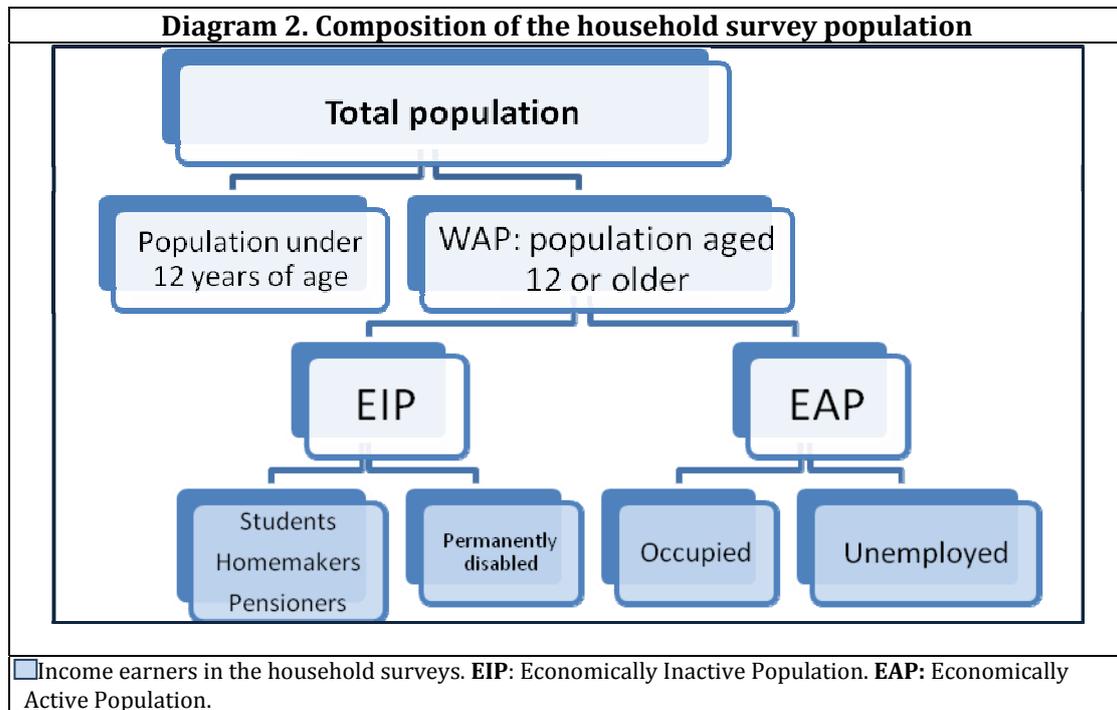
⁴ Pensioners, workers, domestic employees and their relatives are not part of the spending unit.

⁵ It should be clarified that the correction by omission does not necessarily correct all biases in collected data. Nonrandom omission, which is handled by the imputation technique, is only one possible source of bias.

The following sections describe each of the steps for obtaining the per capita income of the spending unit.

2.1 Construction of the aggregate income

Two forms are used for the construction of the aggregate income for the 2002-2010 period. These forms resulted from the changes in the survey between the 2002-2006 and 2006-2010 periods. The changes were indicated in section 2. The purpose is to build the *income* variable for all earners that make up the WAP, taking into account the differences between the groups that comprise it, mainly the disaggregation among the *economically inactive population* (hereinafter EIP) and the *economically active population* (hereinafter EAP), and the division of each of these categories into other groups, as described in Diagram 2⁶.



Source: MESEP's technical team

For the purpose of constructing the *income* variable, the individuals who make up WAP are grouped into four categories of income earners: employee, independent (self-employed and employers), unpaid family workers, as well as unemployed and inactive. The individual income

⁶ Such disaggregation corresponds to that used for the measurement of labor market indicators, based on the recommendations of the International Labour Organization (ILO). In Colombia, the WAP corresponds to the population aged 12 years or older in urban areas and 10 years or older in rural areas.

corresponding to each of these categories is composed of at least one of the following types of income:

1. *Monetary Income of the First Activity (MIFA)*: In the period prior to the second quarter of 2006, this information on MIFA was obtained by a single question on the form; consequently it did not disaggregate monthly monetary income of employees corresponding to: wages and salaries, benefits, overtime, bonuses and travel allowances. As of the second quarter of 2006, the form used in the survey disaggregated this heading according to the implicit items mentioned in the above definition: monthly wage income, overtime, benefits (food, transportation, family or education), bonuses (technical, seniority, climate, public order, etc.), monthly bonuses, and annual bonuses (Christmas, holidays, etc.). For independent workers it includes the net income or professional fees of the first activity.
2. *Income in kind (IiK)*: corresponds to food, housing, transportation, and others such as Sodexo coupons or home appliances⁷. It applies only to employees.
3. *Income of the Second Activity (ISA)*: It applies to all employed persons (employees, independent and unpaid family workers). Includes income in cash or in kind.
4. *Money Income of the unoccupied and inactive (MIUI)*: Includes income from work done by the unoccupied and inactive in periods prior to the one of reference.
5. *Income from other sources (IOS)*:
 - IOS1: Leases (effective rents)
 - IOS2: Investments interests and dividends
 - IOS3: Old age, disability pensions or substitution pensions
 - IOS4: Household benefits⁸ within and outside the country, and of institutions, alimony, in cases of paternity, divorce or separation.
 - IOS5: Severance pay and severance pay interests⁹.
 - IOS6: Occasional profits.

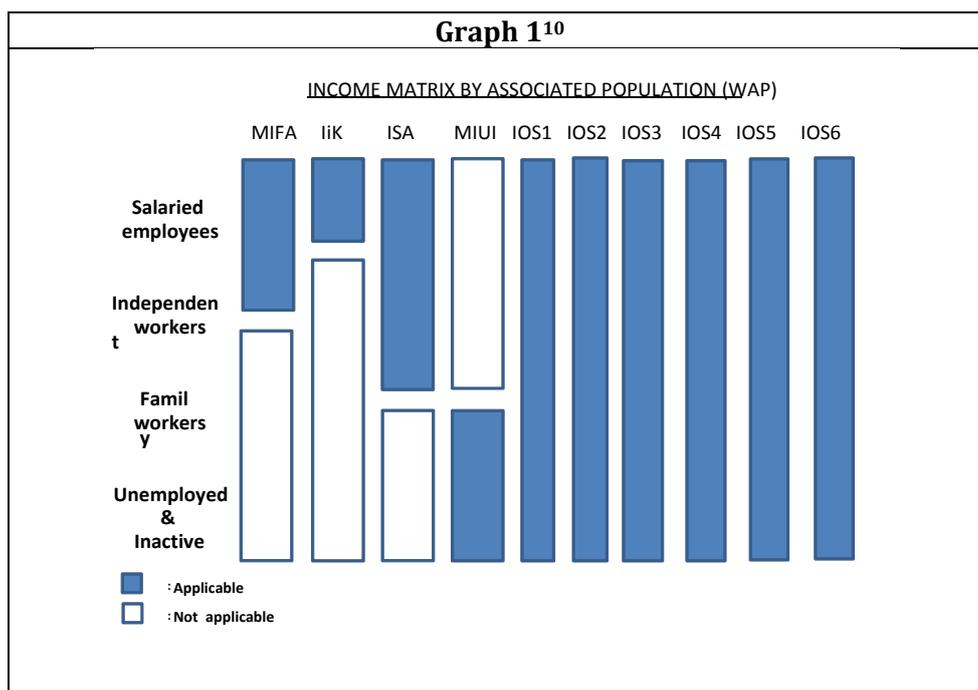
⁷ Before the second quarter of 2006 the question regarding income in kind was variable; many times it only included the concepts of housing and transportation.

⁸ From 2002 up to the second quarter of 2006 this item inquired on benefits received by households for any reason, including government benefits. As of August 2006, the household survey inquired on benefits of households or non-governmental institutions. Information on government benefits was asked in the household module, but its permanence was not continuous, therefore its inclusion was dismissed in the composition of income. In the module of other income for WAP, benefit restrictions of non-governmental organizations were abolished since the third quarter of 2010, and replaced by benefits from institutions in general, without being able to discriminate the source of the money, whether governmental or not.

⁹ These concepts were included only in the GEIH form.

Income from Other Sources (IOS) corresponds to the sum of the six defined income types.

The types of income that are incorporated into the calculation of the total income for each of the types of identified earners are illustrated in the income matrix (refer to Graph 1) and are defined as follows:



Source:

technical team

MESEP's

- a) Salaried Employees: includes workers, employees and domestic workers. Its corresponding formula is:

$$\text{Total Income Salaried Employees (TIS)} = \text{MIFA} + \text{IiK} + \text{ISA} + \text{IOS}$$

- b) Independent workers: Includes all those self-employed workers and employers. It is calculated by using the following formula:

¹⁰ Once a definition of current disposable income of individuals in WAP was established, an income matrix of this population is implicitly specified. In this arrangement, the columns will be the sources that compose the current disposable income of the adopted definition. In turn, the rows of the matrix will be the individuals in WAP. This matrix will need adjustments in some of its cells, because persons did not report any data in them, and/or what they reported as such can be labeled as "aberrant" or "extreme" values versus expected patterns of income. Thus, once the income matrix is defined (WAP), it is necessary to "complete" and "adjust" the information in some of its cells, in order to obtain the "complete" matrix. This procedure is called correction by omission.

$$\text{Total Income Independent workers (TIIW)} = \text{MIFA} + \text{ISA} + \text{IOS}$$

- c) Unpaid family workers: It is calculated according to the formula:

$$\text{Total income unpaid family workers (TIFW)} = \text{ISA} + \text{IOS}$$

- d) Unemployed and Inactive: the corresponding formula is:

$$\text{Total income unemployed and inactive (TIUI)} = \text{MIUI} + \text{IOS}$$

The *current disposable income* (CDI) is equal to the total income (according to the sum of income sources that apply for each of the types of earners described above) less other income items corresponding to *occasional profits and severance interests*.

The aggregation of the CDI of all household earners constitutes *the current disposable income of the spending unit* (ISU):

$$SUI = \sum_{i=1}^n CDI$$

Where i represents each one of the income earners of the spending unit and CDI corresponds to the formula:

$$CDI = TI - \text{occasional profits and severance pay interests}$$

In turn, total income (TI) is expressed as:

$$TI = YIS + TIIW + TIFW + TIUI$$

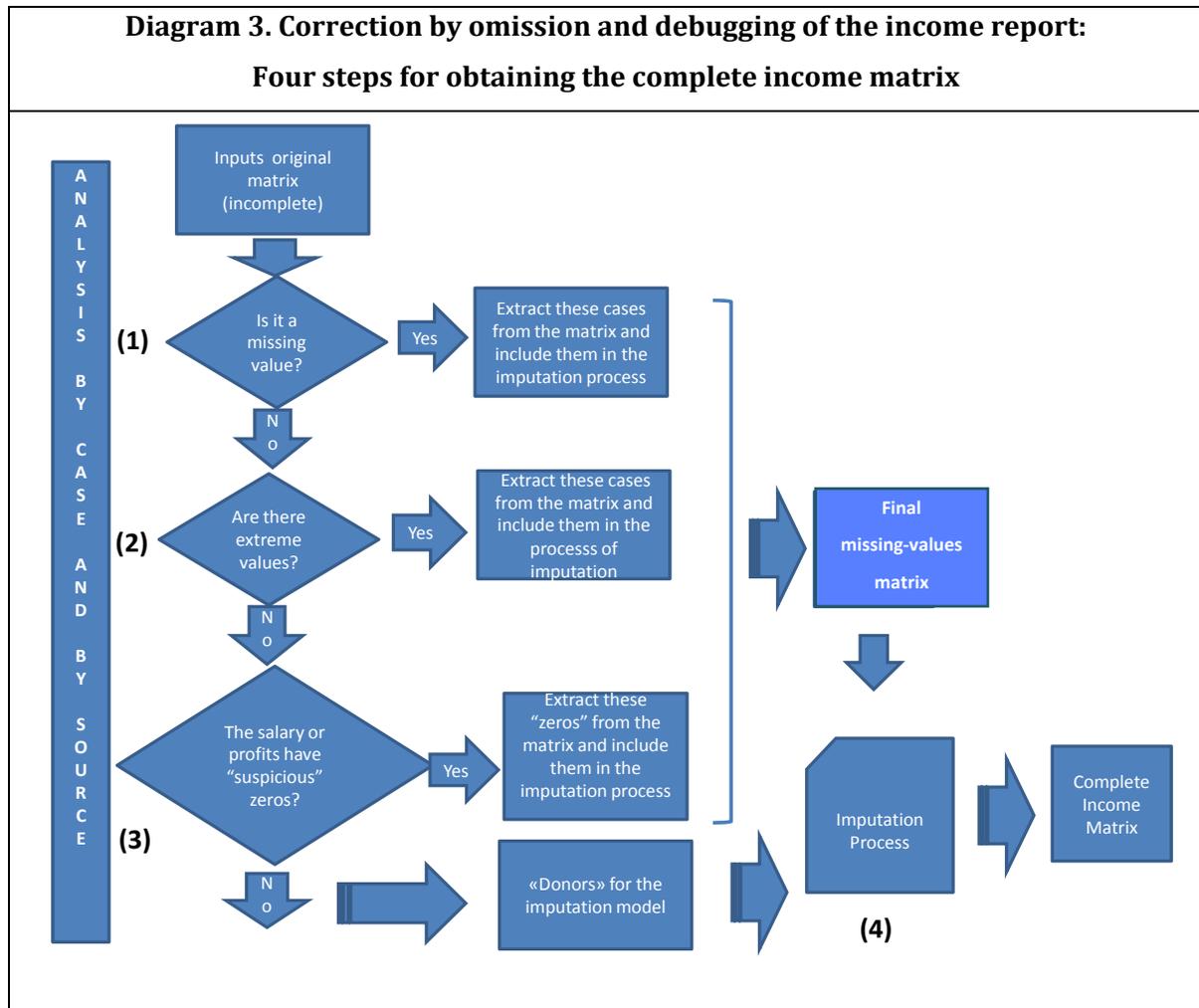
Where the addends correspond to the types of income listed above.

Before calculating the income per capita of the spending unit, some potential biases in the collected data of household surveys must be identified and corrected. This procedure is called correction by omission and it is described below.

2.2 Correction by omission and debugging of the income report for WAP

The correction by omission and debugging of the income report, for each of the individuals in WAP, allows obtaining the complete income matrix and requires four steps that are explained in

Diagram 3: in the first step, *missing values* that come in the original survey are identified; in the second step, the *extreme observations* in each source of income are identified; in the third step, the salaries or profits which have a *suspicious zero value* are identified; and finally, the *missing-data imputation model* is applied to the cases selected in the previous three steps.



2.2.1 Missing Values

In the context of household surveys, a missing value is identified by a code, which indicates that the information to be recorded is not held nor provided by the informant.

In order to establish the records with missing values in the income variables the following criteria are used:

- In all sources, the cases reporting the codes 98, 99, 999, 9999, 99999, 999999, 9999999, 99999999, 999999999, 9999999999, in the field corresponding to the income value.
- If prior to the question pertaining to an income value, it is inquired whether or not the person received such income, the following are taken as missing values: persons who respond

the option *does not know/no answer* (code 9), or those who respond affirmatively (code 1) and do not report information in the field corresponding to the value, or those reporting any of the codes included in the previous point.

- The sources that consist of the sum of various values are considered missing values if any of its components are missing, except for MIFA for employees, in which priority is given to salary; if information on salary was reported, the value is considered as observed even if it has missing values in any of the other components (overtime, bonuses, etc..)

2.2.2 Extreme Values

The procedure for identifying "extreme" observations¹¹ is based on the non-parametric optimization technique named *quantile regression*¹², which is illustrated in Diagram 4. Unlike the least squares technique, which minimizes the sum of *squared distances* between each value and its estimate, quantile regressions minimize the amounts of the *same distances* but at an *absolute value level*¹³ (optimizes an absolute distance between each value and the τ -th quantile)¹⁴. Thus a set of parameters for each quantile regression will be obtained. In other words, there will be a family of curves that optimize the criterion for each quantile. For this exercise, six regressions with the quantiles¹⁵ 10, 25, 50, 75, 85 and 95 were defined.

¹¹ It should be clarified that in this procedure, the values in the range (0.1) are taken directly as extreme values without being assessed through the quantile regression model.

¹² Annex A shows the explanatory variables that were considered for the specification of the regressions for each of the sources of income.

¹³ With the least squares method, a set of parameters $\beta = (\beta_0, \beta_1, \dots, \beta_p)$ that produces the linear combination that minimizes the

distance $\sum_{i=1}^n (y_i - [\beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi}])^2$ for a given set of data $(x_{1i}, x_{2i}, \dots, x_{pi})$ is obtained. Each observation y_i is compared with the linear combination and its squared distances are minimized by a single set of parameters.

By using the quantile regression, the weighted absolute deviation is minimized by asymmetric weights. That is, if the quantile has the proportion τ of observations below the quantile, there will be a ratio $1-\tau$ of observations above the quantile. The asymmetric weights are τ and $1-\tau$ in the expression subject to minimization. Likewise, the expression does not support the conventional method of minimization. Instead, linear programming methods are used.

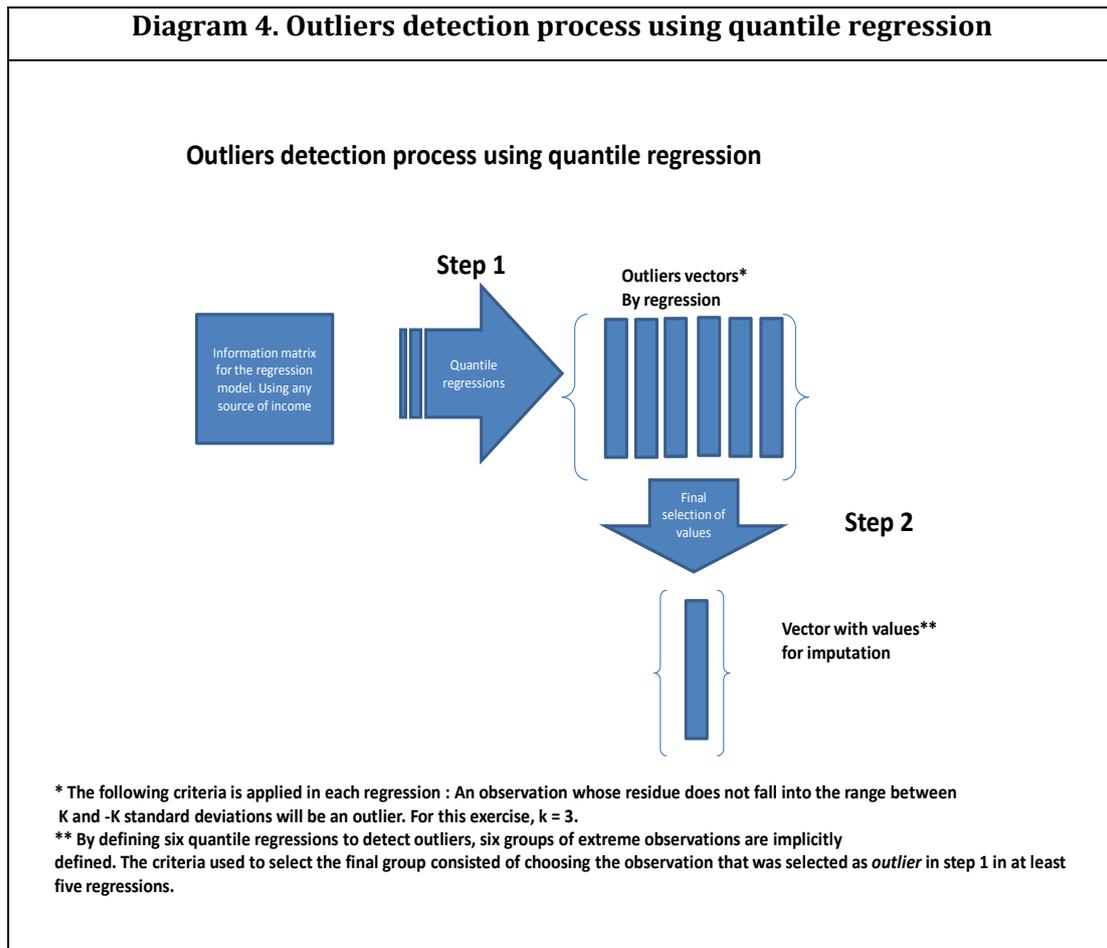
¹⁴ The τ -th quantile regression, $0 < \tau < 1$, is a solution to the minimization problem:

$$\min_{\beta \in \mathbb{R}^p} \left\{ \left[\sum_{i: y_i \leq x_i' \beta} \tau |y_i - x_i' \beta| \right] + \left[\sum_{i: y_i > x_i' \beta} (1-\tau) |y_i - x_i' \beta| \right] \right\}$$

This solution is denoted by $\hat{\beta}(\tau)$. For each quantile τ there will be a set of parameters $\hat{\beta}(\tau)$.

¹⁵ The prevalence of extreme values occurs mainly towards the "tails" of the income distribution. The selection of the number of quantiles should facilitate the identification of such values throughout the entire distribution, with special emphasis on the mentioned areas. After several tests to determine the location and the number of quantiles, it was determined that the selection of the quantiles 10, 25, 50, 75, 85 and 95 was efficient to locate values along the range of the dependent variable.

Diagram 4. Outliers detection process using quantile regression



Source: MESEP Technical Team

The quantile regression technique generates as many outliers and non-outliers data groups (i.e., a group of atypical data) as regressions are defined. The existence of multiple *outliers* groups forces the inclusion of an additional criterion for the final selection. With respect to this procedure, it was decided to consider an outlier or extreme value the cases that resulted as *outliers* in at least *five out of six regressions*.¹⁶

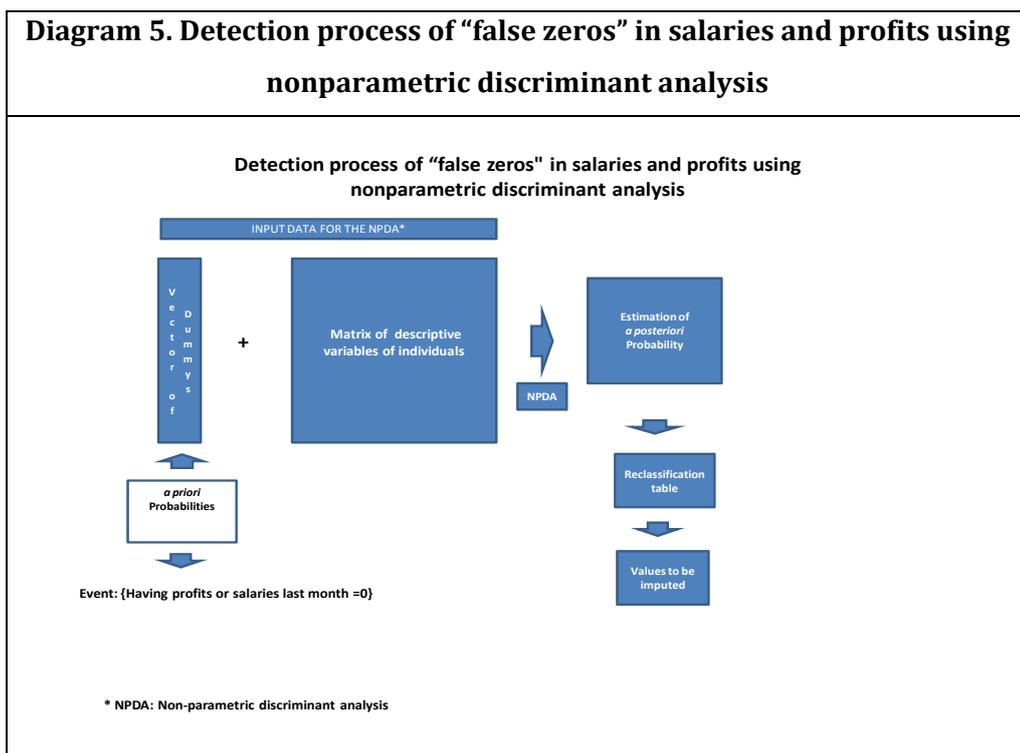
2.2.3 Correction of "False Zeros"

To detect if the values declared as "zero" by respondents that have MIFA, truly correspond to such value, a statistical criterion for classification named non-parametric discriminant analysis¹⁷ (NPDA) per stratum¹⁸ is applied. This statistical criterion divides the population into

¹⁶ With the method of ordinary least squares (OLS) a single set of outliers is obtained, while with the quantile regression methodology (RC) one can obtain as many sets as the number of regressions that are defined. Initially, it was decided that an observation would be classified as an extreme value if it was identified as such in each of the sets generated by the regressions (in this case, six sets). However, for some sources of income, the application of the above criterion did not produce detection of extreme values, i.e., the final set of extreme values was empty. For this reason, the criterion was loosened and those values, which on at least five of the six regressions proved to be extreme values, were considered as outliers in order to include values in the final set.

¹⁷ Annex A presents the explanatory variables used in the analysis.

two analysis groups: those reporting MIFA equaling zero and those reporting MIFA different from zero¹⁹. Initially, the proportions of cases for these two groups (*a priori* information) are obtained. Subsequently, by using the nonparametric model, the *a posteriori*²⁰ probabilities that a case has MIFA income equaling zero given certain information (being it general, personal, work-related, home-related or geographical) are estimated (refer to Diagram 5). As a result of this process, a two-way table is built, which enables identifying to which "observed zeros" corresponds a high probability of being "not zeros" (top right cell in Table 2).



Source: MESEP technical team.

¹⁸ This classification corresponds to the sampling stratum of the ECH survey in 13 major cities and to the sextile of the Living Conditions Index (LCI) for other urban municipal townships and dispersed rural areas. In GEIH, it corresponds to the stratum of electrical power available in the housing module for all domains of analysis. The cases without information on this variable are classified according to the LCI sextile.

¹⁹ This procedure does not require the assumption of normality for vectors of variables.

²⁰ The decision rule can be summarized as follows: The element x_0 will be declared "zero" if:

$$\pi_0 f_0(x_0) > \pi_1 f_1(x_0)$$

Where π_0 and π_1 are the *a priori* probabilities (or proportions observed in each population subgroup) of being "zero" or "different than zero", respectively. On the other hand, f_0 and f_1 are density functions associated with the subpopulations of interest, namely those which are zero and different from zero in MIFA. In this case, since it is a non-parametric analysis, the density functions are replaced by the Kernels. In short, an individual is classified in the group in which it obtains the highest *a posteriori* probability (the probability for a given element x_0 calculated on density functions f_0 and f_1 , i.e. $\pi_0 f_0(x_0) > \pi_1 f_1(x_0)$).

Table 17. <i>A priori</i> and <i>a posteriori</i> distribution of "zero" and "non-zero" income				
			Estimates	
			Zeros	Non-zeros
O B S E R V E D	Zeros	OK	False zeros	
	Non-zeros	Error	OK	

Source: MESEP technical team.

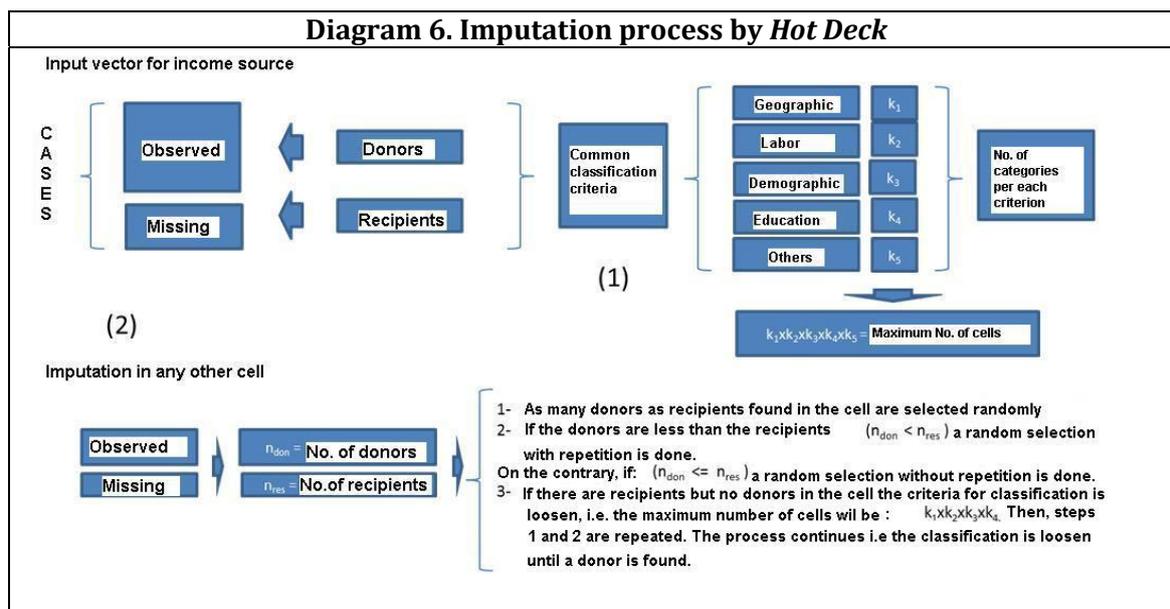
2.3 Correction by omission and imputation of false zeros and extreme values: *Hot Deck*

Considering the information above, the cases that will finally be imputed are found: cases declared as *outliers* (extreme values); cases classified as "false zeros" and the original missing values that came in the survey²¹. To perform this procedure an imputation model is used: the *Hot Deck*. (refer to Diagram 6)²²

²¹ The incidence of missing values within WAP has decreased with the improvement of field operations. In the 2002 ECH, 15.7% of income records of WAP corresponded to missing values. In 2005, the percentage dropped to 10.2%. When starting GEIH, this percentage dropped drastically. In 2008, the missing values represented 3.5% of income records of WAP. In 2010 this incidence was 3.2%. Also, the importance of imputed income in the total income has also declined significantly. Imputed income in 2002 represented 23.4% of total income. In 2010, this percentage dropped to 4.5%.

²² This method was chosen after evaluating five methodologies: Multiple Imputation, Maximum likelihood, Traditional (which was being used by DNP), traditional by stratum and *Hot Deck*. For the selection of the imputation model the following steps were followed:

1. A master database for trials was built (resulting from joining in a single database the records of the years 2008 and 2009 of GEIH) that contained only records of individuals with reported incomes. Households that had at least one member with missing information in any of the sources of income were excluded.
2. On this basis random patterns of missing values per stratum were generated according to the following classification criteria: geographical area (13 metropolitan areas, rural and other municipal townships), occupational status, source of income and design stratum (for the cases without stratum, Living Conditions Index (LCI) sextiles were generated). Thus, a random set of observed income values were replaced by simulated missing values. The reference standard that was followed in order to simulate missing values was the one observed in each of the surveys of 2002, 2003 and 2008.
3. Then, by using the five different alternative imputation models, an approximation of an income value for each one of the cases that were simulated as missing values was made.
4. The difference between the imputed value for each model and the original value observed in the survey was estimated.
5. The selected model was the one that generated the smallest differences between the imputed values and the original values.



Source: MESEP technical team.

This methodology classifies relevant data for each source of income according to categories that differentiate them. For example, MIFA is different if the people who receive it belong to groups that differ by sex, educational level, occupational status, mastery of the survey, socio-economic strata, head of household or age groups.

For any source of income, individuals are classified as "donors" and "recipients".²³ Donors are those who reported their income and that in the two previous steps were not declared *outliers* or "false zeros" (only in the case of MIFA). The recipients are the ones that did not declare income in the survey, or were declared *outliers* or "false zeros" (only in the case of MIFA).

The income of the donors is the input to impute that of the recipients. The procedure is as follows:

Donors and recipients are distributed in the various cells that are defined from variables that differentiate income²⁴. If in a cell the number of donors is higher than the number of recipients, a number of donors equal to the number of recipients will be chosen randomly without

²³ Each source of income has, by definition, a population for which such item applies, which is why the number of donors and recipients is restricted to that definition.

²⁴ The specification of cells for the case of MIFA is as follows. The domains or categories of the survey defined for the exercise are three: 13 metropolitan areas, other municipal townships and remaining rural areas. Likewise, sex has two categories; educational level has three (elementary or less, secondary and higher), head of household has two; socioeconomic stratum has six; occupational status has five categories (workers, employees, domestic service, employers and self-employed workers), and age group has 4 (under 18, 18-24, 25-45, 46 and older). Therefore, there will be $3 \times 6 \times 4 \times 3 \times 5 \times 2 \times 2 = 4320$ cells. An example of a cell can be heads of household, men, workers, with secondary education, who are between 24 and 45 years old, in stratum three of the thirteen metropolitan areas. The specifications for other sources of income other than MIFA are found in Annex B.

repetition and the latter will be assigned the income values of the selected donors. On the contrary, if the number of donors is lower than the number of recipients, the random sample of income from donors will be selected with repetition. If receptors are found in any cell, but not donors, new cells will be generated by the omission of a variable in the classification, and the process of random allocation previously described is repeated.²⁵

Once the cases requiring it have been imputed, the construction of the income aggregate at the spending unit level is made.

As noted above, once some of the potential biases of the various sources of income are corrected, these are added to obtain the current disposable *income of the spending unit* (ISU).

After the estimation of ISU, the housing tenure status by the household must be identified in order to determine if an imputation for ownership of the home has to be made. The process of income imputation for homeownership is described below.

2.4. Imputation of homeownership

The imputation process of homeownership is a homogenization of the income that enables comparisons between households with different housing occupancy status. The payments for rent or for mortgage amortization differentiate the disposable income of a household that makes these payments compared to the disposable income of a household that is a homeowner or lives in usufruct.²⁶

Additionally, it should be considered that the calculation of the incidence of poverty involves two elements: the value of the poverty line and the per capita income of the spending unit, where each one of these must refer to the same items. Among the expenditure items considered for the construction of the poverty line, an amount for imputed rent is included (see section 6.1). For this reason, it is necessary to also consider the inclusion of the amount for rent imputation for owners within the household income in order to ensure consistency when estimating the proportion of persons in poverty conditions.

²⁵ Following the previous example, losing the classification criteria means that for that cell, the heads of household will no longer be considered; in its place, the cases will be classified as men, workers, with secondary education, who are between 24 and 45 years old, in stratum three of the thirteen metropolitan areas.

²⁶ In the case of usufruct, tenure can be assimilated as a transfer from another home.

Depending on the type of tenure, an imputation on ISU in the quality of homeownership is made. Spending units to which this imputation is made are the ones reporting the following categories of housing occupancy²⁷:

- Own home fully paid
- Own home still being paid
- Home in usufruct

According to the stage of the household survey, *two alternatives* were established to assign a value to homeownership for the above categories:

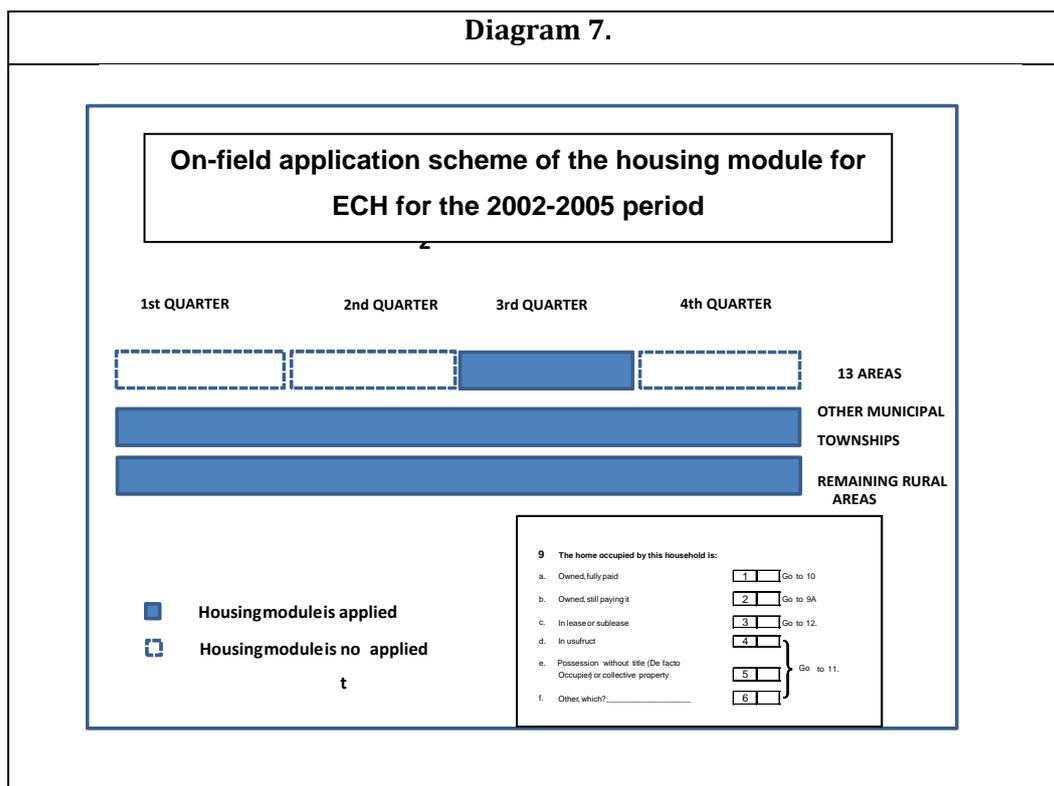
- For ECH for the 2002-2005 period a *Hot Deck* model was used.
- For GEIH for the 2008-2010 period, the question of household data module: "If you had to pay rent for this house, how much do you estimate you would have to pay monthly?" was used.

The stages of ECH do not include the question of expected rent, which is why the *Hot Deck* methodology was used to assign an imputed value per property. Additionally, in the 13 metropolitan areas, ECH for the 2002-2005 period includes only information on the housing module in the third quarter. For the domains *other municipal townships and rural*, the information on the housing module was collected for the four quarters of the year. (Refer to Diagram 7)

²⁷ Of the total households in the 2008 GEIH, 42.8% had their own home fully paid, 1.5% were paying their home and 13.4% was in a home in usufruct. The corresponding percentages in the 2010 GEIH were 42.4%, 2.3% and 13.7%, respectively. These sets of households were those that underwent an imputation for homeownership.

In 2010, the percentage of imputation for homeownership within the total income of the spending unit was around 15% for homeowners who already paid their home, 4% for those that are paying it, and 14% for those living in usufruct. This percentage is similar in the other years of household surveys.

Diagram 7.



Source: MESEP technical team.

In turn, GEIH includes information of such module for all quarters and for the three geographic domains (13 metropolitan areas, other municipal townships and rural).

The absence of the housing module in those 13 areas in the quarters I, II and IV of ECH for the 2002-2005 period involved using a multinomial logistic model to predict the probability of belonging to each category of housing tenure in those quarters (Annex C shows the description of the multinomial logistic model).

Once having the complete information on housing tenure in each of the stages and domains of household surveys, the imputation of homeownership was made according to the two alternatives listed above. The procedure in each case is described below.

Hot Deck model for imputation of homeownership in ECH for the 2002-2005 period: The Hot Deck methodology divides the population into two subgroups: donors and recipients. The information with respect to donors is used as an input to predict that of the recipients. In the case of imputation of homeownership, donors are those records identified as tenants. On the

other hand, recipients are the cases that in the question on housing tenure answered: Own home fully paid, Own home still paying²⁸ and Home in usufruct²⁹.

For the present exercise the following classification criteria of donors and recipients were defined: the metropolitan area, the stratum and the number of rooms. When combining these criteria, homogeneous cells that are heterogeneous among themselves are generated within them. The resulting cells differentiate the values in the rents significantly and allow bias minimization by building a complete data distribution.

From the *Hot Deck* method, the records are located within the generated cells. Subsequently, the number of donors and recipients per cell are compared. If there are more donors than recipients, a random selection without replacement is made to assign the rental value that will correspond to each recipient. If the number of donors is lower than the number of recipients, a random selection with replacement is made. Finally, if there are cells without donors the classification level is loosened and the process with these cases is repeated.

In each cell the imputed rental value that is assigned to recipients (owners and usufruct) will be equivalent to that of the donors (tenants) that have similar characteristics (metropolitan area, strata and the number of rooms).

Expected rental value for imputation of homeownership in GEIH for the 2008-2010 period: In GEIH, the owners are asked for the rent they would pay if they were tenants³⁰. This information was used to impute the rental value to the categories of owners (own home fully paid and own home still being paid) and to the households in usufruct³¹. The fact that this amount comes from the owners, makes it a suitable solution for the adjustment of income.

2.5 Per capita income of the spending unit

Once the ISU has been imputed by homeownership³², the *per capita income of the spending unit* (PCISU) is calculated. This value is the result of dividing the ISU³³ into the number of persons who compose the spending unit³⁴.

²⁸ It is necessary to note that the *Hot Deck* model was also used to impute the value of the mortgage amortization payments of *owners paying* for the periods without housing module.

²⁹ In the absence of the housing module, the classification considered for the records of the 13 largest cities in the quarters I, II and IV of ECH for the 2002-2005 period is the one resulting of applying the multinomial logistic model.

³⁰ The question is: "If you had to pay rent for this house, how much would you be willing to pay monthly?". This question can be found in the chapter of the form corresponding to household data.

³¹ Values, which are less than 10.000 and greater than 3.000.000, both for lease and for the mortgage amortization payments were excluded.

³² This imputation is made at the spending unit level and is exclusive of those that were classified or identified as homeowners or in usufruct condition.

³³ Any reference to ISU refers to the sum of sources of income of all earners of the spending unit.

There are three alternatives for the formation of PCISU:

- a) In the case of tenant households of ECH for the 2002-2005 period and the GEIH for the 2008-2010 period: ISU of the tenant household (or in another tenure condition other than fully paid, being paid or in usufruct) corresponds to the sum of income by source without any type of imputation for homeownership. PCISU will be the result of dividing ISU into the number of members of the spending unit.
- b) In the case of home-owning and home in usufruct households of ECH for the 2002-2005 period: ISU of home-owning and home in usufruct households of ECH for the 2002-2005 period corresponds to the sum of income by source plus the imputed value of homeownership through the Hot Deck model, according to the following criteria:
 - If the home is owned and fully paid, the estimated rental value is imputed.
 - If the home is owned, but is still being paid, the difference between the estimated rental value and the mortgage amortization payment is imputed, provided that this difference is not negative.
 - If the home is occupied in usufruct, the estimated rental value is imputed, provided that no family member had reported their home as income in kind for their work.
- c) In the case of home-owning and home in usufruct households of GEIH for the 2008-2010 period: the ISU home-owning and home in usufruct households of GEIH for the 2008-2010 period corresponds to the sum of income by source plus the expected rental value³⁵. Again, PCISU will be the result of dividing ISU (including imputation for homeownership) into the number of persons that compose the spending unit.

2.6 Income underreporting and the adjustment methodology to National Accounts

As mentioned in Chapter 5 of the *MESEP Primera Fase* (First Phase) document, the Adjustment of Income to National Accounts (ANA) had been used in the methodology for measuring poverty in Colombia in order to correct the problem of underreporting and domestic coverage that is recorded in household surveys. It is important to note that at the time, this methodology was taken from ECLAC (with some variations) in the same manner the methodology for constructing poverty lines was taken in the eighties.

³⁴ As stated at the beginning of the document, the spending unit consists of household members other than the domestic employees and their families, pensioners and other household domestic workers.

³⁵ For GEIH the distinction in the imputation between those who already paid their home, those still paying it (nonnegative difference with the installment) or those you have it in usufruct (only if the home was not reported as income in kind) also applies.

The ANA methodology consisted of comparing the amounts of income by source (salaries, profits and other income) constructed by using household surveys, against the amounts of the same sources of the System of National Accounts (SNA) of the country. As a result of this comparison, adjustment factors were generated that were then applied to the income of the survey, so that the issue of underreporting was corrected to a certain extent.

This procedure faces limitations. Some of the criticism and comments on the methodology are:

- The adjustment of the income to national accounts, being it an adjustment of levels, and by not considering the individual patterns of underreporting, introduces uncontrollable biases, which distort the distribution of income in a non-explicit manner. The adjustment coefficient applies equally to low and upper strata households (cf. Herrera, 2009).
- The values of the coefficients are rather high, which introduces substantial changes in the relative position of households with respect to their income and thus in the identification of the poor and non-poor (cf. Herrera, 2009).
- The ANA, according to Leyva-Parra (2004), rests on the assumption that the differences between national accounts and household surveys is largely due to problems of underreporting in the latter and not of truncation: The truncation in the upper part of the income distribution is a phenomenon that probably explains part of what has traditionally been attributed to underreporting. This implies that if the adjustment to National Accounts is made so that the two components of the discrepancy cannot be distinguished, an amount of money greater than what actually corresponds to the households of the sample will be statistically redistributed among them (cf. Leyva-Parra, Gerardo, 2004).
- "The 'income adjustment' was part of the ECLAC method from the beginning, as a way to correct the underreporting of income in the surveys. At the time, the 'Household Account' of the National Accounts seemed appropriate to assess the income of the surveys. That conclusion does not necessarily currently hold" (cf. Feres, 2010).
- "Some improvements to the correction for income underreporting are being assessed, but also the possibility of not continuing with this practice is being analyzed" (cf. Feres, 2010).

In addition to the above comments, another aspect indicated in MESEP as a disadvantage of the ANA methodology is the delay in the publication of the final household account, which is two years on average. This delay requires assumptions or extrapolations about the structure of the sub-accounts with respect to salaries, profits and other income, so that the amounts to be used as "ceilings" for the corresponding adjustment can be obtained. This made the figures of poverty and inequality that were produced in the most recent years to have a "provisional" character

until the publication of the final SNA. In addition, ANA reduces the inherent variability to the income cycle.

MESEP, considering the disadvantages mentioned above, decided not to continue using ANA as a method of correcting income underreporting. This entails that the figures with respect of income, poverty and inequality that are reported with the new methodology *only include income imputation for non-response (or biases therein) and of homeownership*.

2.7 Summary: Differences with the previous methodology

According to the above, the new methodology for measuring poverty in Colombia *adopts significant changes in the calculation of the income aggregate*³⁶ in three aspects:

- The imputation model seeks to debug primary information in three dimensions: in addition to imputing by missing values (as it was done in the previous methodology), corrections were implemented by extreme values and false zeros.
- Income imputation of homeownership differs in two aspects: i) the previous methodology only considered for this imputation, the owners with homes that were fully paid. The new methodology considers also owners who are paying their home and living in usufruct. ii) The previous methodology used information of National Accounts (gross operating surplus) to make the imputation of homeownership. The new methodology uses a Hot Deck model for the 2002-2005 period and the question on expected rental value for the 2008-2010 period.
- The application of the adjustment to National Accounts, which was implemented in the previous methodology, was discontinued.

Before describing the methodology used to define the values of the new poverty lines, below is a description of the information source selected for calculating them (in Section 3) and a brief overview of the most common methods to perform this procedure (in Section 4).

³⁶ It should be noted that the *income* variable incorporates the splices due to the methodological change in household surveys (from ECH to GEIH starting from the second half of 2006) and the change in the sampling frame (from the 2005 Census starting from 2009). The splicing factors that were used were the same as those applied during Phase I of MESEP (cf. MESEP, 2009). The necessary verifications were carried out in order to assess whether these splicing factors remained relevant to the construction of the income aggregate of Phase II of MESEP (dominance analysis and calculation of the statistical significance of the differences in the indicators of interest). It is important to clarify that in the case of *splicing due to sampling frame change*, unlike the results formalized in Phase I of MESEP (op. cit. MESEP, 2009), splicing factors of Phase II of MESEP were applied on 2002-2008 data. This was done so that the data for those years would incorporate the updated sampling frame. It must be added to this that the splicing factors due to sampling change were applied gradually backwards, which means that this effect diminishes to the extent in which the year for which the estimate is made moves away from the year in which the new framework was introduced in GEIH (2009). In Phase I of MESEP the procedure was different. Splicing factors were applied to the 2009 datum to discount from this the effect of this new sampling frame and to define in this manner a series that was comparable to the years 2002-2008. In this sense, the procedure was contrary to the one incorporated in the new series presented herein.

3. Assessment of the source of information for the construction of the new poverty lines: The National Survey of Income and Expenditure (ENIG) for the 2006-2007 period.

The sources of information used to build the different versions of the poverty lines in Colombia have been the Income and Expenditure Surveys (IESs) (cf. Muñoz and Rivas, 2006: 10-13). It should be noted that the previous methodology was developed from two sources of information. IES for the 1994 – 1995 period was used for the definition of the urban poverty line and the information of the first follow-up to the *Familias en Acción*³⁷ Program (2003) was used for the definition of rural poverty line.

Unlike such methodology, the construction of urban and rural poverty lines, which are presented in this paper, was carried out with the same source of information: ENIG, corresponding to the 2006 to 2007 period. This innovation ensures better urban and rural data comparability, and greater robustness when estimating the incidence of rural poverty³⁸.

ENIG for the 2006-2007 period collected information on the amount and distribution of Colombian household expenditure, as well as the amount of the different sources of their income (DANE, 2009: 15). ENIG for the 2006-2007 period was designed with a sample of 64.119 households³⁹, which is in turn a subsample of GEIH (cf. DANE, 2009: 19). Data collection took place between October 9, 2006 and October 7, 2007. The collected database had a sample of 35.998 households with completed surveys, i.e. 56% of the projected size in its design. Of these households, 32.534 (90%) belong to the urban area. The remaining 3.464 are from the rural area⁴⁰.

Information on household expenditure was collected in four forms⁴¹:

- Form 2 Urban –Daily household expenditure diary.
- Form 2 Rural – Weekly household expenditure diary.

³⁷ Families in Action. Conditional cash transfer program of the Colombian government. For further information refer to: <http://www.dps.gov.co/contenido/contenido.aspx?catID=204&conID=157>.

http://siteresources.worldbank.org/INTCCT/Resources/5757608-1234228266004/PRR-CCT_web_noembargo.pdf

³⁸ The previous rural poverty line faced a restriction with respect to the information that was considered for its design (first follow-up of the *Familias en Acción* (Families in Action) Program - 2003) which only represented a subset of the rural population (the poorest ones, whose condition made them the beneficiaries of the program).

³⁹ This number of surveys enabled having a sample design with representativeness at a national level, at an urban and rural level of the national level, at a six-region level (Atlántica (Atlantic), Bogotá, Central, Eje Bananero (The Banana growers' axis), Oriental (Eastern) and Pacífica (Pacific)) and at a twenty-four-cities level (Armenia, Barranquilla, Bogotá, Bucaramanga, Cali, Cartagena, Cúcuta, Florencia, Ibagué, Manizales, Medellín, Montería, Neiva, Pasto, Pereira, Popayán, Quibdó, Riohacha, San Andrés, Santa Marta, Sincelejo, Tunja, Valledupar and Villavicencio).

⁴⁰ By region, 33% of the sample are households of the Central region (12.045 households), 30% of the Atlantic region (10.686 households), 18% of the Eastern region (6,609 households), 15% of the Pacific region (5.340 households), and the rest are in Bogotá (830, which represents 2%) and in the Banana Axis region (488, representing 1%). Considering the sample by cities, those that have more households in the sample are: Bucaramanga (1.655) Medellín (1.595), Cartagena (1.578) and Ibagué (1.578). Cities with fewer households in the sample are: San Andrés (83), Quibdó (493), Santa Marta (812) and Bogotá (830) (cf. Cortes (2009): 3).

⁴¹ Information on labor market and income of GEIH were captured in Form 1.

- Form 3 – Less Frequent expenditures diary.
- Form 4 – Personal expenditures diary (for income earners).

As explained below, the value of the basic food basket (which corresponds to the value of the extreme poverty line) as well as the value of the poverty line (which results from multiplying the value of the extreme poverty line by the Orshansky coefficient (ratio between total expenditure and food expenditure)) are estimated from the information of ENIG.

Overall, for the construction of the poverty and extreme poverty lines, the information collected in ENIG should meet two conditions:

- It should satisfy the empirical regularity of Engel's Law.
- The estimated amounts of the various items of expenditure are within reasonable levels when compared with other sources of information.

3.1 Internal consistency

This section summarizes the results found by Cortés (2009) pertaining to the consistency of ENIG for the 2006-2007 period, which was the main input used for the second phase of MESEP to determine the usefulness of the survey in the design of these lines.

In order to analyze the information of ENIG for the 2006-2007 period, Cortés (2009) assesses two aspects: i) the internal consistency and ii) the external consistency. To assess the internal consistency, Engel curves, linear expenditure systems and descriptive exercises were estimated. To analyze the external consistency, distributions of various items of expenditure were compared to those obtained in other surveys (IES for the 1994-1995 period, Quality of Life Survey [ECV for its acronym in Spanish] for the year 2003, Bogota ECV for the year 2007 and ECV for the year 2008).

The results obtained by Cortés indicate that the information of ENIG for the 2006-2007 period is internally consistent at a national level as well as for the urban and rural domains. However, in some geographic domains the results are not satisfactory due to the effective sample that was collected: "Among the regions, the most serious cases are the banana axis region and Bogotá. Among the cities, the most serious cases are: San Andrés, Riohacha, Quibdó and Bogotá" (Cortés, 2009: 331). Since the loss of sample was concentrated in particular cities and regions, it is advisable to use the information with a higher degree of aggregation. This aspect was taken into account by MESEP to define the maximum level of disaggregation of information to build the new poverty lines.

The distributions (relative and absolute) of expenditure items⁴² keep statistical sense and are consistent with the economic theory of consumer⁴³: "The averages in all items are sorted correctly according to the income quintile and are statistically different from each other" (Cortés, 2009: VI). This means that for each of the expenditure types, households of lowest income quintiles spend on average less than high-income quintiles.

As a complement, internal consistency was assessed by estimating *linear expenditure systems* with a sample that was restricted to urban households (7.387 households). According to the postulates of economic theory, the estimated coefficients (that represent the marginal increase in the expenditure in each good when the total expenditure is increased) were significant and positive. From these analyses, expenditure on food, housing, health and clothes were identified as necessary. In these cases, the coefficients, which are interpreted as subsistence quantities, were greater than zero. In descending order, Cortés indicates "households require more food, housing services and clothes to survive" (Cortés, 2009: 177). The coefficients were below zero for expenditure on culture and education. This result would indicate that it is not necessary to spend in such item to ensure subsistence.

Expenditure and price elasticities of the Marshallian and Hicksian demand (calculated in the sample mean) were also estimated. The lowest *expenditure elasticities* were food, housing services and health, indicating that these are necessity goods. Expenditure on personal services resulted with an elasticity greater than one, which indicates that this type of expenditure is considered luxurious. In turn, the *price elasticities* were negative. From these results it is concluded that the demand for goods that have more changes with respect to the price are those of culture and education, personal services and transport. The more inelastic goods are food, housing services and clothes. Cortés concludes that "Overall, the obtained results are consistent with economic theory" (Ibid.).

Other than the internal consistency of the results, it is important to verify that the amounts for the various items of expenditure are consistent with those obtained from other sources that also collect information on household budgets (e.g. ECV). Below is the summary of the results found by Cortés regarding this matter.

⁴² The distribution of expenditure was analyzed in the following 10 items: Food, Drinks and Tobacco, Clothes, Housing Services, Furniture, Health, Transportation, Recreation, Education, and Other Services.

⁴³ It refers to Engel's Law (established by Ernst Engel in 1857). According to this law, as total expenditure increases, food expenditure also increases but at a lower rate. The relative importance of food in total expenditure (captured by Engel's Law) indicates the living conditions of households. For this reason, it is important that the results that are captured in the IES meet this empirical regularity. First, because they account for an adequate differentiation of living standards between the households surveyed. Second, because they are the basis for the estimates of poverty and extreme poverty lines. Specifically, the relationship between total expenditure and food expenditure is the basis for the calculation of the Orshansky coefficient (inverse of the Engel coefficient), one of the two components involved in obtaining the poverty line (the other component is the value of the basic food basket or extreme poverty line). Regarding the calculation of Engel curves for different items of expenditure, consult Cortés, 2009, p.28.

3.2 External consistency

Cortés' study (2009) highlights the underestimation of some items of expenditure of ENIG for the 2006-2007 period, compared with other sources that collect information of this nature. According to Cortés, the results of external consistency are not at all satisfactory and, in some cases, the underestimation is around 50% (cf. Cortés, 2009, p. 331). In particular, the author warns about the undervaluation of the item concerning food. When comparing the average expenditure on them of ENIG for the 2006-2007 period and ECV for the 2008 period, the difference is significant: \$ 287,000 pesos for the first one and \$ 654,000 pesos for the second one, in real prices of 2007 in both cases⁴⁴. The only item that does not present undervaluation is the one corresponding to clothes⁴⁵.

The underestimation of some items, particularly food, influences the estimation of the Orshansky coefficient. This parameter can be overvalued when the food component represents an amount lower than expected, compared to other sources of information.

Since the value of the poverty line is the result of combining two elements namely, the value of the basic food basket [extreme poverty line] and the Orshansky coefficient, the overvaluation of the latter leads to a value of the poverty line that is above the expected.

Considering the above, MESEP opted for a set of decisions to construct poverty and extreme poverty lines. These decisions are mentioned in the following section.

3.3 Implications of the consistency assessment of ENIG for the 2006-2007 period

As a result of the consistency analysis (internal and external) of ENIG for the 2006-2007 period, the group of experts of MESEP highlighted the difficulties associated with the definition of an endogenous Orshansky coefficient considering underreported information on food expenditures. Constraints on the source of information led MESEP to propose the following guidelines for the construction of new baskets:

- *Source of information:* Building new poverty and extreme poverty baskets considering information from ENIG for the 2006- 2007 period.

⁴⁴ It is worth mentioning the methodological differences that characterize each investigation. In ENIG, which is applied in urban areas, a *daily record* of weekly food expenditure is made (i.e. in the last 7 days). In rural areas, the registration of food expenditure is *the total expenditure for the week*.

Moreover, in ECV, *total food expenditure made in the week* is recorded in both urban and rural areas. In addition, in ENIG, expenditure is recorded at the article level whereas in ECV it is done by groups of food articles. ENIG also records food expenditure with a frequency higher than weekly.

⁴⁵ At a regional level (Bogotá, Eastern, Central, Atlantic and Pacific), the results confirm what was observed at a national level. In Bogotá, the underestimation of the items of ENIG for the 2006-2007 period varies between 20% and 30%.

- *Domains*: Using urban and rural domains as the maximum level of disaggregation for the information of ENIG for the 2006-2007 period.
- *Orshansky coefficient*: Using an exogenous Orshansky coefficient⁴⁶ for urban areas and, from this point, estimating the coefficient of the rural area by keeping the proportionality that ENIG for the 2006-2007 period reports between the two areas.

⁴⁶ This aspect is analyzed in greater detail in Section 5.8.

4. Alternative methodologies for the construction of poverty lines

Sen (1979) differentiates between two general methods to identify the poor⁴⁷. The first of them is the *direct method*. The set of households or individuals that do not meet a specific group of previously established needs (e.g., housing conditions, education, household demographics, asset ownership, etc.) is identified through this method. The second one is the *income method*. In this case, the identification is done by calculating a minimum threshold, usually associated with a level of income or expenses, below which it is considered that a person cannot meet their basic needs. While the first method is based on the *actual* satisfaction of a set of needs, the second refers to the *possibility* of satisfying them given a minimum level of income⁴⁸.

Examples of the direct method include the Index of Unsatisfied Basic Needs and the new Multidimensional Poverty Index⁴⁹. The income method is the framework for traditional poverty lines.

The *income method* has two alternative methodologies for defining the minimum level of resources required for meeting basic needs: the food energy intake method and the cost of basic needs method. A brief description of these two methods and the methodology used by ECLAC for defining budget-based poverty lines is presented below. A section introducing some alternatives to correct possible inconsistencies in the construction of poverty lines is also included.

4.1 Food Energy Intake Method

The food energy intake method aims at specifying the level of total expenditure/income that would allow meeting minimum nutritional requirements. The basic assumption (see Graph 2) is that *as income (or expenditure) increases, food energy intake also increases, but at a decreasing rate* (cf. Haughton and Khandker, 2009, p. 55). In order to obtain the poverty line, the point where nutritional requirements are met is established in the *calorie income* function. This point will be associated with the amount of income (expenditure) equal to the value of the poverty line (Z). It is worth mentioning that the method implicitly considers food and non-food

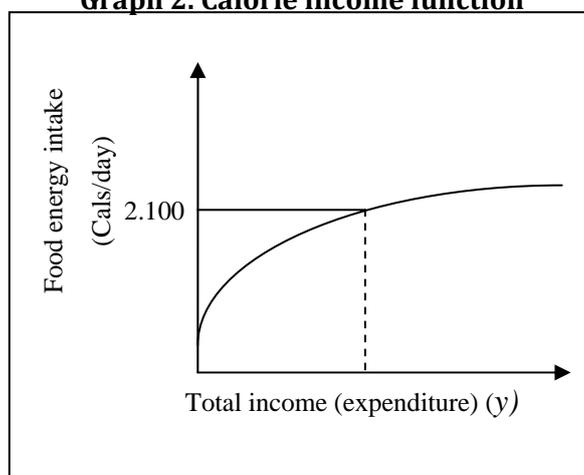
⁴⁷ The same author establishes two steps to assess poverty: identification and aggregation. The first step consists of selecting a criterion to differentiate the poor within the total population. It answers the question: Who are the poor? The second step consists in obtaining a summarized measurement of the extent of poverty. It answers the question: How many persons are poor? cf. Sen 1979: 1. See also Foster and Sen, 2001: 192.

⁴⁸ Sen (1981) illustrates these concepts through a frequently cited example: "*The ascetic who fasts on his expensive bed of nails will be registered as poor under the direct method, but the income method will offer a different judgment in recognition of his level of income, at which typical people in that community would have no difficulty in satisfying the basic nutritional requirements.*" Sen, 1981: 27.

⁴⁹ Angulo, Roberto; Díaz, Yadira; Pardo, Renata. 2011. *Multidimensional Poverty Index for Colombia. (MPI-Colombia) 1997-2010*. Departamento Nacional de Planeación (National Planning Department).

consumption, given that the x-axis represents the total expenditure. This aspect is of great importance when analyzing the consistency of the lines obtained with this method.

Graph 2. Calorie income function



Source: Haughton y Khandker (2009). p. 55.

Despite the advantages offered by this method given that it requires less information (cf. Osmani, (1982); Greer and Thorbecke, (1986); Satya, (1989); Ravallion (1998), 11 (there is no need for spatially differentiated prices nor to distinguish the non-food consumption of the poverty line), estimated lines can be inconsistent because the level of expenditure/income is not the only factor determining energy consumption (cf. Ravallion, 1998)⁵⁰. Other aspects such as preferences, activity level, relative prices or publicly-provided goods also affect calorie consumption⁵¹, as shown below:

- *Preferences*: Urban households tend to consume foods that have higher prices per calorie. For example, they tend to consume protein-rich foods such as meat instead of carbohydrate-rich foods such as flour. When comparing two persons with similar actual expenditures, it cannot be concluded that the one that consumes food with the higher prices per calorie is the poorest.
- *Relative prices*: Unlike rural areas, in urban areas the prices of some non-food goods tend to be relatively lower than those of food. It may be that, for a given level of income, the demand for food in urban areas is lower than in the rural area. However, this does not mean that urban households are poorer than rural ones.
- *Activity level*: Usually, the jobs carried out in urban areas demand fewer calories than those carried out in the rural area. For this reason and for a given level of income, food

⁵⁰ On consistency, Ravallion argues that poverty lines should be absolute in the space of welfare, but relative in the space of goods and prices: "I will argue that for the purposes of informing anti-poverty policies, a poverty line should always be absolute in the space of welfare. Such a poverty line guarantees that the poverty comparisons made are consistent in the sense that two individuals with the same level of welfare are treated the same way". (Ravallion (1998), p. 5)

⁵¹ "The relationship between food energy intake and income will shift according to differences in tastes, activity levels, relative prices, publicly-provided goods or other determinants of affluence besides consumption expenditure". Ibid. p. 11.

consumption tends to be lower in urban areas without implying that such households are poorer.

When in the definition of poverty line aspects such as those mentioned above are not taken into account, inconsistent results can occur that prevent an adequate comparison of welfare between individuals. For example, the incidence of poverty in urban areas may be higher than in the rural area⁵².

Another characteristic of this method is that while it determines the level of expenditure/income in which the average food energy intake requirement is met, it does not enable the identification of the food consumption pattern of the population that meets such requirements. This is a limitation given that there are multiple ways of meeting the food energy intake requirements. With the food energy intake method it is not possible to identify whether or not the requirements are “adequately” being met.

4.2 Cost of Basic Needs Method

The cost of basic needs (CBN) method consists of calculating the cost of a basic consumption bundle. This approach has its origin in the seminal studies of Rowntree (1901, 1936 and 1951). In his 1901 work, Rowntree distinguishes between two types of poverty: primary and secondary. Primary poverty is defined as the situation in which families live, where income is insufficient to cover the minimum requirements to maintain physical efficiency. Secondary poverty is described as the situation of those families whose income may be sufficient to cover the minimum requirements, but a portion of it is spent on other type of consumption. Both concepts are closely related to the budget approach from which extreme poverty and poverty lines are defined. In general, the steps below⁵³ are followed to calculate the basic consumption bundle (Haughton and Khandker, 2009: 49):

- To define the nutritional requirement of the food component. Usually, the value estimated by FAO is taken as reference (about 2,100 calories per day per person), considering differences in age, sex or activity⁵⁴.
- To estimate the cost of meeting the nutritional requirement, guaranteeing a diet that respects the dietary habits of the reference stratum⁵⁵ (e.g. the population in the lowest or

⁵² Bidani and Ravallion (1994) conducted a study in Indonesia with information from the 1990 Household Survey. They found that the incidence of poverty in urban areas was higher than in rural areas when using the food energy intake method to construct poverty lines.

⁵³ Some of these issues will be discussed in more detail in subsequent sections.

⁵⁴ The latest information with respect to energy requirements is from 2001: FAO, WHO and UN (2001). Colombia had been using FAO's information from 1985. The new methodology for the construction of the extreme poverty line incorporates information from 2001. ECLAC also uses information estimated by FAO as a reference criterion to define the nutritional requirements. Another source of information is the one of the World Health Organization, 1985 (cf. Ravallion 2010, p.8).

second-lowest quintile, or those consuming between 2,000 and 2,200 Calories). This will be the food component equal to the value of the extreme poverty line.

- To obtain the value of the poverty line. The Orshansky coefficient is used to establish this value. The coefficient reflects the reciprocal of the food expenditure share with respect to the total expenditure. Thus, the value of the poverty line will be the sum of the food and non-food component⁵⁶.

The CBN method reports the minimum income that offers the possibility of acquiring the basic bundle, but without guaranteeing its actual consumption. It indicates if persons have enough resources to buy the bundle regardless of whether they consume it or not. Therefore, in this analysis it is essential to have information on the prices of the consumer goods in the bundle. If they are not available, analysts rely on the food energy intake method.

The CBN method is not exempt from the limitations mentioned in the description of the food energy intake method. Differences attributed to relative prices, activity levels and preferences can also generate inconsistencies in the results. However, unlike the food energy intake method, the CBN method does include information regarding the contents of the food basket. Because of this, it is possible to establish the consumption pattern of the population that meets the requirements.

4.3 Alternatives to fix inconsistencies

Ravallion proposes some alternatives to overcome the limitations described above⁵⁷. Particularly, the author proposes the construction of a Spatial Price Deflator (hereinafter SPD)⁵⁸ to fix inconsistencies associated with relative prices. With this procedure, an expenditure adjusted by cost-of-living may be formed, expressing prices from different domains of analysis as a function of a reference domain (e.g., rural prices in terms of the urban prices). Once the expenditure has been adjusted by cost-of-living, the reference stratum can be selected to observe its consumption pattern. To a certain extent, a reference population defined in this way will not be affected by the differences in prices.

⁵⁵ The reference stratum is the segment of the population in which two aspects are met: i) the fulfillment of minimum nutritional requirements and ii) the representativeness of a household consumption pattern. Extreme poverty and poverty lines are built with the information of this subset of households. For Muñoz and Rivas, "The idea is to take a population segment that represents the consumption of food of households that fulfill the nutritional requirements and therefore, can be taken as the most representative segment of the poor population" (Muñoz and Rivas 2006, p.10).

⁵⁶ See Orshansky, Mollie (1965 and 1966). These works initiated the use of the Orshansky coefficient, which is still current. As noted by Haughton and Khandker: "In 1963 and 1964, Mollie Orshansky of the U.S. Social Security Administration computed the cost of an 'adequate' amount of food intake. Then, to obtain the value of the basic basket (with the food and non-food components), Orshansky multiplied such cost by 3. At the time, the average food share for all consumers in the United States was 1/3." (cf. Haughton and Khandker 2009 : 50).

⁵⁷ Deaton, Angus and Tarozzi, Alessandro (2000); Ravallion, Martin (1998), (2006) and (2010).

⁵⁸ For more information refer to Section 5.2

As it was already mentioned, the food energy intake method cannot collect information on prices directly from the source to build the line. In such cases, a secondary source can be used to deflate the total expenditure. For the CNB method, the information on prices is directly available. In both cases, the differences in the consumption capacity caused by differences in prices can be reduced as much as possible through SPD.

Moreover, as some studies have acknowledged, differences attributed to age, sex, climate or level of activity significantly influence the household consumption pattern: "The family life cycle affects significantly the share of expenditure destined to, for example, transportation, education, health and housing" (Feres and Mancero 2001:57). Proposals in this sense include: i) considering nutritional requirements according to the type of work done by individuals⁵⁹ and other demographic characteristics (age and sex) (cf. Ravallion 1994 and 1998) or ii) using equivalence scales.

An equivalence scale is an index that enables expressing the relative acquisition cost of a set of goods and services in terms of a typical household (depending on its size and composition). For Feres and Mancero:

"The concept of equivalence scales simultaneously groups two elements. On one hand, the scale considers the different needs of household members by age, sex, or other demographic characteristics or type of activity. On the other, this index enables taking into account the existence of economies of scale, characterized by decreasing marginal costs so as to achieve the same level of welfare given the addition of a new member to the household. This last point is related to the existence of public goods that can be shared within the household without incurring in a reduction in the welfare of the people" (Feres and Mancero, 2001: 53).

With respect to the preferences that characterize consumption in different regions (urban/rural), geographically differentiated consumption patterns can be established, as indicated by Ravallion: "A simple method of allowing for substitution is to set a bundle of goods in each region, (say) which is the average consumption of a reference group fixed nationally in terms of their income or expenditure" (Ravallion, 1998:16). This proposal establishes a single definition of the reference population (national reference stratum), from which a set of goods that characterizes the preferences of each region can be selected (based on the consumption pattern of the persons selected as reference stratum).

The new methodology for measuring monetary poverty in Colombia uses: a SPD, a single reference population at the national level (although with a pattern differentiated by regions given the information of the reference population) and a calculation of the average food energy

⁵⁹ IESs that have been conducted in Colombia do not include information from which the level of activity of respondents can be established.

intake requirements by domain, taking into account differences by age and sex. It was not possible to take into account differences attributed to activity level since ENIG 2006-2007 does not include information on this matter ⁶⁰.

4.4 Methodology applied by ECLAC

According to Herrera (2010), the methodology applied by ECLAC is a combination of the two methods described above. The reference stratum in which the food energy intake requirements are met is determined in order to obtain the extreme poverty line⁶¹, and subsequently, using the Orshansky coefficient, the value of the poverty line is obtained, as estimated utilizing the method of the basic consumption basket. ECLAC's Methodology (1991) for estimating poverty lines in 10 countries in Latin America is described below.

The first step is to organize the population according to its per capita income. Based on this organization, a reference population stratum is selected for which a consumption pattern is obtained, i.e., the list of goods with the specification of the value spent on each of them. The reference stratum is selected based on two criteria: i) "that the consumption habits of the group are the expression of household decisions in a context presumably free from a significant restriction of resources, and ii) that the size of the group is such that its pattern of expenditure may be considered representative enough" (ECLAC, 1991: 20). The selected reference stratum is the one that meets the representativeness criterion, and that, in average, slightly exceeds the level of minimum nutritional requirements. In general, for the Latin American countries analyzed by ECLAC (1991), the reference stratum was ranked among the 20th and 50th percentiles of the income distribution.⁶²

The consumption pattern of the reference stratum is inferred from the detailed record of food expenditure. Food items with a significant share within each food subgroup are selected, and those food items with less participation are located in the "others" category. To determine the quantities associated with each of the food expenditures, ECLAC used consumer prices reported by the statistical offices of each country (the same involved in the construction of the Consumer Price Index [CPI]). The amounts deduced in this manner are expressed in grams/day per person (dividing the respective expenditure by its price⁶³). The calories and nutrients of each product

⁶⁰ Including information on the activity level of respondents is an important recommendation for future applications of ENIG.

⁶¹ Unlike the Food Energy Intake method in which the value of the poverty line is directly obtained, including the food and the non-food component.

⁶² For the particular case of Colombia, this study ranked the reference stratum between the 26-50 percentiles in Bogotá, and the 36-60 percentiles for other urban areas.

⁶³ "In the case of the 'others' category, which includes a set of products, its price and therefore its quantity, was estimated in the base of the weighted average of the values of the individual goods comprised in the respective subgroup". Ibid. p. 21.

are obtained from the nutritional data tables.⁶⁴ Physical quantities of each food are adjusted in such way that the caloric content of each basket matches exactly the average calorie requirements by country and region. Three adjustments make these baskets to be considered as adequate in terms of calorie intake: i) adjustment to the minimum requirements in terms of energy and protein, ii) suitability according to food availability and iii) exclusion or substitution of items according to cost considerations.⁶⁵

Additionally, the basket is subject to an assessment in terms of the nutritional quality to meet certain dietary standards (source of calories and protein quality). A diet is considered as acceptable if it comprises at least 10% of calories from protein and 15% to 25% calories from fat. In turn, cereals and legumes must not provide more than 60% of total calories. With respect to the protein quality it is considered as appropriate if at least 35% of them are of animal origin. Quality analysis also takes into account the content of calcium, iron, vitamin A, thiamine, riboflavin, niacin and vitamin C.

The cost of the food basket is calculated by using the prices collected to construct the CPI. In order to establish the cost of meeting the basic non-food needs, a normative relationship is adopted between the expenditure on food and other consumption expenditure. Two conceptual implications arise from this procedure: i) it is assumed that households that meet food requirements in an adequate manner, also meet the requirements associated with other needs and ii) the access to public goods and services is not included in the consumption pattern; only private consumption is considered.

Finally, in order to obtain the value of the poverty line for urban areas, ECLAC (1991) considered appropriate to adopt an Orshansky coefficient equal to two (2). This implies that the value of the private consumption doubles the budget destined for food. Annex D describes the changes that ECLAC is evaluating in order to update the methodology for the construction of extreme poverty and poverty lines.

⁶⁴ "Given that there was no information available on pricing and nutritional composition of the expenditure on drinks and food 'outside of the home', it was necessary to define a special procedure to determine its equivalence in physical quantities and, particularly, the respective contribution in terms of calories and protein: it was assumed, based on the limited information available, that the cost of the nutrients provided by those items was equal to three times the average cost of the nutrients contained in food consumed 'within the home' and that the composition in terms of goods of that expenditure component resembled the latter". Ibid. p. 21.

⁶⁵ Once the consumption pattern of the reference population was identified, some items that have a high price per calorie, low impact on food expenditure or those that are dispensable from a nutritional standpoint were excluded. The result is a food basket that respects consumer habits and minimizes the "raise-the-price" effect of the resulting diet.

5. Methodology applied for the construction of new poverty lines

The update of the methodology for the construction of poverty lines for Colombia followed the method proposed by Ravallion (1998), with some variations, which are the result of the analysis and empirical tests conducted by MESEP based on special recommendations of ECLAC that were presented during the MESEP sessions. In particular, the construction of a *short* food basket⁶⁶ and some considerations for the estimation of the Orshansky coefficient were adopted from ECLAC.

The procedure that was followed to build the new poverty lines (PL) and extreme poverty or indigence⁶⁷ lines (IL) included the following steps (cf. Herrera 2001 and 2010):

1. Construction of per capita current expenditure at the spending unit level.
2. Construction of the Spatial Price Deflator (SPD).
3. Sorting of households by deflated per capita current expenditure percentile.
4. Application of the iterative method for the selection of the reference population (starting the iterative process with an *a priori* reference population that is between 30-59 percentiles).
5. Construction of the basic food basket.
6. Regulatory adjustment of quantities of the basic food basket in order to meet the caloric requirements.
7. Assessment of the adjusted food basket; its value corresponds to the indigence or extreme poverty line (IL).
8. Transition from IL to PL by means of the Orshansky coefficient.
9. Estimation of the incidence of poverty⁶⁸.

In the following each of these steps are explained in more detail.

5.1 Construction of per capita current expenditure at the spending unit level

To build the new poverty and extreme poverty lines, *current expenditure* of households was used as a welfare variable. Three determinants of expenditure defined by the International Labour Organization (ILO) were considered:

⁶⁶ It is called short food basket because a subset of items that meets three criteria is chosen: 1) It is consumed by 30% or more of households, 2) It represents at least 1% of total food expenditures and 3) It provides at least 1% of the total of calories consumed. The objective is to obtain a basket that includes food that is frequently acquired by persons and that considers their consumption habits.

⁶⁷ The terms *extreme poverty* and *indigence* are used interchangeably in this document.

⁶⁸ This last step verifies that the incidence of poverty is consistent with the reference population that was selected in step 4. That is, that the value of the incidence is located in the interval used as the reference population (following Ravallion 1998).

- *Acquisition*: It comprises the value of consumer goods and services received by households during a given period, whether or not they have been paid, or consumed in the same period or further.
- *Payment*: It equals the sum of the payments made in cash or in kind during the reference period for purchases of consumer goods and services, regardless of when they were received or will be received, and the time in which they will be used or consumed.
- *Consumption*: It refers to the value of goods and services used during a certain period, regardless of when they were acquired or paid. These goods and services become "extinguished" when not consumed. With respect to consumer durables, their "use" should be measured.

The starting point to form *current expenditure* was the definition of *total expenditure* in ENIG for the 2006-2007 period (DANE, 2009: 25). From the latter, some modifications were made to reach the final version that was used in the definition of the lines (Annex E shows the modifications made in detail). The *items included* to form current expenditure were:

- *Out-of-Pocket expenses*: Food, transportation, communication, recreation and culture, included in Forms 2 and 4 of ENIG;
- *Less frequent expenses*: The items included in Form 3, except for the following items: mortgage amortization payments for dwellings that are for household use but different from the dwelling inhabited; purchase of dwellings, dwelling expansion or subdivision; purchase of land, lots, premises, farms, plots and the like; payment pertaining to taxes of dwellings that are occupied by households, the purchase of new and used vehicles for private use; registration fees for new vehicle and used vehicle handover; certificate of gases and other reviews mandated by law; payment of SOAT (mandatory auto liability insurance); vehicle taxes; monthly installments for acquisition of real estate other than housing; loan installments for vehicle acquisition; income tax and other taxes;
- Monthly rent payment;
- Payment for food received by children of the spending unit that are under three years of age (community center, day care or preschool) and persons aged three years or older who are studying;
- Repair and maintenance of furniture, flooring accessories and materials;
- Flooring accessories and materials;
- Purchase of books;
- Imputed rent;
- Materials and workforce for housing maintenance;
- Reconnections of various public utilities (water, gas, telephone);

- Credit card payment.

The *items discounted* for obtaining current expenditure were:

- Expenditure on purchase of furniture;
- Purchase of vehicles;
- Purchase of bicycles;
- Durable consumer goods;
- Housing expansion (investment);
- Hand and electric fans;
- Equipment and accessories for camping;
- Encyclopedias, works and collections;
- Radios for vehicles.

As a result of the deliberations of the experts group, the deductions made to the occupied (salaried and independent): health, pension, ICA (Industry and Commerce Tax and Supplementary Notice and Boards), VAT, withholding of tax at source, official diary⁶⁹ and pension solidarity discount, were not taken into account within the definition of current expenditure to construct poverty lines. One of the reasons for not considering these deductions is connected to the difficulty of estimating their value (especially in the case of independent workers). The level of recall of these values by the respondents is low. Likewise, in the case of payment of ICA, VAT, withholding of tax at source, and official diary, the survey does not explain the periodicity and the number of contracts to which the paid value refers. This information would allow showing the deduction on a monthly basis.

5.2 Construction of the Spatial Price Deflator (SPD)

Once the items that compose the current expenditure are defined, the next step is ordering households according to this aggregate. This requires removing the effect that regional prices may have on the household ordering, this is done in order to not affect the comparisons of the level of welfare among households.

Building SPD is an alternative that enables the adjustment of expenditure in order to eliminate differences attributable to relative prices. With this deflator, the magnitude of the difference in

⁶⁹ Payment made to the National Press of Colombia by natural or legal persons who have signed a contract with the State so to publish the contract in the Single Official Journal of Government Procurement. For further information refer to: http://www.gobiernoenlinea.gov.co/web/guest/home;jsessionid=11E1F3229A2C435061E68906E11370E1?p_p_id=GovBuilderPdfServicesViewer_WAR_govservicesviewerportlet&p_p_lifecycle=0&p_p_state=exclusive&p_p_mode=view&GovBuilderPdfServicesViewer_WAR_govservicesviewerportlet_serviceid=1386

the cost of living between regions can be determined. Following the recommendation of Deaton and Tarozzi (2000), information of IES is used for the construction of SPD.

The value of SPD depends on a basket of goods that represents the common expenses of all domains. Ideally, the SPD basket should contain goods that represent food and non-food. However, it is not easy to access information on quantities and expenditure on goods other than food, such as housing and clothes (ENIG for the 2006-2007 period does not capture quantities in expenditure items other than food and they cannot be calculated in an implicit manner either).

For this reason, MESEP decided to build SPD from a basket comprised exclusively of food⁷⁰. The criteria used to choose food included in the SPD basket are:

- It should be acquired by at least 30 households per domain.
- It should be acquired in all domains.

These criteria follow the empiric recommendations of other similar studies (cf. Romero, 2005 and Brandt and Holz, 2005). The first criterion is applied in order to have a minimum of observations per food item. The second criterion is used to ensure that included food is part of the daily consumption of all domains. After using these two criteria, the resulting basket contains 220 food items⁷¹.

The valuation of this basket can be obtained from two sources: either from the Consumer Price Index (CPI) (exogenous source), or from unit values extracted from ENIG for the 2006-2007 period (endogenous source). In this case, unit values were used to estimate SPD.

Unit values. Once the base basket for the country is defined for the domains of study selected, that is, for municipal townships (urban) and remaining areas (rural), the unit values are obtained. A unit value is a proxy of the real price that each household pays for an item. It is obtained by dividing the amount paid by the quantity purchased.

Exclusion of outliers. Following Deaton and Tarozzi (2000), outliers are excluded from the sample given that the unit values received by each family are different, even though, in theory, they should be the same as long as there is no quality distortion. For this, those values whose logarithm is 2.5 standard deviations above or below the mean of the distribution of unit values of each item are discarded at a domain level.

Weights. Once the unit values of each item included in the SPD basket are obtained, the average percentage of each good in total expenditure of each domain has to be calculated. These

⁷⁰ Other countries using SPD are Perú and Paraguay. They also build the deflator considering only food information.

⁷¹ Annex F contains the list of food items considered in the SPD basket.

proportions are called weights. There are two alternatives for calculating them: the democratic weights method and the plutocratic weights method. In a democratic weight, all households weigh the same within a domain. By contrast, in the plutocratic weight, households that incur higher expenses weigh more. In order to estimate the index with democratic weights, an arithmetic mean of the proportions of total expenditure that each household allocates to each good is calculated, taking into account all households within the domain. In turn, in order to calculate the plutocratic weight, the level of expenditure of all households for a particular good is added and divided by the total expenditure on all goods. In the same manner as for the democratic weight, the calculation is performed for each domain of analysis. To calculate the average percentage of each good in total expenditure per domain, MESEP adopted the democratic weights method.

Spatial Price Indices: alternatives. For the selection of the reference population on which consumer habits will be examined, Ravallion argues that households belonging to that group must have similar income or expenditure levels. That is, they should have an equivalent level of welfare. The level of welfare is affected, among other factors, by differentials in: i) the levels of relative prices, ii) income and iii) preferences between regions. In order to achieve an equivalent level of welfare between different regions or domains, Ravallion proposes the construction of a single reference population that is nationally defined and takes into account the existing price differentials between regions. The equivalence of prices is achieved with the application of SPD, which is explained below.

Spatial price indexes summarize differences in the prices of K products or goods in the value added of domain 0 to domain 1 . Once a representative basket is chosen for the two domains under consideration, the price level of domain 1 is calculated in relation to domain 0 . This calculation corresponds to the ratio of the cost of the basket in domain 1 with respect to the cost of the same basket in domain 0 .

There are several price indices that can be used to create a SPD. In particular, the Laspeyres, Paasche, Fisher and Törnqvist indices were assessed by using democratic and plutocratic weights. According to the obtained results, a significant difference between the resulting values of these indices was not evidenced. For this reason, the Törnqvist index with democratic weights was chosen as SPD. The Törnqvist index meets desirable theoretical properties featuring obtained estimates that are unbiased, symmetric and robust to outliers.

The results of SPD for Colombia by domain are shown in Table 3. By taking the urban area as reference domain, these results indicate that prices in rural areas are 3.3% lower than in urban areas⁷².

Table 18. Spatial Price Deflator (SPD)	
<i>Domain</i>	<i>SPD – Törnqvist index with Democratic Weights</i>
Urban	1
Rural	0.966

Source: MESEP based on DANE's ENIG for the 2006-2007 period.

This deflator is applied to per capita current expenditure of each household depending on its domain, in order to match all domains in terms of prices. This ensures consistency of the extreme poverty baskets.

5.3. Household sorting by percentile of deflated current expenditure

Once SPD has been applied, population is sorted by level of total per capita expenditure, expressed in prices of a domain considered as reference. From this sorting, the reference population (on which the consumption pattern is recognized) is identified.

5.4 Application of the iterative method for the selection of the reference population

In order to determine the composition of the basic food basket, whose cost corresponds to the value of the extreme poverty line, a segment of the population sorted by its current spending must be selected, after applying the price deflator as explained in the previous section. This subset of households of the sample, which serves as basis for the construction of the lines, is called *reference population*. Both lines (poverty and extreme poverty) are obtained from the consumption pattern of the population. For the construction of the extreme poverty line (IL) the consumption structure of the group of food items is used, and for the construction of the

⁷² In Colombia and other Latin American countries these type of estimations have been made . In the case of Colombia, Romero (2005) finds differences of up to 26% at the level of the cost of living between the domains of Colombia's IES for the 1994-1995 period. However, this cost of living index has a fairly significant non-food component. Even in one of the appendices of his work, Romero presents the results of the index considering food alone and reaches regional differences of very similar magnitudes to the exercise presented herein. The case of Peru is also considered as example. According to the 2001 National Household Survey (ENAHO) conducted by the National Institute of Statistics and Information Technology- INEI, the estimations for this Andean country consider as reference domain the city of Lima and find differences in urban prices between departments ranging from less than 17% compared to those of the reference domain (in the case of the city of Arequipa) up to 13% above those of the reference domain (in the case of the department of Madre de Dios). In general, rural prices between departments are lower than those of the reference domain (Lima), except the departments of Madre de Dios (15% higher than those of Lima) and Tumbes (1.2% higher), being those of Arequipa (20% below those of Lima) the lowest. For more information on the Peruvian case, refer to the document: *¿Cómo se mide la pobreza monetaria en Perú?* (How is monetary poverty measured in Perú?) National Household Survey. Presentation. Lima, October 2009.

poverty line (PL) the Orshansky coefficient is estimated, that is the inverse of the proportion of food expenditure with respect to total expenditure.

The purpose when selecting the reference population is to prevent the construction of the lines with information of the population that is located at the ends of the distribution of expenditure (very poor or very rich population). In the first case, consumption may be of low quality (e.g., poor diet, due to conditions of deprivation); and in the second case, consumption is associated with high prices (inefficient diet).

Thus, poverty and extreme poverty lines will be consistent in the sense of considering two households that are from different areas with the same per capita expenditure as poor, after applying SPD.

In order to obtain the reference population, a percentage of initial poverty for the national total must be assumed, around which such population would be centered⁷³. For Colombia, a 45% was considered based on measurements made in Phase I of MESEP. Thus, the initial reference population is the set of households belonging to the interval ranging from the 30th to the 59th percentile of the distribution of per capita expenditure.

Once the reference population is chosen *a priori*, its validity is verified. For this, verification of whether the incidence of poverty⁷⁴ is *within or outside* the range initially selected is done (in this case, it is verified whether the incidence of poverty is between 30% and 59%).

If the incidence of poverty falls *within* the range of the reference population, the process ends. Otherwise, a new population must be chosen and it will be located within the range having the incidence of poverty calculated in the immediately preceding iteration as its center⁷⁵.

The process of construction of poverty lines will be completed after ensuring the validity of the selected reference population.

⁷³ This method was proposed by Pradhan, Suryahadi, Sumarto and Pritchett (2001).

⁷⁴ This incidence will be calculated from the poverty lines constructed as described in Sections 5.5 to 5.6 below.

⁷⁵ The verification of this process for the selection of the reference population is described in Section 5.11.

5.5 Construction of basic food baskets

Basic food baskets for each domain are built with the information of household consumption habits of each domain (urban and rural) belonging to the national reference population. This step consists of obtaining a short list of food items (or a subset of the total consumed food items) selected by applying three inclusion criteria⁷⁶: i) It is consumed by 30% or more of households; ii) It represents at least 1% of total food expenditure; and iii) It provides at least 1% of the total of calories consumed⁷⁷. Food items that are part of the diet of most households are selected under the first criterion. The second criterion ensures the inclusion in the basket of food items that have an important weight in monthly household expenditure. The last criterion includes food items that provide more calories in the diet of Colombian households. This includes all food items that meet at least one of the three criteria.

After applying those criteria, the basic basket of the urban area is composed of 39 food items and the rural basket of 42 food items. Annex G lists the items that are part of each one of the baskets.

5.6 Adjustment of the basic food basket to meet calorie requirements

Once the basic food basket is obtained, the calorie content per capita of the two baskets (urban and rural) is calculated by using the observed quantities in the reference population, for then comparing the calorie content with the calorie standard shown in Table 4. The standard calorie domain was obtained from the population structure by age and sex of the ENIG for the 2006-2007 period when assigning nutritional recommendations provided by FAO, WHO and the United Nations (2001). That is, the allocation of calorie requirements for each person registered in the survey was made in order to obtain an average for each domain (urban and rural). It should be noted that compared to the previous methodology for measuring poverty (cf. Muñoz and Rivas, 2006), urban calorie requirements are lower by 10%, due to a reduction of calorie requirements in the tables appearing in FAO *et al* (2001) that range between 9% to 19% for different age groups and sex⁷⁸.

The adjustment of quantities consists of applying the factor resulting from dividing the average calorie requirement per domain and the observed calorie content. The resulting adjustment

⁷⁶ Three of the four criteria used by Muñoz and Rivas (2006) for the selection of food items were followed. The authors applied as an additional criterion for the inclusion of food the one which represented 0.5% or more of the total weight of food acquired by households. Similar criteria are used by ECLAC as part of its methodology for measuring monetary poverty. See Section 4.4 of this document.

⁷⁷ The allocation of energy intake of each food was taken mainly from food composition tables from ICBF (Instituto Colombiano de Bienestar Familiar- the Colombian Institute for Family Welfare) for the years 2000 and 2005, and other sources such as FAO and DANE.

⁷⁸ Calorie requirement by sex and age can be found in detail in Annex H.

factor was 1.97 in the urban domain and of 0.92 in the rural domain⁷⁹. Thus, the adjusted basic food basket of each domain enables meeting the corresponding calorie requirement.

Domains	Calories
Urban	2090
Rural	2049

Source: MESEP, based on ENIG for the 2006-2007 period and FAO *et al* (2001).

5.7 Assessment of the adjusted basic food baskets

After establishing the items and quantities that compose the adjusted basic food basket in each domain, its cost is determined in order to obtain the respective extreme poverty lines. Implicit prices that were calculated from data of ENIG 06/07⁸⁰ are used to assess the baskets (the medians of unit values). The cost of the adjusted food basket by domain corresponds to the value of the extreme poverty line in Colombian pesos of March 2007 and is presented in Table 5.

Domains	per capita IL	Average Household IL⁸¹
Urban	\$73,984	\$295,936
Rural	\$60,968	\$243,872

Source: MESEP based on ENIG for the 2006-2007 period.

5.8. Transfer from IL to PL using the Orshansky coefficient

To obtain the value of the poverty line, the values of extreme poverty lines are multiplied by the Orshansky coefficient (OC).

A fraction of total household expenditure is allocated to food and non-foods. The proportion of non-food expenditure can be deduced by knowing the share of food expenditure in the total expenditure of the reference population. Thus, the Orshansky coefficient is estimated as the ratio of total expenditure and food expenditure of the reference population⁸². Thus, the value of

⁷⁹ This factor must not be interpreted as an indicator of calorie deficit of the reference population, since it pertains to the calorie content of a food subset of the total that is actually consumed by the population.

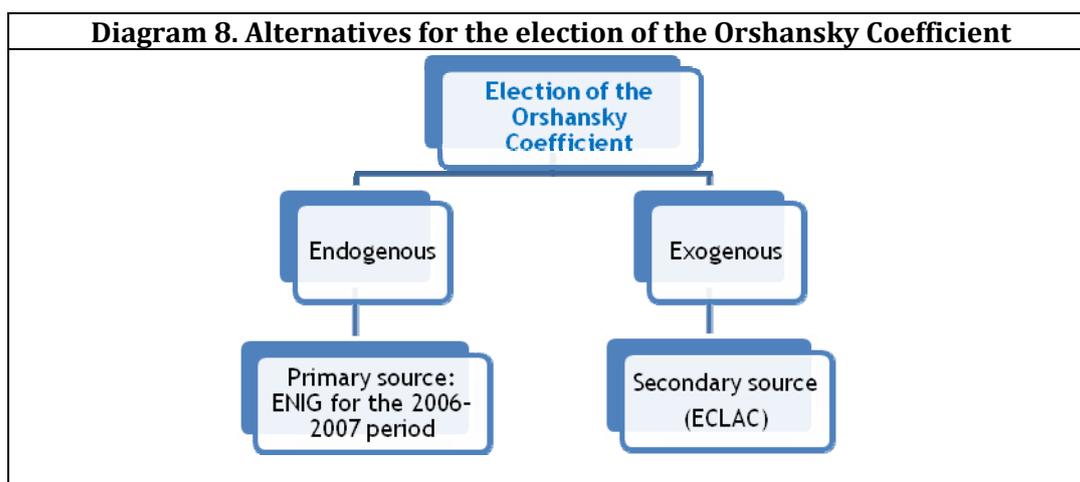
⁸⁰ The median of implicit prices was considered because it is a measure of central tendency that is robust to outliers.

⁸¹ A household consisting of 4 individuals (according to DANE's household surveys, the average number of persons per household in Colombia is about 3.7).

⁸² It has already been mentioned that the establishment of the pattern of food consumption is possible because there is information with respect to the expenditure on each food item and the quantity acquired in each case. For non-food expenditure, only information on

the poverty line is obtained by multiplying the value of the extreme poverty line by the Orshansky coefficient: $PL=IL*OC$.

As mentioned above, there are two possibilities for estimating the Orshansky coefficient: endogenously or exogenously (Diagram 8). In the first case, the calculation is made based on survey information for the design of the lines (in the case of Colombia, ENIG for the 2006-2007 period). The second case consists in using secondary information. For this second alternative, MESEP analyzed data estimated by ECLAC for Latin American countries.



Source: MESEP technical Team

The values of the endogenous Orshansky coefficients (ENIG for the 2006-2007 period) are presented in Table 6. Estimates of Orshansky coefficient for the nineties (cf. Muñoz and Rivas, 2005) that are based on IES for the 1994-1995 period indicate that the average in the cities and rural areas is of 2.5 and 2.2 respectively.

Domains	Endogenous Orshansky
Urban	2.87
Rural	2.09

Source: MESEP based on ENIG for the 2006-2007 period.

In the revision that is being performed by ECLAC to the methodology of poverty measure and the values of the lines, a preliminary estimation of the endogenous Orshansky coefficients for

expenditure on each good is available, but not on the amount associated. The value of the pattern of consumption of non-food goods of the reference population, which has a food basket that meets calorie requirements, can be deduced by using the Orshansky coefficient.

each country in Latin America was made based on recent income and expenditure surveys. In urban areas, the endogenous Orshansky coefficient varies between 1.9 and 3.5, with an average of 2.4, whereas in rural areas it varies between 1.8 and 3.2, with a simple average of 2.1.

From these results, and following the purpose of applying a standard methodology for all countries in Latin America, ECLAC is considering the alternative of using fixed coefficients but slightly higher than the current ones (2.0 and 1.75); possibly 2.4 in urban areas and 2.1 in rural areas.

Both estimates from ENIG (2006-2007) and the ones resulting from IES (1994-1995) exceed the endogenous values estimated by ECLAC. These overestimations of Orshansky coefficient arise in part because IESs tend to underestimate the total food expenditure. This is largely due to the change in frequency of purchase, changing to longer periods (biweekly, monthly, bi-monthly instead of weekly), to the purchase of groceries and the consumption of meals outside the home. To capture all the items of expenditure for food in the presence of the above-mentioned changes constitutes the biggest challenge that IESs are currently facing in various Latin American countries.

In order to use a Orshansky coefficient that was more comparable to the average of the countries in the region, MESEP decided to *adopt an exogenous Orshansky coefficient of 2.4 for the urban area, equal to the Latin American average* (see Table 7). For rural areas⁸³, MESEP took the exogenous Orshansky coefficient as a starting point, but kept the proportionality urban / rural observed in the endogenous Orshansky coefficients. The resulting value of the Orshansky coefficient is 1.74⁸⁴. The new values of the poverty lines for Colombia (in Colombian pesos of March 2007) are presented in Table 8.

⁸³ Making the most of the information in ENIG (2006-2007) was the reason why a OC that kept the proportionality found from estimates of endogenous coefficients (1.74) was chosen for the rural areas.

⁸⁴ This value is obtained by dividing the value of the urban exogenous OC (2.4 average for Latin America) by the value of the ratio urban / rural of endogenous OCs (ENIG 2006-2007).

Table 22. Endogenous Orshansky coefficients for Latin American countries, ECLAC		
Countries	Urban area	Rural area
Argentina 1996	2.6	
Bolivia 2003	2.2	1.9
Brazil 2002	3.5	2.7
Chile 2007	3.0	
Costa Rica 2004	2.8	2.3
Ecuador 2003	2.2	
Guatemala 2006	2.5	2.0
Honduras 2004	2.2	2.0
México 2006	2.8	2.5
Nicaragua 2005	1.9	1.7
Panamá 2007	2.8	
Paraguay 1997	2.4	1.9
Perú 2008	2.2	1.8
Dominican Republic 1998	2.3	2.1
Uruguay 2006	3.4	3.2
Average with debugged pattern	2.4	2.1

Source: Information provided by members of the ECLAC. ECLAC's preliminary estimations. The average with debugged pattern refers to the exclusion of some non-essential items by ECLAC.

Table 23. Exogenous Orshansky coefficients and poverty lines by month (in Colombian pesos of March 2007)			
Domains	Exogenous Orshansky	Per capita PL	Average household PL⁸⁵
Urban	2.4	\$ 177,562	\$710,248
Rural	1.74	\$ 106,084	\$424,336

Source: MESEP based on ECLAC and ENIG 2006-2007.

⁸⁵ A household composed of four persons is taken as reference.

5.9. Considerations with respect to the construction of the new poverty lines

Why use data from ENIG (2006/2007) to build the basic food basket and not for the endogenous Orshansky coefficient?

Cortés' analysis shows the internal consistency of ENIG (2006/2007), this enables the use of survey data for the construction of the basic food basket (collected data follow the empirical regularity of Engel's Law). Cortés also found that there is an underestimation of food expenditure (external consistency), which is certainly a limitation. In the case of the construction of the basic food basket, this limitation is mitigated by the regulatory adjustment of quantities, unlike the calculation of the endogenous Orshansky coefficient, for which the underestimation of food expenditure is, by construction, an overestimation of the Poverty Line. For this reason, MESEP considered it appropriate to choose an exogenous coefficient for the calculation of the poverty line.

What countries use exogenous Orshansky coefficient⁸⁶?

According to Medina and Galván, from 17 countries in Latin America, 7 of them use this pattern in an exogenous manner for the assessment of the poverty line. Other countries with an Orshansky that is exogenously determined are: Ecuador, El Salvador, Venezuela, Brazil, Chile and Honduras (see Table 9).

⁸⁶ When the exogenous OC is mentioned, it refers to the coefficient used in the urban area. In the rural area, the coefficient used was the one calculated according to what was previously described, that is the rural area coefficient that keeps the urban/rural proportionality found in the endogenous coefficients that were calculated based on the information of ENIG for the 2006-2007 period.

Table 24. Orshansky Coefficients by country

Country	Orshansky
Argentina*	Endogenous
Bolivia*	
Costa Rica*	
México*	
Nicaragua	
Panamá	
Paraguay*	
Perú	
Dominican Rep.	
Uruguay*	
Ecuador	Exogenous
El Salvador*	
Venezuela*	
Brazil*	
Chile*	
Honduras*	
Colombia* ⁸⁷	

*Countries that use income as welfare variable, for the estimation of the incidence of poverty.

Source: Information provided by members of the ECLAC.

Calculations by Medina, Fernando; Galván, Marcos.

5.10. Estimation of the incidence of poverty

Finally, in order to obtain the incidence of extreme poverty and poverty for a given year, the monthly per capita income of the spending unit (PCISU)⁸⁸, calculated from household surveys of the respective year, is compared with the values of the poverty and extreme poverty lines updated due to inflation. Households are classified into three categories: *not poor*, if income is higher than the PL; *poor*, if income is less than the PL, and *extremely poor* if the per capita income is below the IL⁸⁹. The incidence of poverty is equal to the proportion of the population living in a situation of poverty and the incidence of extreme poverty is equal to the proportion of the population living in a situation of extreme poverty.

- Incidence of Extreme Poverty (%)

⁸⁷ It is worth reiterating that the OC is exogenous only in the urban area. The rural OC is obtained by multiplying the ratio of urban / rural endogenous OCs by the urban exogenous OC:

$$\frac{\text{rural endogenous OC ENIG}^{\frac{86}{87}}}{\text{urban endogenous OC ENIG}^{\frac{86}{87}}} = \text{urban exogenous OC}^{\frac{86}{87}}$$

⁸⁸ Described in Section 2.5.

⁸⁹ The poor and extremely poor categories are not exclusive. The extremely poor are included in the poor category.

$$\frac{\text{Extremely poor population}}{\text{Total Population}}$$

- Incidence of Poverty (%)

$$\frac{\text{Poor population}}{\text{Total population}}$$

5.11. Validity verification of the estimates and the selected reference population

As explained in Section 6.4, the method proposed by Ravallion to construct poverty lines follows an iterative process. The starting point is the selection of an *a priori* population of reference, centered on a value corresponding to the incidence of poverty that is presumed to be the most feasible (it should be remembered that the value initially assumed was 45%, which corresponds to the estimated incidence in the first phase of MESEP for the year 2009). Subsequently, the described steps are developed to build the extreme poverty line (IL), the Orshansky coefficient and the total poverty line (PL) are implemented. Once their values are determined, their corresponding incidences are estimated.

To prove the validity of the obtained estimates and the selected reference population, the resulting incidence is compared with the initially assumed percentage (45%). If the new percentage of poor falls within the interval defined for the *a priori* reference population (between 30-59 percentiles), this shows that the poverty and extreme poverty lines are consistently built. If the calculated incidence is out of the interval, the process should be repeated, selecting a reference population centered in the new incidence percentage found with the new lines. In the case of Colombia, the process was done once, obtaining an incidence of 49%. Since this incidence was within the 30th to 59th interval, additional iterations were not necessary.

Since the 2006-2007 ENIG was conducted between October 2006 and September 2007, the values of the lines are centered in March 2007. In order to estimate the incidence of poverty for each year, centered values in March 2007 must be updated (backward and forward) to carry out the calculations. The methodology to update the values of the lines due to inflation is described below.

5.12. Update of the poverty lines due to monthly inflation

As part of the new methodology for measuring poverty in Colombia, the monthly update process of poverty and extreme poverty lines was changed. This section presents the new update methodology, which allows taking into account price changes of each domain.

New domains

Given the expansion of the geographical disaggregation levels implemented by DANE with GEIH starting from 2006 (from 13 to 24 cities), MESEP decided to expand⁹⁰, in the same amount, the number of domains in order to update the poverty and indigence lines, which resulted in 11 new geographic domains, for a total of 26 domains that are listed in Table 10:

	13 traditional areas	New cities of GEIH	Other geographic domains
	Medellín Barranquilla Bogotá Cartagena Manizales Monteria Villavicencio Pasto Cúcuta Pereira Bucaramanga Ibagué Cali	Tunja Florencia Popayán Valledupar Quibdó Neiva Riohacha Santa Marta Armenia Sincedejo San Andrés	Other municipal townships Rural

Source: MESEP technical team

Base year and month

Since the last IES conducted by DANE was between October 2006 and September 2007, the average expenditure and income were determined approximately in the month of March 2007. This is the base period of the new poverty lines.

Assigning initial values

The assignment of initial values in the various domains was done as follows: for the 25 urban domains, \$73,984 and \$177,561.6 were taken as values of IL and PL respectively. For the rural area, \$60,968 and \$106,084.3 were the values designated for IL and PL respectively.

⁹⁰ The decision to expand the number of domains for the updating obeys to the availability of information from CPI.

Price indices used

Once the values of the poverty and extreme poverty lines are obtained for the period in which ENIG was applied, the values of these lines for the other years involved in the analysis are calculated (i.e., the years before and after ENIG for the 2002-2005 period and the 2008-2010 period). MESEP decided to use, in domains where the indicator exists, the total and food (low income group) consumer price indices (CPI) to update the monthly values of the PL and IL⁹¹ respectively. In domains where there is no price index (*other municipal townships and rural areas*), using a weighted average of the cities of Manizales, Montería, Neiva, Villavicencio and Pereira for each index (for food and total) was decided. These cities were considered because they were the same cities considered by *the Mission for the Design of a Strategy for Reducing Poverty and Inequality (MERPD)* and for being the smallest cities among the thirteen major cities. For the latter aspect, these cities represent a proxy of the behavior of price changes in the *Rural and other municipal townships domains*. The weights were obtained from the total current expenditure in the first 50 percentiles of the five cities, in order to approximate the low-income range of CPI. Table 11 shows the weights that were used.

Table 26. Price index weights for <i>Remaining municipal townships and Rural domains</i>	
City	Weight
Manizales	21.4%
Montería	10.4%
Neiva	14.9%
Villavicencio	22.1%
Pereira	31.2%
Total	100.0%

Source: MESEP technical team

It should be taken into account that DANE started measuring the CPI for twenty-four cities as of December 2008: before this date only information on CPI for thirteen cities existed. Therefore, to obtain the values of IL and PL in this period, the same weighted average of food and total indices of the eleven additional cities was used. As a result of these updates due to monthly inflation, official values for the new poverty and extreme poverty lines from 2002 to 2010 for urban and rural domains were obtained (see Annex I).

⁹¹ This procedure was being applied in the former methodologies for measuring poverty in Colombia.

5.13. Summary: Main characteristics of the new poverty line

Comparison of the new and old methodologies

The main differences of the new methodology when compared with the previous one are:

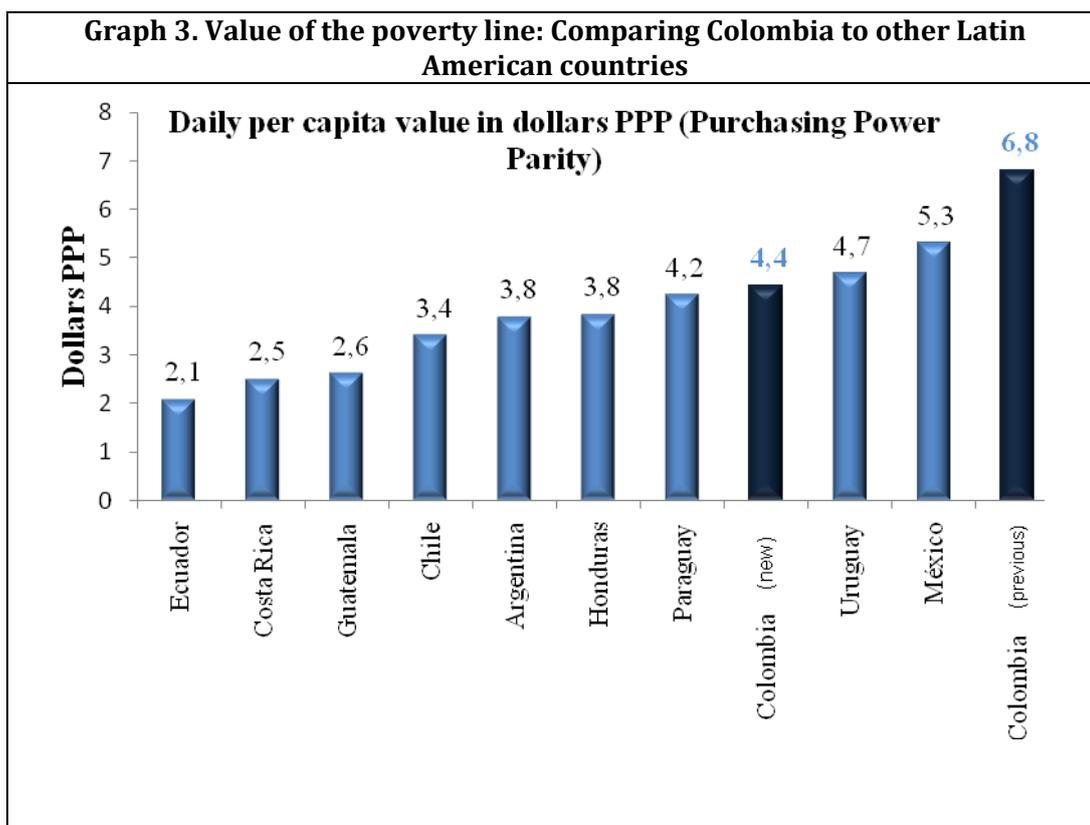
- For the construction of the new methodology, a more updated source of consumption habits of the population is considered (ENIG 2006-2007. Vs. ENIG 1994-1995).
- There are less domains: two domains for the new line (Urban and Rural), which contrasts with fifteen domains of the previous one (thirteen Metropolitan Areas, other Municipal Townships and Rural Area).
- For the first time the same survey is used (ENIG for the 2006 - 2007 period) to build the Urban and Rural ILs. In the previous methodology, information from the first follow-up of the Familias en Acción (*Families in Action*) - 2003 program was used for the construction of the rural poverty line.
- Of the five criteria established for inclusion with respect to food items in the previous methodology, only three criteria remain: in the new methodology, the following criteria is applied: included food should be consumed by a representative set of households (30%), or represent at least 1% of current expenditure, or represent at least 1% of consumed calories. The criteria of minimum contribution to protein and grams of consumed food are no longer applied.
- The reference population of the new methodology does not include the extremely poor or the wealthier households and seeks to represent the habits of the population of interest that is closest to the cut-off point of poverty (30-59 percentiles). The previous methodology focused on the habits of the population in extreme poverty (first quartile of expenditure).
- The new methodology uses a spatial price deflator to sort *the entire population* under the same criterion of per capita expenditure. The previous methodology did not use spatial deflators. In consequence, the new methodology uses *a single reference population*, whereas the previous methodology used 13 reference populations.
- The implicit prices for the assessment of the adjusted basic basket: the *median* was used in the new methodology and the *mean* in the previous methodology.
- The urban Orshansky coefficient is exogenous in the new methodology (2.4) and was endogenous in the previous methodology (2.5 urban average). The rural Orshansky coefficient (1.74) combines information of endogenous Orshansky coefficients calculated in ENIG for the 2006-2007 period and the urban exogenous Orshansky

coefficient based on ECLAC's average. In the previous methodology, the rural Orshansky coefficient was endogenous (2.2).

- The update of IL and PL due to price inflation is equal in both methodologies: the PL is updated with CPI for the low-income group and the IL with the food price index for the same income group.

Value of the new line in the regional context

Colombia's new poverty line is the third highest line compared to that of nine other Latin American countries and to the corresponding value of the previous methodology. As shown in Graph 3, Colombia's new poverty line is 4.42 PPP dollars per person, per day, and surpasses the PL values for Ecuador, Costa Rica, Guatemala, Chile, Argentina, Honduras and Paraguay. It is only surpassed by the lines of Uruguay and México that reach values of 4.7 and 5.3 PPP dollars per person, per day respectively. The same graph also shows how the previous PL of Colombia surpassed all the countries in the region.



Source: Data from Colombia is taken from DNP's DDS* and SPSCV** based on the values of MESEP's lines (Phase II) and the PPP dollar value based on World Bank's information. Data from other Latin American countries are taken from Azevedo *et al* (2009), p. 18.

Note: In order to make comparisons between proximate years, data for Colombia corresponds to the year 2005. Data for the other countries are for the year 2006.

* Social Development Division

** Social Promotion and Quality of Life Subdivision

6. Levels of poverty and inequality with the new methodology

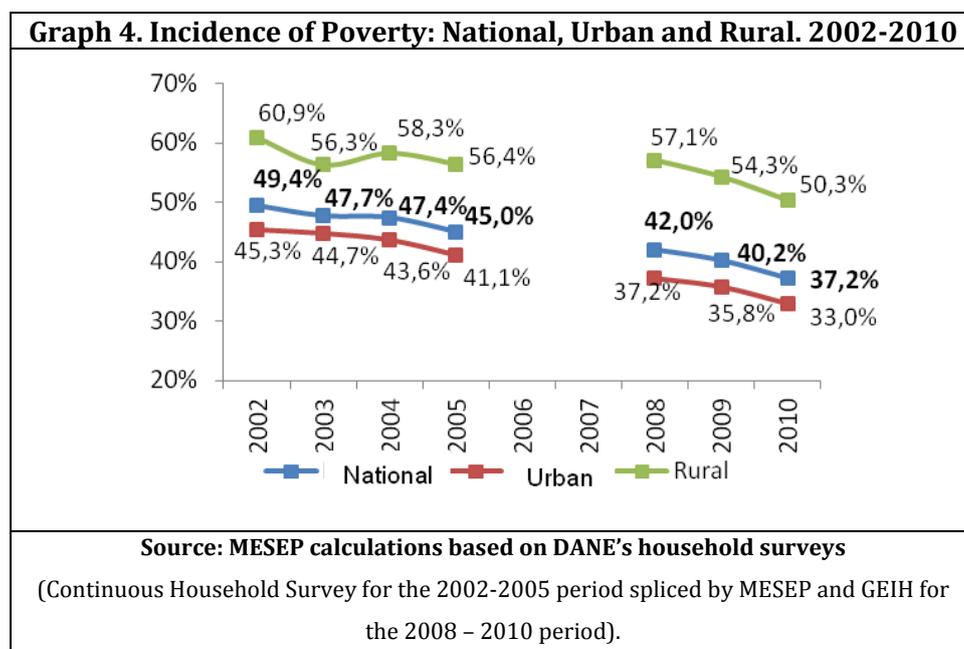
6.1. Poverty for the 2002-2010 period

By the year 2010, national poverty reached an incidence of 37.2%; while urban and rural poverty incidence were of 33% and 50% respectively (see Table 12). In the same year, *extreme* poverty was 12.3%, 8.1% and 25.5% respectively. Graph 4, which presents the evolution of poverty over the last decade (2002 to 2010), shows a reduction in the incidence of the national poverty of 12.2 percentage points (from 49.4% to 37.2%)⁹². For the last two years of the series, there is a decrease of three percentage points of the incidence of national poverty (in 2009, the percentage of poor people was 40.2%).

	National	Urban	13 MA	Remaining urban areas	Rural
Poverty	37.2%	33.0%	23.2%	47.3%	50.3%
Extreme poverty	12.3%	8.1%	4.6%	13.2%	25.5%

Source: MESEP calculations based on DANE's household surveys

(ECH for the 2002-2005 period spliced by MESEP and GEIH for the 2008 -2010 period)



⁹² The series does not include data for the years 2006 and 2007. See MESEP (2010).

Poverty has decreased more in urban areas than in rural areas. While in the urban area the reduction of poverty was 12.3 percentage points in the period 2002-2010 (went from 45.3% to 33%), poverty in the rural area decreased by 10.6 percentage points (from 60.9% to 50.3%). The differential between the two areas has widened: while in 2002 the incidence of rural poverty was equivalent to 1.3 times urban poverty, in 2010 the incidence of rural poverty is 1.5 times urban poverty. Despite these contrasts, the greater dynamism registered in rural areas during the last three years should be noted. Between 2008 and 2010, the incidence of rural poverty decreased from 57.1% to 50.3% (minus 6.8 percentage points). In the same period, urban poverty declined from 37.2% to 33% (minus 4.2 percentage points).

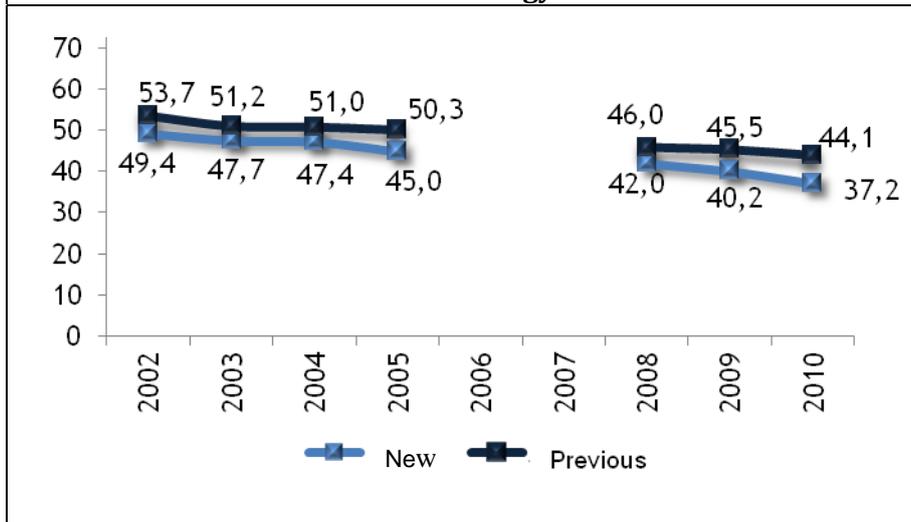
While in urban areas the reduction is significant, there are marked differences between the 13 major cities (23.2% in 2010) and other urban townships (47.3% in 2010). In 2010 the incidence of poverty in urban townships other than the 13 major cities is close to that of the rural area. These cities register the highest reduction in percentage points (minus 12.7 percentage points between 2002 and 2010) while, at a time showing the lowest incidence of poverty of all geographic domains. Bucaramanga and Bogotá have the lowest incidences (10.8% and 15.5% in 2010, respectively). The highest were recorded in Pasto (42.6%), Barranquilla (39.5%) and Montería (39.5%)⁹³.

Comparison of methodologies

The comparison between the poverty series of 2002-2010, which resulted from the two methodologies shows very similar trends in all domains, with lower levels of poverty under the new methodology. Graphs 5 and 6 show how the new series in all domains (except other municipal townships or remaining urban areas) are below the series that are based on the previous methodology. By the year 2010, poverty levels with previous methodology surpassed the measurements of new methodology in 6.9 percentage points for the national domain, 5.4 percentage points for the urban domain, and 12.2 percentage points for the rural domain. However, these differences are reduced in 2002, except for other municipal townships (Remaining urban areas).

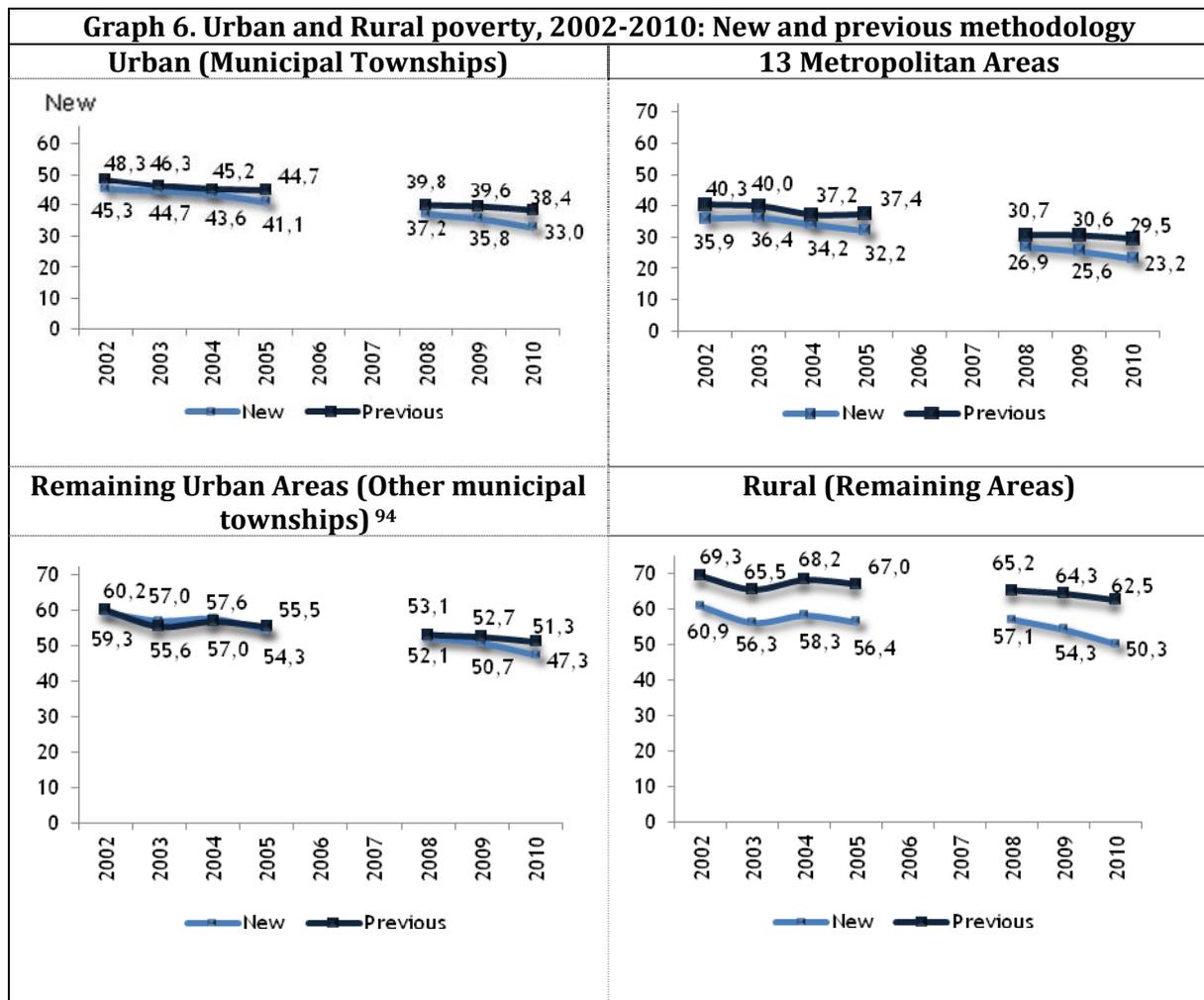
⁹³ Refer to Annex J. For information on departmental figures see Annex K.

Graph 5. National Poverty 2002-2010: New and previous methodology



Source: MESEP based on DANE's ECH for the 2002-2005 period and GEIH for the 2008-2010 period.

Graph 6. Urban and Rural poverty, 2002-2010: New and previous methodology



Source: MESEP based on DANE's ECH for the 2002-2005 period and GEIH for the 2008-2010 period

⁹⁴ The domain "Other municipal townships" refers to municipal townships other than the thirteen metropolitan areas.

On the other hand, poverty measurements with the new methodology indicate steepest declines in the 2002-2010 period for all domains. As shown in Table 13, national and urban poverty reductions under new methodology exceed by more than two percentage points the reductions registered with previous methodology⁹⁵. In rural areas the difference is close to four percentage points.

Table 28. Changes in poverty level 2002-2010: Comparison between new and previous methodology

Changes 2002-2010	National	Urban	Rural
	<i>(percentage points)</i>		
Previous Methodology	-9.6	-9.9	-6.8
New Methodology	-12.2	-12.3	-10.5

Source: MESEP based on DANE's ECH for the 2002-2005 period and GEIH for the 2008-2010 period

6.2. Extreme Poverty

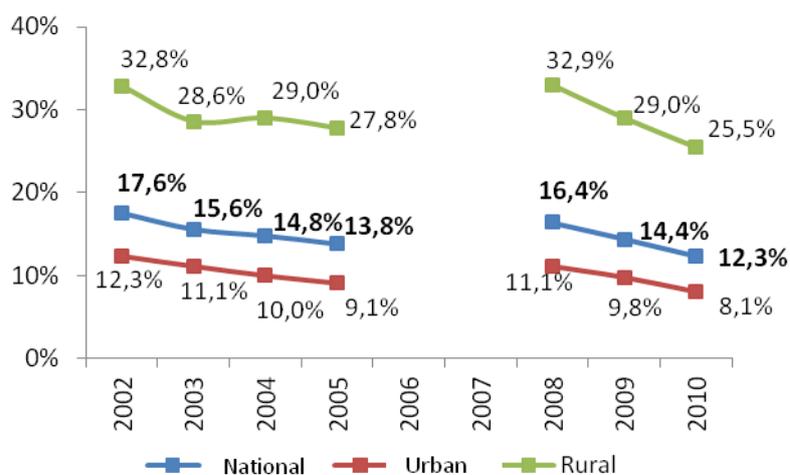
The reduction of extreme poverty has been less pronounced than that of poverty. During the last decade, extreme poverty fell 5.3 percentage points between 2002 and 2010 (from 17.6% in 2002 to 12.3% in 2010), and between 2009 and 2010 this decline was of 2.1 percentage points (from 14.4% to 12.3%). Between 2002 and 2010 the reduction of urban extreme poverty was less than the reduction of the rural one (minus 4.2 percentage points vs. minus 7.3 percentage points). However, the differences between the two areas in terms of incidence remain large and are more marked than those of poverty. While in 2002 the incidence of extreme rural poverty represented 2.6 times the incidence of extreme urban poverty (32.8% vs. 12.3%), in 2010 this ratio increased to 3.1 times (25.5% vs. 8.1%). There is concern regarding the fact that extreme poverty has increased in all domains between 2005 and 2008. (Refer to Graph 7)

With respect to the 13 largest cities, the percentage of extreme poor population fell by approximately 40%. Between 2002 and 2010 the incidence decreased from 7.4% to 4.6% (minus 2.8 percentage points). In other urban townships, the reduction was 6.4 percentage points (19.6% in 2002 and 13.2% in 2010). Currently, the incidence of extreme poverty of other urban townships is 2.8 times the incidence of poverty in the 13 cities. In the same manner as with the incidence of poverty, Bucaramanga and Bogotá register the lowest incidence of extreme poverty (1.2% and 2.6% in 2010, respectively), whereas Cucuta and Pasto register the highest incidence (11.3% and 8.4% in 2010, respectively)⁹⁶.

⁹⁵ A decomposition of the differences between the two methodologies is presented at the end of this section.

⁹⁶ Refer to Annex J.

Graph 7. Incidence of extreme poverty: National, Urban and Rural. 2002-2010



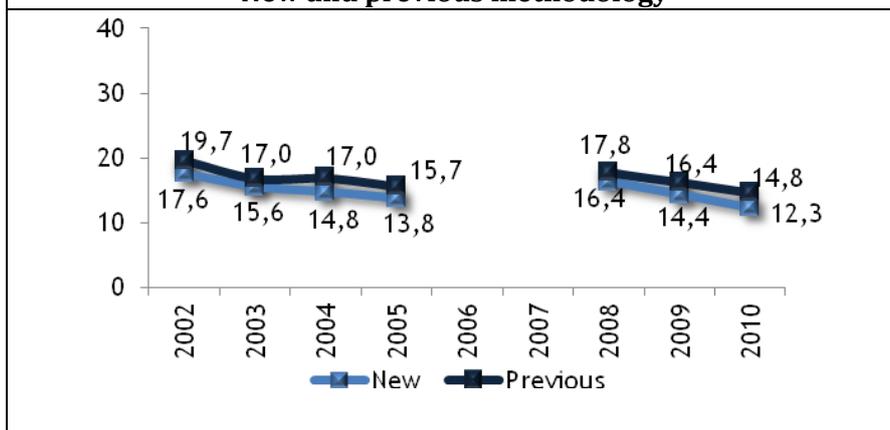
Source: MESEP's calculations based on DANE's household surveys

(ECH for the 2002-2005 period spliced by MESEP and GEIH for the 2008 - 2010 period).

Comparison of methodologies

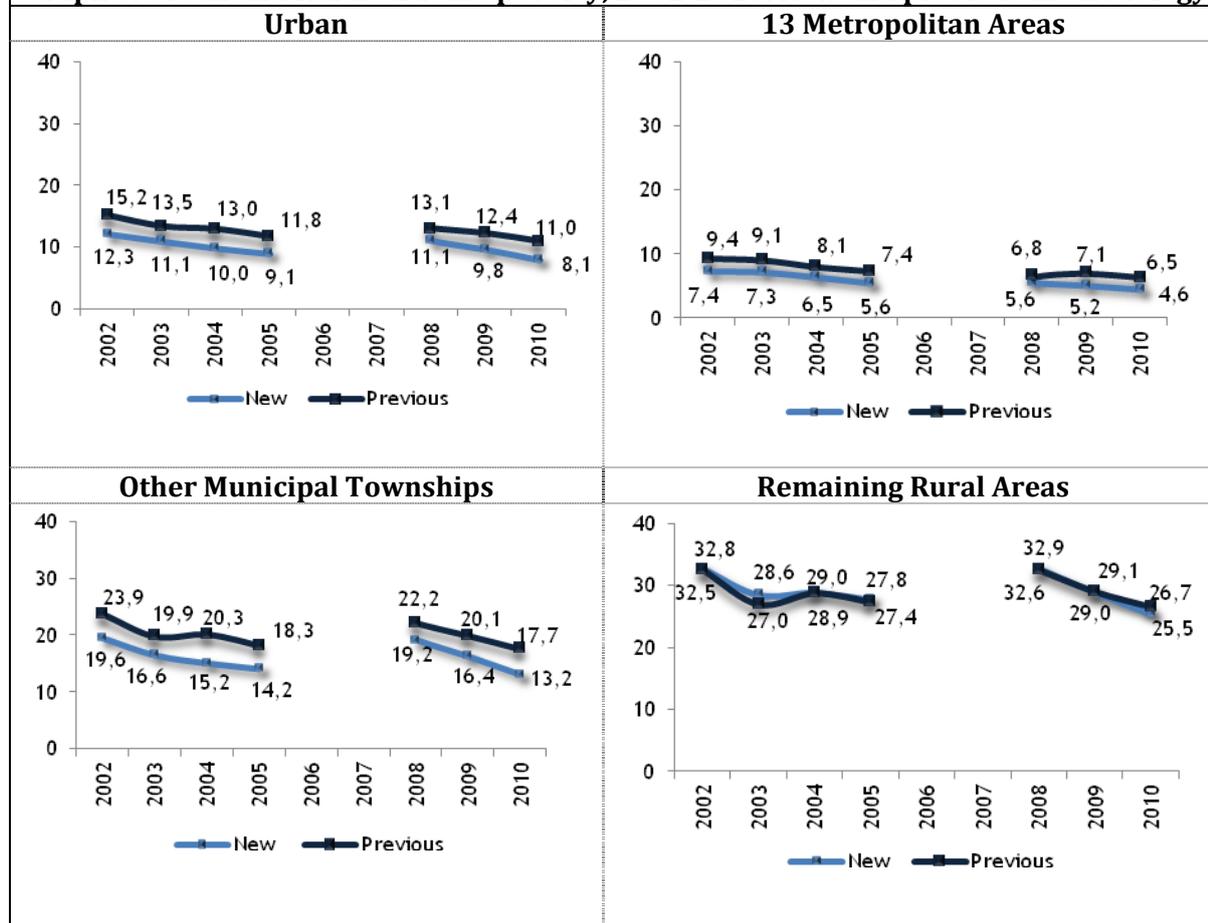
As observed for the measures of poverty, the comparison of extreme poverty series with both methodologies shows very similar trends, with lower levels in all domains under the new methodology except for the rural domain. In effect, Graphs 8 and 9 show how National and Urban extreme poverty levels in 2010 with the new methodology are 2.3 and 2.9 percentage points below the measurements with the previous methodology. However, the difference in the measurement of rural extreme poverty for the same year barely reaches 0.3 percentage points.

**Graph 8. National extreme poverty for the 2002- 2010 period:
New and previous methodology**



Source: MESEP based on DANE's ECH for the 2002-2005 period and GEIH for the 2008-2010 period

Graph 9. Urban and Rural extreme poverty, 2002-2010: New and previous methodology



Source: MESEP based on DANE's ECH for the 2002-2005 period and GEIH for the 2008-2010 period

A rural-urban asymmetry can be observed when comparing the dynamics of the reduction of extreme poverty with both methodologies. While the reductions of urban extreme poverty in the 2002-2010 period are slightly more pronounced with previous methodology (4.2 vs. 4.0 percentage points), the largest declines for the Rural domain are registered with the new methodology (5.9 percentage points vs. 7.2).

Table 29. Changes of the extreme poverty level for the 2002-2010 period: Comparison of new and previous methodology

Changes 2002-2010	National	Urban	Rural
	(Percentage points)		
Previous Methodology	-4.9	-4.2	-5.8
New Methodology	-5.3	-4.2	-7.3

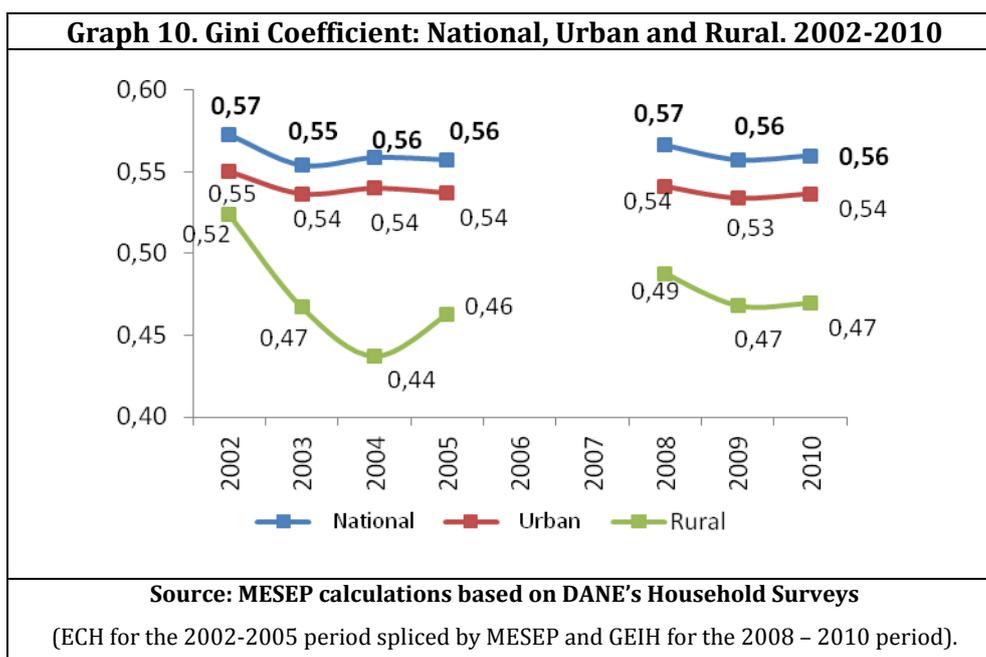
Source: MESEP Technical Team

6.3. Income inequality (Gini coefficient)

In Colombia, income inequality has been particularly high and under the new methodology, the Gini coefficient, which measures such inequality, hardly shows a slight decrease for the period under analysis (decreased from 0.573 to 0.560. See Graph 10). However, income inequality in Colombia remains one of the highest in the Latin American context⁹⁷. In urban areas, the level and evolution of the Gini coefficient has been very similar to that of the national level: for the 2002-2010 period, the coefficient fell 1.3 points (from 0.550 to 0.537). In the 13 metropolitan areas domain, a similar level of the coefficient can be observed, but with a more pronounced reduction (three points between 2002 and 2010). It is worth pointing out the increase the indicator had in the last year (from 0.522 (2009) to 0.538 (2010)).

When decreasing the degree of urbanization, the Gini coefficient is lower, a result that is characteristic of areas where the incidence of poverty is higher. In urban townships other than the 13 major cities, the coefficient had a slightly decreasing behavior between 2002 and 2005 (from 0.493 to 0.482). From 2005 to 2008 the increase was significant, reaching a value of 0.521. As of 2008, the coefficient resumed its downward trend, and in 2010 reached a value of 0.502. A significant decrease can be observed in the Gini of the rural area between the 2002 and 2004 period (from 0.524 to 0.437). As of this last year, the trend is reversed and the coefficient reaches a value of 0.488 in 2008. Thereafter the indicator decreases and lies at levels of 0.47.

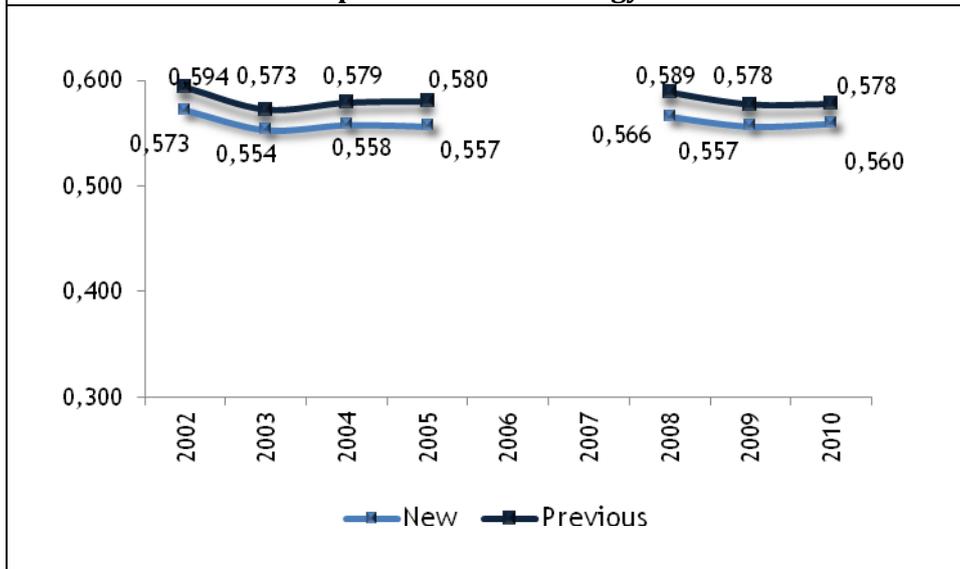
⁹⁷ According to ECLAC data, Colombia is the country with the highest inequality in recent years (Gini coefficient of 0.578 in 2010). Similarly, data from the World Bank (calculations based on the Socio-Economic Database for Latin America and the Caribbean - SEDLAC) puts Colombia in the first place with a Gini coefficient of 0.554 in 2010. The difference between the two data is due to the methodology used by each agency.



Comparison of methodologies

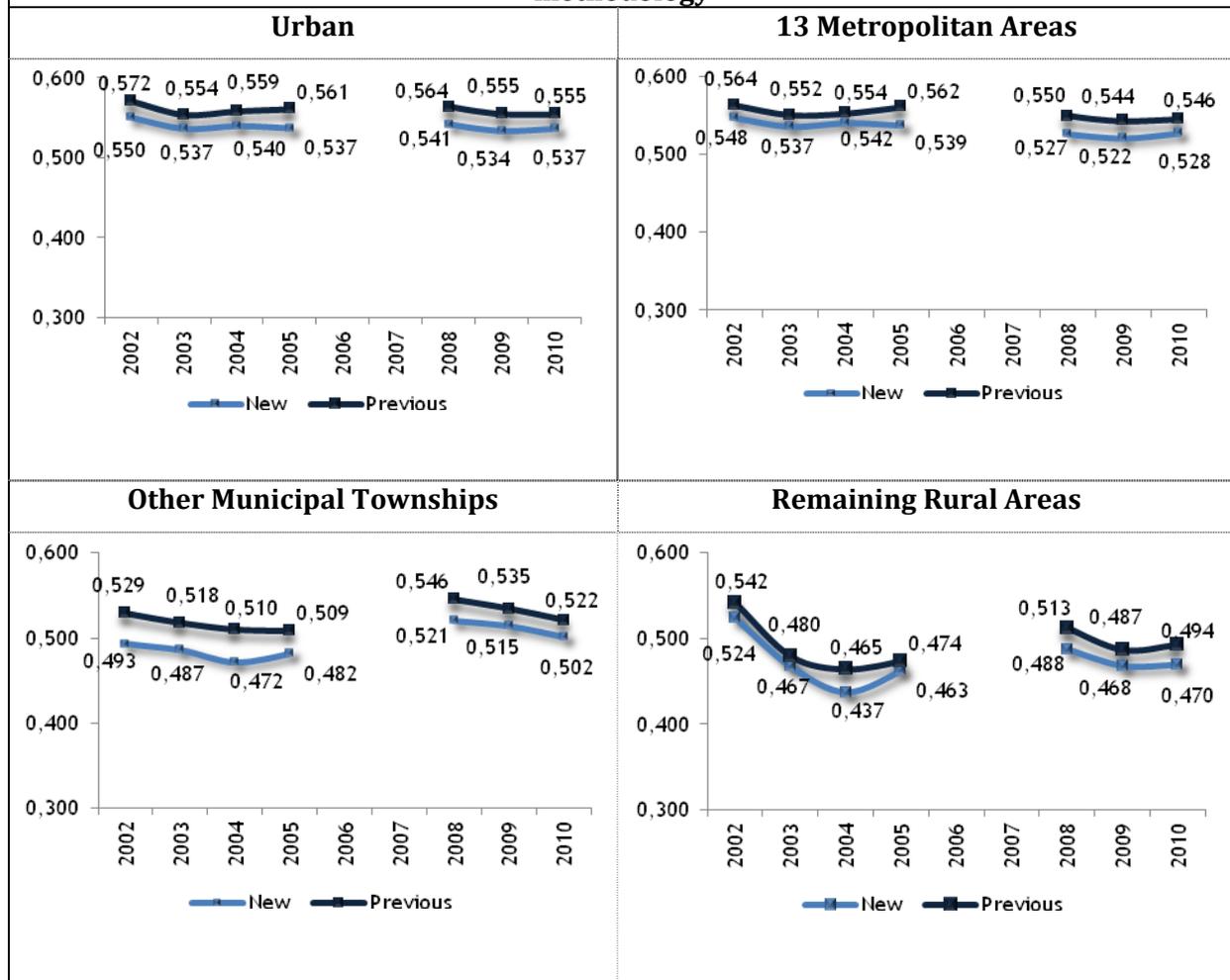
The series of unequal income distribution also show similar trends with both methodologies, but they register slightly lower levels with the new methodology (refer to Graph 11). By 2010, these differences are about two-hundredths in all domains (refer to Graph 12). With both methodologies, a reduction in income inequality can be observed. This reduction is slightly more pronounced in the rural area as compared to the 13 metropolitan areas, which show a slightly lower reduction.

Graph 11. Income inequality (Gini coefficient) 2002-2010: New and previous methodology



Source: MESEP based on DANE's ECH for the 2002-2005 period and GEIH for the 2008-2010 period

Graph 12. Urban and rural income inequality, 2002-2010: New and previous methodology



Source: MESEP based on DANE's ECH for the 2002-2005 period and GEIH for the 2008-2010 period

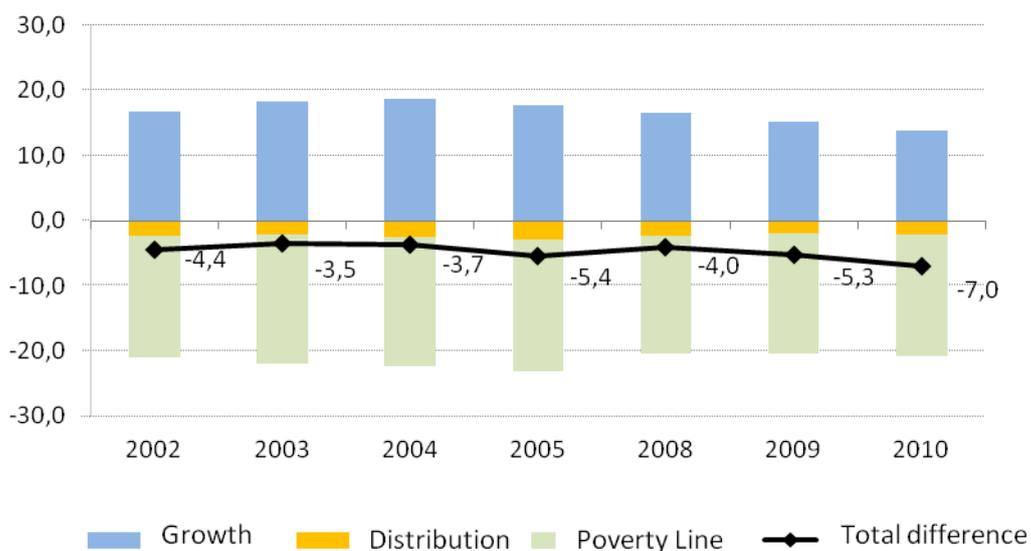
7. The impact of methodological changes on the measurement of poverty

Methodological changes have generated significant differences in the measurement of poverty and extreme poverty. As explained in the previous sections (see Sections 3 and 6), the new methodology introduced changes both in the poverty line as well as in the construction of the income (elimination of adjustment to National Accounts, imputation of income by homeownership and others). The question then arises: Which methodological changes resulted in a reduction or an increment in the measurement of poverty in Colombia? The decomposition exercises presented below indicate that the methodological changes have effects that are partially offset. Otherwise, the differences between the poverty series obtained from the two methodologies would be three or four times larger. In fact, while the new poverty line reduces the poverty measurement, changes in the construction of the income aggregate increase the poverty measurement (due to the reduction of the average income) and produce moderate reductions of poverty (via distributional effects).

Graphs 13 and 14 and Table 15 present the results of the Shorrocks and Kolenikov decomposition of the differences in the measurements of poverty and extreme poverty for the 2002-2010 period under the two methodologies. Graph 13 shows how the change in the poverty line (green bar) generates a *reduction* of 18 to 20 percentage points in the poverty level, which in turn is partially compensated by the reduction of the mean income resulting in an *increase* of 13.9 to 18.6 percentage points in the poverty level (blue bar). The effect of the changes in the distribution of income is added to the previously described components. Such effect *reduces* the measurement of poverty 2 to 3 percentage points, resulting in a *net difference* that oscillates between 4 and 7 percentage points (black line).

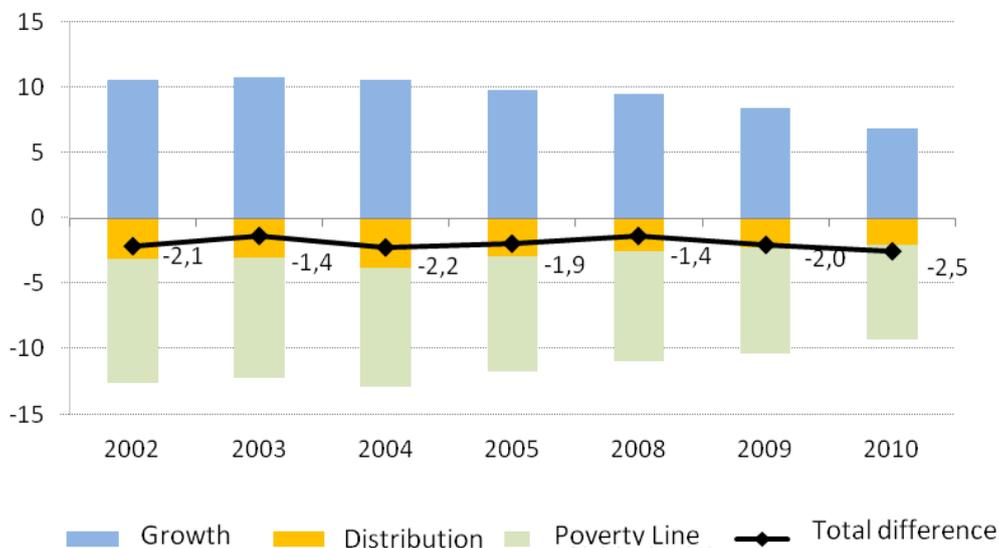
Similarly, Graph 14 shows that the change in the line (green bar) produces a *reduction* in the extreme poverty level of 8.2 to 9.5 percentage points, which is largely compensated by the reduction in the average income, resulting in an *increase* in the extreme poverty level (blue bar) of 6.9 to 10.8 percentage points. The final result is defined by the effect of the change in the distribution of income, which *reduces* the measurement of extreme poverty (orange bar) between 2.0 and 3.8 points and produces a *net difference* that oscillates between 1.4 and 2.5 percentage points (black line).

Graph 13. Decomposition of poverty measurement: Differences between the two methodologies



Source: Azevedo, 2011.

Graph 14. Decomposition of extreme poverty measurement: Differences between the two methodologies



Source: Azevedo, João. 2011. Technical Note for MESEP. World Bank, Washington, DC. (mimeo).

Table 30. Decomposition of extreme poverty and poverty measurements: Differences between the two methodologies

		2002	2003	2004	2005	2008	2009	2010
Extreme poverty	Total difference	-2.1	-1.4	-2.2	-1.9	-1.4	-2.0	-2.5
	Growth	10.6	10.8	10.6	9.8	9.5	8.4	6.9
	Distribution	-3.1	-3.0	-3.8	-2.9	-2.5	-2.2	-2.0
	Poverty line	-9.5	-9.2	-9.1	-8.8	-8.4	-8.2	-7.3
Poverty	Total difference	-4.4	-3.5	-3.7	-5.4	-4.0	-5.3	-7.0
	Growth	16.8	18.3	18.6	17.7	16.6	15.2	13.9
	Distribution	-2.3	-2.2	-2.6	-3.0	-2.4	-2.0	-2.1
	Poverty line	-18.8	-19.7	-19.7	-20.1	-18.1	-18.5	-18.7

Note: (*) Shorrocks and Kolenikov decomposition

Source: Azevedo, João. 2011. Technical Note for MESEP. World Bank, Washington, DC. (mimeo).

8. Conclusions and Recommendations

The new methodology for measuring monetary poverty and income inequality, provides an updated, more accurate and more comparable measurement of poverty, compared to that of other Latin American countries. Six aspects differentiate this new methodology:

1. The use of more current information on consumer habits;
2. From the statistical point of view, a more robust rural poverty line;
3. A more relevant reference population for the measurement of poverty;
4. An exogenous urban Orshansky coefficient equal to the Latin American average;
5. The elimination of the adjustment to National Accounts for the calculation of the household income;
6. The use of a more refined aggregate income, with more efficient imputation models that solve problems of outliers, missing data and suspicious zeros, and that include money transfers.

In the following, each of these aspects are explained in more detail.

Contemporary information on consumption patterns: The use of ENIG for the 2006-2007 period enabled using the latest information on the consumption patterns of the Colombian population for the construction of the new poverty lines. This facilitated the incorporation of changes in the household patterns of the last twelve years.

The decision of using the information of the ENIG for the 2006-2007 period for the construction of the new poverty lines underwent an assessment by Cortés (2009) and a subsequent analysis by the MESEP's expert group. Two aspects should be emphasized from the analysis of the source of information for the design of the new methodology:

- ENIG is internally consistent and meets the basic requirements for the construction of poverty lines, in particular, with the empirical regularity of Engel's Law.
- Some items of the 2006-2007 ENIG expenditure are underestimated, such as food items. This limitation of ENIG supposes restrictions for calculating an endogenous Orshansky coefficient. Consequently, MESEP adopted the use of an exogenous Orshansky coefficient in urban areas.

A statistically more robust rural poverty line: The rural poverty line of the previous methodology was estimated based on the first survey for monitoring the *Familias en Acción* program. This information only represented a subset of the rural population (the poorest, which

given their condition were the beneficiaries of *Familias en Acción*). The new methodology overcomes this limitation by including a rural poverty line constructed with the same source of information for urban lines (the 2006-2007 ENIG). This innovation guarantees better comparability of urban and rural data, and greater statistical robustness when estimating the incidence of rural poverty.

A more relevant reference population for the measurement of poverty: In contrast with the previous methodology, which took as reference the poorest 25% of the population, the new methodology builds on the consumer profile of a population around the median of the national population. Therefore, the population was taken as a reference to define the nutritional intake of a population that was neither the poorest nor the richest, thus avoiding the use of patterns that do not meet minimum quality standards or nutritional balance. It also avoids the use of luxury items that do not represent the consumer patterns of the population in the calculation of calorie requirements. This new methodology chooses a population that is placed between the two extremes.

An exogenous urban Orshansky coefficient equal to the Latin American average: The urban and rural extreme poverty lines resulted from the application of the Ravallion methodology (with variations), to the 2006-2007 ENIG. For the calculation of urban poverty, the poverty line was calculated by multiplying the extreme poverty line by an exogenous Orshansky coefficient equal to the urban average for Latin America (2.4). For the calculation of rural poverty, an Orshansky coefficient was applied (1.74), which took into account the proportionality between the urban and rural domains in the 2006-2007 ENIG survey.

Elimination of the process of adjustment to National Accounts: Previous methodologies included the adjustment to National Accounts, a technique that aimed to correct household income underreporting in household surveys with factors derived from the National Accounts aggregates. MESEP decided to discontinue the adjustment to National Accounts because it considered that such adjustments to the different items of income introduced uncontrollable bias to the distribution of income and the urban-rural composition of poverty. MESEP considered that the assumptions underlying the adjustment to National Accounts would only be met in exceptional circumstances and that the correction of income aggregates did not effectively differentiate between households that actually underreported and those who did not, resulting in an artificial upward adjustment for the latter.

The use of a more refined aggregate income: The new measurement of poverty is based on an aggregate household income that includes only public cash transfers, but does not constitute an appropriate instrument to measure the long-term impact of cash transfer programs. The new methodology incorporates public cash transfers to the definition of income, taking into account

the national representativeness of GEIH, even if this survey is not representative of the set of specific beneficiaries of certain public programs. The best manner to measure the impact of these transfers on poverty in the short-term is by using poverty depth indicators. And the proper manner to measure the long-term effects of such programs on poverty is to turn to independent studies to measure their impact and effectiveness, conducting surveys designed for such specific purpose. It is important to note that the new methodology excludes non-monetary or in kind government transfers, due to the fact that the valuation by beneficiaries of the received services is problematic for the measurement of poverty.

The new methodology for measuring poverty adopted changes both in the definition of the poverty line and in the construction of aggregate household income. By 2010, the poverty rates were 37.2%, 33% and 50% for the national, urban and rural domains, respectively. For the same year, the extreme poverty rates were 12.3%, 8.1% and 25.5%, respectively. The national income inequality, measured by the Gini coefficient, was 0.56. For the 2002-2010 period, poverty declined 12.2 percentage points and extreme poverty declined 5.2 percentage points. However, inequality remained stable and is currently the highest in the Latin American context.

Finally, it is important to mention that the measurements of poverty based on the new methodology show similar trends to those of the previous methodology with two important differences, in terms of levels: a lower level of poverty and a similar but slightly lower level of extreme poverty.

9. Institutional recommendations

First, in order to guarantee continuous, accurate and timely monitoring of the living conditions of the population, MESEP proposed the creation of a permanent Technical Committee for the Measurement of Living Conditions, to assist DANE as the entity responsible for the calculation of official poverty figures. MESEP recognizes the importance of the joint work between the technical teams of DNP and DANE (MESEP) to strengthen their technical capacity in the production of monetary poverty indicators using an updated methodology and with statistical rigor. In particular, strengthening DANE with respect to its responsibility of providing comprehensive and timely access to official statistics.

In order to consolidate these institutional developments, MESEP proposes: a) DANE to be the responsible entity for the calculation of official monetary poverty and inequality figures, and b) the creation of a Technical Committee for the Measurement of Monetary Poverty. This permanent committee, coordinated by DANE, would be composed of three to five qualified persons, academically recognized in the social sciences and would also have one DNP representative. Members of the Technical Committee would be appointed for a fixed period and would participate especially with technical criteria. Experiences in other countries with similar institutions (such as Perú and México) have shown the benefits of this type of institutional arrangement for the credibility, transparency and timeliness of measurements of monetary poverty, income inequality and living standards. MESEP recommends that, as part of its agenda, this committee should complete the splicing of the measurements of poverty by using DANE's household surveys prior to 2002.

Second, MESEP recommends the publication of the poverty series with both methodologies for a period of two years in order to guarantee comprehensive and transparent information with respect to the measurements of poverty. However, MESEP believes that the new methodology provides a more accurate measurement of extreme poverty, poverty and income inequality in Colombia and therefore, recommends its use to guide public policies.

Third, MESEP recommends access to microdata of household surveys to guarantee transparency and replicability of the measurement of poverty. Moreover, MESEP supports DANE's decision of making available to the users the set of household surveys that MESEP used for the calculation of poverty. Thus, a rule of open access to statistical information is adopted enhancing the transparency, reliability and replicability of poverty estimates.

Bibliography⁹⁸

- Azevedo, Joao Pedro; Molina, Ezequiel; Newman, John; Rubiano, Eliana and Saavedra, Jaime (2009).** *How Has Poverty Evolved in Latin America and How is it Likely to be Affected by the Economic Crisis?* Latin American Poverty and Gender Group. The World Bank.
- Bidani, Benu and Ravallion, Martin (1994).** *How robust is a Poverty Profile?* World Bank Economic Review 8 (1): 75-102.
- Brandt, Loren and Holz, Carsten (2005).** *Spatial Price Differences in China: Estimates and Implications.* Microeconomics 0512001, EconWPA.
- Cortés, Darwin (2009).** *Análisis de los Gastos de los Hogares Colombianos 2006-2007. Informe Final. (Analysis of the Colombian Household Expenditure 2006-2007. Final Report).* Departamento Administrativo Nacional de Estadística (*National Administrative Department of Statistics*) (DANE), Universidad del Rosario (Colombia)
- Deaton, Angus and Tarozzi, Alessandro (2000).** *Prices and poverty in India.* Working Papers 213, Princeton University, Woodrow Wilson School of Public and International Affairs, Research Program in Development Studies.
- Departamento Administrativo Nacional de Estadística (National Administrative Department of Statistics) DANE (2009).** *Metodología Encuesta Nacional Ingresos y Gastos 2006-2007.* Colección Documentos No. 84. (*The 2006-2007 National Income and Expenditure Survey Methodology. Collection of documents No. 84.*)
- Economic Commission for Latin America and the Caribbean (ECLAC) (1991).** *Magnitud de la pobreza en América Latina en los años Ochenta.* Estudios e Informes de la CEPAL (*Magnitude of poverty in Latin America in the Eighties. Studies and reports of ECLAC*). United Nations. Santiago de Chile.
- FAO, WHO and UNU (2001).** *Human energy requirements: Report of a Joint FAO/WHO/UNU Expert Consultation.* Roma 17-24 October.
- Feres, Juan Carlos (2010).** *Medición de la pobreza a partir del ingreso: Avances y desafíos.* Presentación para el Seminario Internacional “Colombia en las Nuevas Tendencias de Medición de la Pobreza y la Igualdad de Oportunidades” (*Measuring poverty from income: Progress and Challenges. Presentation for the International Seminar: “Colombia in the New Measurement Trends of Poverty and Equal Opportunities”*). Bogotá, July 28-29, 2010.
- Feres, Juan Carlos and Mancero, Xavier (2001).** *Enfoques para la medición de la pobreza. Breve revisión de literatura. (Approaches to measuring poverty. Brief literature review).* ECLAC.
- Foster, James and Sen, Amartya (2001).** *La Desigualdad Económica. (Economic Inequality)* Fondo de Cultura Económica (*Economic Culture Fund*). México. Translation by Eduardo Suárez Galindo. Originally published as: *On Economic Inequality.* Oxford University Press. 1997.

⁹⁸ The translation of the Bibliography is for reference purposes only.

Greer, Joel and Thorbecke, Erik (1986). A Methodology for Measuring Food Poverty Applied to Kenya. In: *Journal of Development Economics* n. 24. p. 59-74.

Grosh, M, & Glewwe, P. (s. f.). *Designing Household Survey for Developing Countries. Vol. 1. The World Bank.*

Haughton, Jonathan y Khandker, Shahidur (2009). *Handbook on Poverty and Inequality.* The World Bank. Washington, D.C.

Herrera, Javier (2001). *Nuevas estimaciones de pobreza en el Perú 1997-2000* (New estimates of poverty in Perú). Instituto Nacional de Estadística e Informática (INEI) - (National Institute of Statistics and Information Technology) Institut de recherche pour le développement (Institute of research for development) (IDR). Lima, Perú. 1997-2000.

_____ **(2009).** *Medición de pobreza en Colombia: comentarios a la luz de debates recientes.* (Measuring poverty in Colombia: comments in light of recent debates). Presentación realizada en Taller sobre medición de pobreza (Presentation held at Workshop on poverty measure) DANE. November 9-19, 2009.

_____ **(2010).** *Medición de la Pobreza Monetaria: Desarrollos Recientes.* Presentación para el Seminario Internacional "Colombia en las Nuevas Tendencias de Medición de la Pobreza y la Igualdad de Oportunidades" (Monetary Poverty Measurement: Recent Developments. Presentation for the International Seminar "Colombia in the New Measurement Trends of Poverty and Equal Opportunities"). Bogotá, July 28-29, 2010.

Leyva-Parra, Gerardo (2004). *El ajuste del ingreso de la ENIGH con la contabilidad nacional y la medición de la pobreza en México.* (ENIGH's income adjustment with national accounting and poverty measurement in México) Secretaría de Desarrollo Social (Secretariat of Social Development) (SEDESOL) Documentos de investigación. (Research papers). November.

Misión para el Empalme de las Series de Empleo, Pobreza y Desigualdad. (Mission for the splicing of employment, poverty and inequality series). (MESEP) (2009). Resultados Fase I, Resumen Ejecutivo (1st phase results, Executive Summary). DNP-DANE.

Muñoz, Manuel and Rivas, Guillermo (2006). *Construcción de las Canastas Normativas de Alimentos para Trece Ciudades, Resto Urbano y Zona Rural.* (Construction of normative food baskets for thirteen cities, remaining urban areas and the rural area). Departamento Nacional de Planeación (National Planning Department) (DNP) – Misión para el Diseño de una Estrategia para la Reducción de la Pobreza y la Desigualdad- (Mission for the design of a strategy for reducing poverty and inequality) MERPD.

Orshansky, Mollie (1965). *Who's Who Among the Poor: A Demographic View of Poverty.* Social Security Bulletin, July, pp.3-33

_____ **(1966).** *Recounting the Poor – A Five Year Review.* Social Security Bulletin, April, pp.2-19

Osmani, Siddiqur (1982). *Economic Inequality and Group Welfare.* Oxford University Press. Oxford.

Pradhan, Menno; Suryahadi, Asep; Sumarto, Sudarno; Pritchett, Lant (2001). *Eating Like Which "Joneses"? An Iterative Solution to the Choice of a Poverty Line Reference Group.* Review of Income and Wealth, Series 47, Number 4, December 2001.

Ravallion, Martin (1994). *Poverty Comparisons*. Harwood Academic Press, Chur, Switzerland.

_____ (1998). *Poverty Lines in Theory and Practice*. Living Standards Measurement Study Working Paper No. 133. The World Bank. Washington, D.C.

_____ (2010). *Poverty Lines Across the World*. Policy Research Working Paper 5284. The World Bank. Washington D.C. April 2010.

Ravallion, Martin and Lokshin, Michael (2006). *Testing Poverty Lines*. Review of Income and Wealth, Series 52 No. 3, pp.399-421.

Romero, Julio (2005). *¿Cuánto cuesta vivir en las principales ciudades colombianas? Índice de Costo de Vida Comparativo. (How much does it cost to live in the major Colombian cities? Comparative Cost of Living Index)*. Documentos de Trabajo sobre Economía Regional No. 57. Centro de Estudios Económicos y Regionales (CEER). Banco de la República. (*Working Papers on regional economy No. 57. Center for Economic and Regional Studies (CEER). Central Bank*). Cartagena.

Rowntree, Benjamin (1901). *Poverty: A Study of Town Life*, MacMillan, London.

_____ (1936). *Poverty and Progress: A Second Social Survey of York*. London: Longmans, Green.

_____ (1951). *Poverty and the Welfare State: A Third Social Survey of York*. London.

Satya, Paul (1989). *A Model of Constructing the Poverty Line*. In: *Journal of Development Economics* n. 30. p. 129-144.

Sen, Amartya (1979). Issues in the Measurement of Poverty. In: *Scandinavian Journal of Economics*. Vol. 81.

_____ (1981). *Poverty and Famines*. Oxford University Press.

Shorrocks, Anthony and Kolenikov, Stanislav (2003). *A Decomposition Analysis of Regional Poverty in Russia*. Working Papers UNU-WIDER Research Paper, World Institute for Development Economic Research (UNU-WIDER).

Annexes

Annex A. Variables included in the *Extreme values* and *False Zeros* models

VARIABLES IN THE EXTREME VALUES MODEL (1)	INCOME SOURCES (2)															
	INCOME FROM WORK					INCOME FROM OTHER SOURCES										
	SALARIES		AND			EMPLOYED					UNEMPLOYED AND INACTIVE					
	MIFA	IK	ISA	MIUI	IOS1	IOS2	IOS3	IOS4	IOS5	IOS6	IOS1	IOS2	IOS3	IOS4	IOS5	IOS6
AGE	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
AGE2	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
HOURS NORMALLY WORKED PER WEEK IN THE FIRST ACTIVITY	X	X			X	X	X	X	X	X						
HOURS NORMALLY WORKED PER WEEK IN THE SECOND ACTIVITY			X													
APPROVED YEARS OF EDUCATION	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
DUMMY OTHER TOWNSHIPS	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
DUMMY BOGOTA	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
DUMMY : MEDELLIN, CALI AND BARRANQUILLA	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
DUMMY WOMEN	X	X	X	X	X	X	X	X	X	X						
DUMMY BLUE COLLAR WORKERS	X	X	X		X	X	X	X	X	X						
DUMMY DOMESTIC WORKERS	X	X	X		X	X	X	X	X	X						
DUMMY UNEMPLOYED				X							X	X	X	X	X	X
DUMMY STUDENTS				X							X	X	X	X	X	X
DUMMY OTHER ACTIVITY OF WAP				X							X	X	X	X	X	X
DUMMY HOME MAKER				X							X	X	X	X	X	X
DUMMY SELF-EMPLOYED	X		X		X	X	X	X	X	X						
DUMMY EMPLOYERS	X		X		X	X	X	X	X	X						
CURRENT JOB DURATION (MONTHS)	X	X	X		X	X	X	X	X	X						
DUMMY FOR HEAD OF HOUSEHOLD	X	X	X		X	X	X	X	X	X	X	X	X	X	X	X
No. OF PERSONS IN THE HOUSEHOLD OF THE PERSON OF WAP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
No. OF SALARY EARNERS IN THE HOUSEHOLD OF THE PERSON OF WAP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
No. OF SELF-EMPLOYED PERSONS IN THE HOUSEHOLD OF THE PERSON OF WAP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
No. OF UNEMPLOYED PERSONS IN THE HOUSEHOLD OF THE PERSON OF WAP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
No. OF PERSONS UNDER 5 YEARS OLD IN THE HOUSEHOLD OF THE PERSON OF WAP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
No. OF ADOLESCENTS IN THE HOUSEHOLD OF THE PERSON OF WAP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
No. OF SENIOR CITIZENS IN THE HOUSEHOLD OF THE PERSON OF WAP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
No. PERSONS WITHOUT EDUCATION IN THE HOUSEHOLD OF THE PERSON OF WAP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
No. OF PERSONS WITH HIGHER EDUCATION IN THE HOUSEHOLD OF THE PERSON OF WAP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
AVERAGE YEARS IN THE HOUSEHOLD OF THE PERSON OF WAP	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
No. OF PERSONS AFFILIATED TO THE SSS IN THE HOUSEHOLD OF THE PERSON OF WAP	X															

(1) THESE VARIABLES WERE THE SAME FOR DETECTING FALSE ZEROS

(2) MIFA: MONETARY INCOME OF THE FIRST ACTIVITY

ISA: INCOME OF THE SECOND ACTIVITY

IK: INCOME IN KIND

MIUI: MONEY INCOME OF THE UNEMPLOYED AND INACTIVE

IOS1: INCOME FROM LEASES

IOS2: INCOME FROM INTERESTS

IOS3: INCOME FROM PENSIONS

IOS4: INCOME FROM BENEFITS

IOS5: INCOME FROM SEVERANCE PAY INTERESTS

IOS6: INCOME FROM OCCASIONAL PROFITS

Annex B. Criteria for the definition of cells for the Hot Deck imputation model, by income source

VARIABLES IN THE HOT DECK IMPUTATION MODEL	INCOME SOURCES (1)															
	INCOME FROM WORK				INCOME FROM OTHER SOURCES											
	SALARIES AND PROFITS		AND		EMPLOYED						UNEMPLOYED AND INACTIVE					
	MIFA	HK	ISA	MIUI	IOS1	IOS2	IOS3	IOS4	IOS5	IOS6	IOS1	IOS2	IOS3	IOS4	IOS5	IOS6
DOMAIN OF STUDY 1= 13 AREAS 2 = OTHER TOWNSHIPS 3=REMAINING RURAL STRATUM (1-6)	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
AGE (GROUPS) (2)	X															
YEARS OF EDUCATION (GROUPS) (3)	X		X		X	X	X	X	X	X						
RECODED OCCUPATIONAL POSITION (4)	X	X			X	X	X	X	X	X						
SEX 1=MEN 2=WOMEN	X	X														
HEAD OF HOUSEHOLD 1= HEAD OF HOUSEHOLD 0 =NOT A HEAD OF HOUSEHOLD	X															
WEEKLY HOURS WORKED (5)			X		X	X	X	X	X	X						
ACTIVITIES OF THE UNEMPLOYED AND INACTIVE (6)				X							X	X	X	X	X	X

(1) MIFA: MONETARY INCOME OF THE FIRST ACTIVITY

ISA: INCOME OF THE SECOND ACTIVITY

HK: INCOME IN KIND

MIUI: MONEY INCOME OF THE UNEMPLOYED AND INACTIVE

IOS1: INCOME FROM

IOS2: INCOME FROM

IOS3: INCOME FROM PENSIONS

IOS4: INCOME FROM AID

IOS5: INCOME FROM SEVERANCE PAY INTERESTS

IOS6: INCOME FROM OCCASIONAL PROFITS

(2) 1=UNDER 18 YEARS OLD

2=FROM 18 TO 24 YEARS OLD

3=FROM 25 TO 45 YEARS OLD

4=46 YEARS OLD AND OLDER

(5) 1=LESS THAN 25 HOURS

2=25 HOURS OR MORE

(3) 1=UNTIL

2=SECONDARY AND MIDDLE SCHOOL

3=HIGHER EDUCATION

ELEMENTARY

(6) 1=UNEMPLOYED

2=STUDENTS

3=HOMEMAKER

4=OTHER ACTIVITY

(4) 1=UNPAID FAMILY WORKERS

2=BLUE COLLAR WORKERS

3=EMPLOYEES

4=DOMESTIC WORKERS

5=SELF-EMPLOYED WORKERS

6=EMPLOYERS

Annex C. Multinomial logistic model to predict housing tenure

The approach of this model assumes the existence of $k + 1$ categories (whose order has no explicit meaning) in y . The probabilities P_1, P_2, \dots, P_{k+1} are associated with these $K + 1$ categories. The form of this model is:

$$\log\left(\frac{\Pr\{Y = i/x\}}{\Pr\{Y = k + 1/x\}}\right) = \alpha_i + x_i' \beta, \quad i = 1, 2, \dots, k$$

Where every $\alpha_1, \alpha_2, \dots, \alpha_k$ is the intercept of the equation assigned to each category and β is the "pending" parameter vector.

For a "given" vector of explanatory variables x_i , a linear predictor of the form $\varphi_i = \alpha_i + x_i' \beta$ is defined.

The probability P_{ij} that the j th individual makes the choice i is given by:

$$P_{ij} = \frac{e^{\alpha_i + \beta' x_{ij}}}{1 + \sum_{t=1}^k e^{\alpha_t + \beta' x_{ij}}} \quad i = k + 1$$

The expression above indicates that the parameters of the linear predictor α_i, β , are estimated first and then, $P_{k+1,j}$ or the probability of the "base" category is estimated (i.e., when $i = k + 1$) by using the previous estimated parameters of α_i, β , and data observed in x_i .

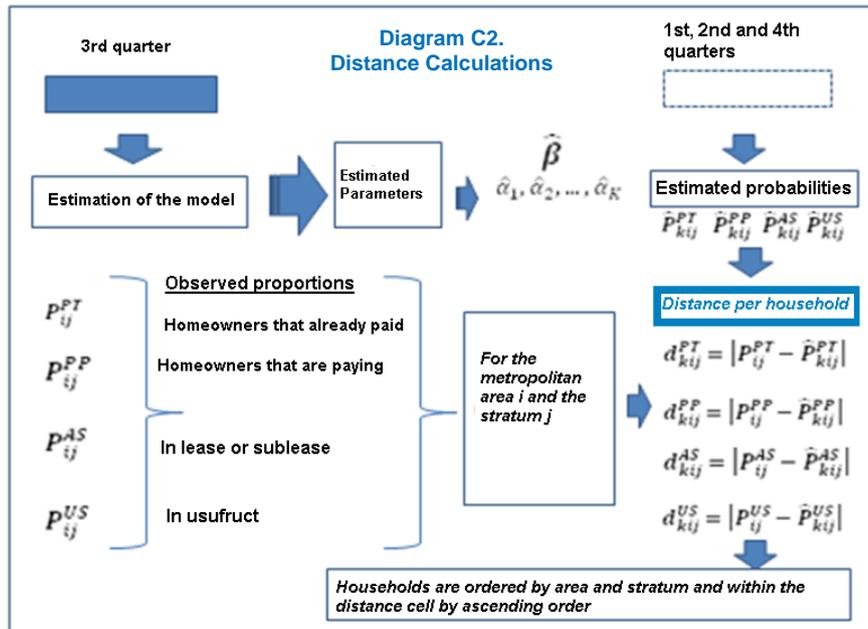
With this value ($P_{k+1,j}$) and the estimated parameters α_i, β , other tenure categories that are not the base category ($1 \leq i \leq k$), but that depend upon its value, where $k + 1$ is the base category, are then estimated. In the case of ECH for the 2002-2005 period, the categories are:

1. Owned, fully paid
2. Owned, still being paid
3. In lease or sublease
4. In usufruct

The chosen base category was "home owners with fully paid homes."

The model included variables to the right of the equation, i.e. x_i , that can be divided into groups as follows:

- a. Personal characteristics of the household head (age, dummy variable for sex, age squared, dummy variable for civil status).
- b. Educational characteristics of the household head (years of education)



Source: MESEP Technical Team

The distance criterion was considered to choose the event that would be assigned to each household, by using each of the four probabilities that were estimated for each one of them⁹⁹. This criterion enabled the comparison between the *observed proportions* for each category in the 3rd quarter and the *estimated probabilities* in each household of the other quarters. In order to determine the categories to be assigned to households, these were ordered according to their distance in ascending order and the first ones completing the proportion observed in the 3rd quarter in that category were selected. Households were chosen by Metropolitan Area and stratum (Diagram C2).

On the other hand, the order for assigning categories remained the same as the one appearing on the ECH form. That is, first, homeowners that already paid their homes were chosen, second, those who were still paying, third, those who were paying rent, fourth, those with homes in

99

\hat{p}_{kij}^{PT} : Estimated probability that the household k is a homeowner who fully paid a home in the metropolitan area i of the stratum j

\hat{p}_{kij}^{PP} : Estimated probability that the household k is a homeowner who is still paying a home in the metropolitan area i of the stratum j

\hat{p}_{kij}^{AS} : Estimated probability that the household k is in lease or sublease in the metropolitan area i of the stratum j

\hat{p}_{kij}^{US} : Estimated probability that the household k is in usufruct in the metropolitan area i of the stratum j

usufruct and lastly the category *others* was assigned. The number of selected households was controlled with the observed proportion of households in the third quarter.

Annex D. Review of ECLAC's methodology to construct poverty and extreme poverty baskets

ECLAC is currently reviewing the methodology for the construction of poverty and extreme poverty lines. Among the reasons for conducting this review are (cf. Feres, 2010, page 3):

- *Obsolescence of existing lines:* Estimates being calculated by ECLAC were made from poverty and extreme poverty lines that were built more than 20 years ago.
- *Availability of new sources of information:* There are new applications of IES in different Latin American countries.
- *Improvement of the methodology:* Recent improvements for the identification and characterization of poverty should be incorporated into the methodology that typically has been applied.

In particular, the update of the methodology implemented by ECLAC considers:

- *New estimates of calorie requirements:* One of the elements involved in the definition of the extreme poverty line are the calorie requirements. At this respect, ECLAC is considering the inclusion of estimates of requirements and to replace the use of the average requirements by the one corresponding to each household (cf. Ibid, p. 17).
- *Modifying the reference stratum selection:* The selection of the reference population is crucial since the cost per kilocalorie is established from it. The traditional method for selecting the reference stratum consists in identifying the percentile of the distribution of per capita income that meets the calorie requirement. In this regard, ECLAC considers that "calorie intake is not necessarily a good indicator for stratum selection". Two reasons justify this consideration: i) "The increasing availability of high-calorie and low cost products indicates that this indicator is less representative of a certain level of welfare than before," and ii) "Calorie intake is not measured directly, but is inferred from information on expenditure and acquired quantities collected by surveys. The accuracy of that measurement may not be sufficient to generate adequate results". "In this respect, ECLAC's proposal consists of defining exogenous reference strata according to the relative position of each country within the regional context, taking into account variables such as per capita GDP, life expectancy, infant mortality, maternal mortality, illiteracy of the population aged 15 or older, access to clean water and sanitation (cf. Ibid, p. 8).

- *Changing the Orshansky coefficient:* ECLAC has kept the same value of the Orshansky coefficient (calculated on the IESs that the countries of the region implemented in the eighties) for its estimates. The update of the information on this coefficient, probably with a value that is higher than the one currently in use is foreseen. As is the case today, the value of this coefficient is the same for all Latin American countries for which ECLAC performs calculations.
- *Including factors of economies of scale:* In order to establish the cost of the basic non-food basket, ECLAC is considering the implementation of a factor reflecting economies of scale.
- *Improvements to the correction for income underreporting:* As part of the construction of the *income* variable, ECLAC adjusted its underreporting when using the Household Account of the National Accounts as a reference framework. The procedure to correct the underreporting is being assessed and even abandoning the practice of adjusting to National Accounts is being considered. An additional aspect that is being considered for modification with respect to the construction of the *income* variable is the non-inclusion of imputed rent as a part of it, and use differentiated lines instead.

Annex E. Construction of the *per capita expenditure* variable in ENIG for the 2006-2007 period

This appendix presents details of the changes made to reach the final version of the variable known as *total current expenditure* of the spending unit, which was used in the definition of the lines.

1. Initially included items to form current expenditure were the same as those used to publish the results of ENIG for the 2006-2007 period¹⁰⁰:
 - *Expenses forms*: Out of pocket expenses in food, transport, communications and recreation, registered in Forms 2 and 4 of ENIG for the 2006-2007 period; less frequent expenses registered on Form 3 except for mortgage amortization payments for dwellings that are for household use but different from the dwelling inhabited; home purchase, home expansion or subdivision; purchase of land, lots, premises, farms, plots and the like; payment pertaining to taxes of homes that are occupied by households, the purchase of new and used vehicles for private use; registration fees for new vehicle and used vehicle handover; certificate of gases and other reviews mandated by law; payment of SOAT; vehicle taxes; monthly installments for acquisition of real estate other than housing; loan installments for vehicle acquisition; income tax and other taxes.
 - *Income and characterization form*: Monthly lease payment and payment for food received by children of the spending unit that are under three years of age (in the community center, day care or preschool) and persons aged three years old or older who are studying.
2. From this first definition of expenditure, new categories were included: *less frequent expenditures with annual periodicity*, registered on Form 3 and that consider three criteria: i) durability of the good, ii) value of the good and iii) purchase frequency.

The following expenses were included: Furniture repair and maintenance, flooring accessories and materials; purchase of books; imputed rent, materials and workforce for housing maintenance; and the reconnection of various public utilities (water, gas, telephone). Goods and services such as expenditure on transportation, recreation and

¹⁰⁰ Purchase on credit of goods and services and the acquisition ways different from purchasing, for example, gifts, exchanging or bartering are excluded.

culture, which were included in the definition of non-current expenditure used in the ENIG survey for the 1994-1995 period were also included.

The following items were excluded from current expenditure considering the same criteria above (durability, value and frequency of purchase): furniture expenses; purchase of vehicles, bicycles, consumer durables; home expansion (investment), electrical and hand fans; camping equipment and accessories; encyclopedias; works and collections; and vehicle radios.

3. The third step for obtaining the definite *current expenditure* consisted in reviewing the deductions of the occupied persons (employees and self-employed) in its different variables (health, pension, ICA, VAT, withholding of tax at source, official diary, pension solidarity discount). Thus, current expenditure decreased from \$ 13.3 billion (if it includes social security) to \$ 12.8 billion (if it excludes social security). The structure of *current expenditure* with and without social security health resulted as follows (Table E1):

CURRENT EXPENDITURE EXCLUDING SGSSS*PAYMENTS		CURRENT EXPENDITURE INCLUDING SGSSS*PAYMENTS	
Expenditure structure	NATIONAL TOTAL	Expenditure structure	NATIONAL TOTAL
Food and non-alcoholic beverages. Includes Chapter C of ENIG 06/07** and meals outside the home	29,7	Food and non-alcoholic beverages. Includes Chapter C of ENIG 06/07** and meals outside the home	28,5
Alcoholic beverages and tobacco	1,5	Alcoholic beverages and tobacco	1,4
Clothing and footwear	5,1	Clothing and footwear	4,9
Housing, water, electricity, gas and other fuels	25,8	Housing, water, electricity, gas and other fuels	24,8
Furniture and household goods	3,2	Furniture and household goods	3,1
Health	2,1	Health	2,0
Transportation	8,8	Transportation	8,4

Communications	3,5	Communications	3,4
Recreation and culture	2,8	Recreation and culture	2,7
Education	3,5	Education	3,4
Restaurants and hotels	1,6	Restaurants and hotels	1,6
Miscellaneous goods and services (excluding SGSSS payments or other payments)	12,5	Miscellaneous goods and services (including SGSSS payments)	15,8
Source: DANE - ENIG for the 2006-2007 period.			
*SGSSS: General System of Social Security in Health			
** Chapter of the survey that collected information on expenditure in 25 food groups that were acquired with a periodicity greater than a week			

The expenditure group “*miscellaneous goods and services*” showed the greatest change, increasing from 12.5% to 15.8%. The other groups do not show greater variation since they exclude social security payments.

Purchase on credit for some goods that were registered on Form 3 was excluded while the value of the credit card payment was included in order to complement the concept of *current expenditure*.

Annex F. Food items for constructing the Spatial Price Deflator (SPD)

Table F1. Food basket for constructing the Spatial Price Deflator (SPD)

Code	Name
1110101	Rice for main dish
1110102	Rice for soup
1110103	Brown Rice
1110201	Pasta or noodles for soup: letters, shells, etc.
1110202	Pasta for main dish: spaghetti, ravioli, macaroni, noodles, etc.
1110301	Oatmeal
1110302	Oat flakes
1110401	Pearl barley
1110402	Barley cuchuco
1110501	Wheat cuchuco
1110502	Wheat flour
1110503	Wheat bran
1110601	Maize: white, yellow, threshed, <i>peto</i> , <i>porva</i> , <i>pira</i> and <i>millo</i>
1110602	Maize cuchuco
1110603	Maize flour
1110604	Pre-cooked cornmeal
1110605	Cornstarch
1110701	Flour for cakes, natilla, buñuelos, etc.
1110702	Flour for cornstarch soup
1110704	Cereals: Corn Flakes, Rice Krispies, Frosted Flakes, Snow Flakes, Fruit Loops, Trix
1110705	Baby cereal
1110706	Sesame, soybean
1110707	Doughs for the preparation of bakery or pastry products (of wheat, corn, soybean)
1110708	Raw or pre-cooked arepas
1110709	Pre-cooked wrapped corn cakes (<i>bollos</i> and <i>envueltas</i>), quiches, tamales
1110710	Wheat germ, lecithin, brewer's yeast (naturist food)
1110711	Precooked meals prepared with a basis of cereals (lasagna, pizza, empanadas)
1110712	<i>Carve</i> (soy-based protein)
1110801	Bread: common, with cheese, seasoned, mogolla, stuffed, with coconut, sliced, pita; rusk (<i>calado</i>).
1110802	Baguette

1110803	Whole wheat bread
1110901	Salty cookies
1110902	Sweet cookies (from bakeries and packed)
1110903	Whole wheat cookies (from bakeries and packed)
1110904	Other bakery products not previously mentioned: brownies, wafers, cakes, biscuits (<i>puñines</i> in the Atlantic Coast)
1120101	Boneless beef
1120102	Beef with bone
1120103	Beef bone
1120104	Ground beef
1120105	Beef tongue
1120106	Liver, kidney, heart, brain, guts and other beef viscera
1120107	Beef legs, tail and head
1120201	Boneless pork
1120202	Pork with bone
1120204	Pork bone (spine and other bones)
1120205	Bacon and hoof
1120206	Liver, kidney, heart, brain, guts and other pork viscera
1120301	Live chicken or hen
1120302	Chicken or hen meat (whole or in pieces)
1120303	Poultry: turkey, squab, <i>zuro</i> , pigeon, duck, goose, quail meat
1120304	Chicken or hen offals
1120402	Lamb, goat, kid
1120403	Capybara
1120502	Mortadella
1120503	Sausages
1120504	Ham and canned ham
1120505	Chorizo
1120508	Other sausages (<i>longaniza</i> , <i>butifarra</i> , <i>genovas</i> , <i>salchichón</i> , salami, <i>cabanos</i> , pâtes, etc)
1120509	Canned processed meats
1130100	Freshwater fish, fresh or frozen
1130201	Saltwater fish, fresh or frozen
1130202	Shrimp, fresh or frozen
1130203	Canned fish and crustaceans (sardines, tuna, salmon, mussels, etc.)

1130204	Other seafood, crustaceans and mollusks, fresh or frozen
1130205	Other fish products and derivatives
1140101	Raw milk
1140102	Pasteurized Milk (whole, skimmed, Semi-skimmed, lactose-free, Infant formula and fortified)
1140103	Long-life milk(whole, skimmed, Semi-skimmed, lactose-free, Infant formula and fortified)
1140104	Powdered milk for infant use
1140105	Powdered milk (whole, skimmed, Semi-skimmed, infant formula)
1140201	Cottage cheese
1140202	Double Cream cheese
1140203	Other cheese: <i>Pera</i> , <i>Mozarella</i> , <i>Paipa</i> , <i>Parmesan</i> , etc.
1140204	Curd, <i>Quesillo</i>
1140303	Yoghurt
1140304	Koumiss
1140305	Costeño sour cream
1140501	Chicken and other bird eggs
1140502	Iguana and tortoise eggs
1150101	Sunflower-seed, maize, soya, palm oil
1150102	Olive oil
1150103	Other edible oils: Canola, ground-nut, coconut, almond, Diet, Etc.
1150200	Butter derived from cow milk
1150300	Margarines
1150401	Shortening
1150402	Pig lard
1150403	Other fats and oils (fish oil and oil from other marine animals, peanut butter)
1160101	Oranges
1160102	Lemons
1160103	Mandarines
1160107	Bananas
1160108	Apples
1160109	Pears
1160110	Melons
1160111	Mangos
1160112	Guavas
1160113	Soursop
1160114	Sapotes
1160116	Tamarillos
1160117	Blackberries
1160118	Banana passionfruits
1160119	Passion fruits
1160120	Lulos
1160121	Watermelons
1160122	Granadillas
1160123	Pineapples
1160124	Papayas

1160125	Mountain papaya
1160126	Grapes (Green and red)
1220700	Concentrates for preparing beverages
1160127	Strawberries
1160128	Coconuts
1160131	Peaches
1160132	Peach-palms
1160133	Plums
1160134	Avocados
1160140	Other fresh fruits
1160200	Frozen fruit or fruit pulp
1160301	Almonds and other nuts (peanut, ground-nuts, Pistachio nuts, etc.)
1170101	Tomatoes
1170102	Bulb onion
1170103	Scallions or green onions
1170104	Chives
1170105	Carrots
1170106	Beets
1170107	Cucumbers
1170108	Bell pepper
1170109	Caigua (cucumber for stewing or stuffing)
1170110	Turnips
1170112	Garlic
1170113	Chillies
1170114	Lettuce
1170115	Spinach
1170116	Chard
1170117	Cabbage, Brussels sprouts
1170119	Cauliflower
1170120	Broccoli
1170121	Coles (Stalks)
1170122	Celery
1170125	Pumpkin
1170126	Winter squash (acorn squash)
1170127	Summer squash(zucchini)
1170128	Green beans in pods
1170129	Shelled green beans
1170130	Green peas in pods
1170131	Shelled green peas
1170132	Fresh Lima beans in pods
1170133	Shelled Fresh Lima beans
1170134	Runner Beans
1170135	Corn and sweet corn
1170136	Eggplant
1170137	Mushrooms and other edible fungi

1170138	Other vegetables and legumes, fresh: watercress, <i>chachafruto</i> , edible seaweed, etc.
1170139	Fresh vegetables, tied (<i>macitos</i> in the Atlantic Coast)
1170140	Chopped vegetables in tray or bag (green mix, <i>macitos</i>)
1170201	Beans, dried: common bean, <i>cargamanto</i> , <i>guarzo</i> , pigeon pea, lima bean, white bean, etc.
1170202	Peas, dried
1170203	Broad bean, dried
1170204	Lentils
1170205	Chickpeas
1170301	Canned vegetables and legumes: peas, beans, corn, asparagus, carrots, green beans, salads
1170302	Olives, capers, pickled cucumbers and other pickles in jar
1170303	Dried legumes and vegetables
1170401	Plantain, green
1170402	Plantain, ripe
1170403	<i>Colicero</i> , <i>guineo</i> , <i>mataburro</i> (different types of banana/plantain)
1170501	Common potatoes: <i>pastusa</i> , <i>tocarreña</i> , <i>R12</i>
1170502	Phureja potatoes/baby dutch yellow potatoes (<i>papa criolla</i>)
1170600	Cassava
1170700	<i>Arracacha</i> (Andean tuber)
1170800	Yam
1170900	<i>Ullucos</i> , <i>cubios</i> , <i>hibias</i> , <i>chuguas</i> , <i>rubas</i> (tubers of the Andean region)
1171000	Other roots and tubers: <i>majafa</i> , <i>achira</i> , <i>batata</i> , etc.
1171101	Potatoes, cassava and plantain, pre-cooked
1171102	Plantain, cassava and canna indica (<i>achira</i>) flours; mashed potatoes
1180101	Refined sugar
1180102	Raw or brown sugar
1180200	<i>Panela</i>
1180301	Jellies and marmalades
1180302	Jams
1180303	Guava jelly
1180307	Gelatins, flans and puddings in cups ready for consumption
1180308	Honey, molasses and syrups
1180400	Chocolate butter cream and chocolate icing
1180500	Candies, confectionery, caramels, bonbons, chocolates, bubble-gum, marshmallows
1180600	Ice cream, cones, frozen bars, ice cream cups, frozen drinks and ice cream cakes
1180700	Gelatins, flans and puddings in powder for food preparations
1190101	Ketchup
1190102	Mayonnaise
1190103	Mustard, Pink sauce (mayo-ketchup dipping sauce)
1190105	Other sauces: soy sauce, Chinese sauce, tartar sauce, chili sauce, hot pepper sauce, etc.
1190106	Vinegar
1190200	Salt
1190301	Cumins

1190302	Pepper
1190305	Food coloring
1190306	Garlic in powder or paste, cinnamon, cloves, nutmeg
1190400	Culinary herbs: parsley, cilantro, rosemary, oregano, etc.
1190501	Soups: of vegetables, meat, chicken, fish, instant soups, soups for babies
1190502	Broth and stock concentrates (bouillon powder and cubes)
1190503	Seasoning mix for preparing meats and other food
1190600	Packaged fried snacks: potatoe chips, Jack's Snacks <i>Chitos</i> ™, Ramo <i>Maizitos</i> ™, plantain chips, <i>Besitos</i> ™ (<i>round cheese snacks</i>), etc.
1190800	Precooked foods: canned <i>lechona</i> and <i>ajiaco</i> , and the like
1210101	Ground coffee
1210102	Instant coffee in powder or granules
1210103	Coffee beans
1210104	Other beverages prepared with instant coffee: cappuccino, mochaccino, latté
1210201	Chocolate tablets (sweetened and unsweetened)
1210202	Instant chocolate in powder or granules
1210203	Cocoa and <i>chucula</i> (Colombian chocolate beverage)
1210204	Powder to prepare chocolate drinks
1210205	Chocolate drinks, ready for consumption
1210301	Tea
1210302	Aromatic herbs (dried and fresh), mate
1220100	Mineral water (personal bottles)
1220200	Canned water or water from tanker
1220300	Ice
1220400	Soft drinks
1220500	Maltas (soft drink)
1220600	Packaged beverages (juices and sports drinks)

Source: MESEP technical team

Annex G. Composition of the Urban and Rural Basic Baskets

Table G1. Food items that compose the urban basic basket

Food items	Unit value median	Calories per-capita	Quantities per-capita
Rice for main dish	1.6	291.4	83.0
Pasta or noodles for soup: letters, shells, etc.	3.2	6.1	1.7
Pasta for main dish: spaghetti, ravioli, macaroni, noodles, etc.	3.2	13.0	3.6
Raw or pre-cooked arepas	2.1	16.7	9.6
Bread: common, with cheese, seasoned, <i>mogolla</i> , stuffed, with coconut, sliced, pita; rusk (<i>calado</i>),.	3.7	67.1	19.3
Boneless beef	9.5	29.3	19.5
Beef with bone	7.0	8.4	3.3
Boneless pork	8.6	8.0	2.8
Chicken or hen meat (whole or in pieces)	6.0	34.3	19.2
Freshwater fish, fresh or frozen	6.0	2.2	4.6
Pasteurized milk (whole, skimmed, semi-skimmed, lactose-free, infant formula and fortified)	1.5	20.6	41.3
Long-life milk (whole, skimmed, semi-skimmed, lactose-free, infant formula and fortified)	1.6	13.7	27.5
Cottage cheese	7.1	3.4	4.0
Chicken and other bird eggs	3.3	28.6	20.2
Sunflower-seed, maize, soya, palm oil	4.0	164.5	18.6
Butter derived from cow milk	8.0	4.9	0.6
Margarines	6.4	4.0	0.5
Tomatoes	1.9	1.9	11.2
Bulb onion	2.0	2.3	6.6
Beans, dried: common bean, <i>cargamanto</i> , <i>guarzo</i> , pigeon pea, lima bean, white bean, etc.	3.6	33.8	9.9
Peas, dried	1.6	9.9	2.9
Lentils	2.0	28.9	8.4
Plantain, green	1.2	19.4	19.4
Plantain, ripe	1.2	8.5	9.2
Common potatoes: <i>pastusa</i> , <i>tocarreña</i> , <i>R12</i>	0.9	51.8	56.0
Cassava	1.0	19.4	15.9
Refined sugar	2.0	46.4	11.7
Raw or brown sugar	2.5	36.3	9.6
<i>Panela</i>	1.9	72.8	19.7
Salt	0.7	-	6.7
Cumins	16.7	-	0.1
Food coloring	10.0	0.2	0.2
Culinary herbs: parsley, cilantro, rosemary, oregano, etc.	6.1	0.2	0.4
Soups: of vegetables, meat, chicken, fish, instant soups, soups for babies	15.0	1.4	0.4
Broth and stock concentrates (bouillon powder and cubes)	17.7	0.6	0.3
Ground coffee	14.0	1.7	1.5
Instant coffee in powder or granules	40.0	0.8	0.2
Chocolate tablets (sweetened and unsweetened)	9.0	7.1	2.1
Instant chocolate in powder or granules	11.9	1.3	0.3
Total			

Source: MESEP technical team

Table G2. Food items that make up the rural basic basket

Food items	Unit value median	Calories per-capita	Quantities per-capita
Rice for main dish	1.6	574.7	163.7
Pasta or noodles for soup: letters, shells, etc.	3.2	28.3	8.0
Pasta for main dish: spaghetti, ravioli, macaroni, noodles, etc.	3.2	30.4	8.4
Oatmeal	2.4	11.5	3.0
Oat flakes	2.6	8.9	2.6
Wheat flour	1.6	35.4	10.1
Maize: white, yellow, threshed, <i>peto</i> , <i>porva</i> , <i>pira</i> and <i>millo</i> .	2.0	29.1	9.0
Bread: common, with cheese, seasoned, <i>mogolla</i> , stuffed, with coconut, sliced, pita; rusk (<i>calado</i>).	3.7	65.7	18.9
Salty cookies	7.2	12.4	3.0
Boneless beef	8.0	47.5	31.7
Beef with bone	7.0	25.4	9.9
Boneless pork	8.0	10.5	3.7
Chicken or hen meat (whole or in pieces)	6.0	22.5	12.6
Freshwater fish, fresh or frozen	5.0	4.0	8.3
Raw milk	1.0	16.2	76.5
Cottage cheese	6.5	6.1	7.3
Chicken and other bird eggs	3.3	51.1	36.1
Sunflower-seed, maize, soya, palm oil	4.0	310.5	35.1
Butter derived from cow milk	8.0	4.9	0.7
Margarines	6.4	7.1	1.0
Shortening	5.0	92.5	10.6
Tomatoes	1.9	3.9	22.7
Bulb onion	2.0	4.5	12.8
Scallions or green onions	1.8	1.7	12.9
Carrots	1.5	4.2	11.9
Beans, dried: common bean, <i>cargamanto</i> , <i>guarzo</i> , pigeon pea, lima bean, white bean, etc.	3.6	71.5	21.0
Peas, dried	2.0	21.2	6.2
Lentils	2.0	54.4	15.8
Plantain, green	1.2	54.5	54.5
Common potatoes: <i>pastusa</i> , <i>tocarreña</i> , <i>R12</i>	0.9	111.0	120.0
Cassava	1.0	45.2	37.0
Refined sugar	2.0	88.9	22.4
Raw or brown sugar	2.5	75.7	20.1
<i>Panela</i>	1.9	261.4	70.6
Salt	0.7	-	34.8
Cumins	16.7	-	0.2
Food coloring	10.0	0.6	0.5
Broth and stock concentrates (bouillon powder and cubes)	17.7	0.8	0.4
Ground coffee	14.0	4.8	4.4
Instant coffee in powder or granules	40.0	1.6	0.4
Chocolate tablets (sweetened and unsweetened)	9.0	15.1	4.4
Instant chocolate in powder or granules	11.9	1.6	0.4
Total	239.9	2,217.2	933.6

Source: MESEP technical team

Annex H. Calorie requirements

The table of nutritional requirements that was used to update the food baskets was generated from the *Calculating Population Energy Requirements and Food Needs* software, developed by FAO. The requirements were estimated by considering the demographic structure by age, sex and weight. The following table shows the calorie requirement according to the established criteria.

Table H1. Calorie requirement by age, weight and sex					
Age	Weight	Men	Age	Weight	Women
<1	7,47	621	<1	6,91	573
1	11,43	943	1	10,79	864
2	13,51	1.129	2	13,00	1.048
3	15,67	1.249	3	15,06	1.154
4	17,69	1.359	4	16,81	1.242
5	16,71	1.306	5	16,02	1.204
6	18,46	1.401	6	17,81	1.295
7	20,37	1.504	7	19,76	1.393
8	22,55	1.620	8	22,09	1.505
9	25,00	1.747	9	24,82	1.627
10	28,45	1.920	10	28,21	1.769
11	31,68	2.076	11	32,36	1.927
12	34,38	2.205	12	36,93	2.081
13	38,63	2.396	13	40,71	2.193
14	43,96	2.618	14	43,22	2.257
15	49,87	2.844	15	45,96	2.318
16	55,21	3.027	16	47,66	2.349
17	58,64	3.132	17	49,47	2.383
18-29	60,26	2.607	18-29	52,11	2.107
30-59	60,26	2.550	30-59	52,11	2.068
60 and more	60,26	2.108	60 and more	52,11	1.844

Source: MESEP technical team

The value of daily required calories went from 2,297 in 1988 to 2,068 in 2005 due to the following reasons in the methodology change as follows (cf. FAO et al (2001), p. 1-2):

- Calculation of energy requirements for all ages, based on measurements and estimates of total daily energy expenditure and on energy needs for growth, pregnancy and lactation;

- In the light of new data, modification of the requirements and dietary energy recommendations for infants and for older children and adolescents, in order to correct previous overestimations for the former and underestimations for the latter;
- Proposals for differentiating the requirements for populations with lifestyles that involve different levels of habitual physical activity, starting as early as six years of age;
- Reassessment of energy requirements for adults, based on energy expenditure estimates expressed as multiples of basal metabolic rates;
- Classification of physical activity levels based on the degree of habitual activity that is consistent with long-term good health and maintenance of a healthy body weight;
- Recommendations for physical activity for children and adults to maintain fitness and health and to reduce the risk of developing obesity and co-morbid diseases associated with a sedentary lifestyle;
- An experimental approach for factorial estimates of energy needs during pregnancy and lactation;
- Distribution in the two last trimesters of pregnancy of the recommendations for additional dietary energy needs.

Annex I. Values of extreme poverty and poverty lines by domain

Table I1. Average monthly value per person of extreme poverty and poverty lines for Colombia. 2002-2010 (current pesos)								
Domain		2002	2003	2004	2005	2008	2009	2010
National	IL*	51,316	55,274	58,416	62,180	80,197	83,020	83,581
	PL*	120,392	129,576	137,428	145,336	174,753	182,599	187,079
Urban	IL	53,890	57,977	61,257	65,154	83,746	86,748	87,401
	PL	134,733	144,749	153,296	161,841	193,701	202,200	207,005
Rural	IL	44,009	47,458	50,055	53,285	69,134	71,263	71,392
	PL	79,687	85,698	90,728	95,965	115,703	120,790	123,502

***IL=Extreme poverty Line; PL=Poverty line**

Source: MESEP Technical Team.

Annex J. Poverty and Extreme Poverty Incidence rates for the 13 metropolitan areas.

Table J1. Incidence of Poverty by metropolitan area 2002 - 2010

City	2002	2003	2004	2005	2008	2009	2010
Barranquilla*	43.1%	48.8%	45.9%	44.1%	43.3%	42.5%	39.5%
Bucaramanga	33.5%	34.0%	31.3%	31.1%	19.1%	13.9%	10.8%
Bogotá	31.3%	31.8%	28.7%	26.6%	19.6%	18.3%	15.5%
Manizales*	37.0%	39.5%	40.5%	36.5%	31.3%	27.3%	23.9%
Medellín	36.1%	34.5%	31.6%	29.4%	25.0%	23.9%	22.0%
Cali	33.1%	33.4%	31.6%	30.1%	28.4%	28.2%	26.1%
Pasto	44.7%	46.8%	46.2%	45.7%	39.5%	42.3%	42.6%
Villavicencio	33.9%	35.7%	32.3%	34.8%	26.7%	27.6%	25.4%
Pereira	32.1%	29.6%	29.6%	28.7%	27.5%	28.4%	26.6%
Cúcuta	52.9%	57.0%	55.2%	55.2%	42.1%	38.3%	39.2%
Cartagena	47.7%	43.6%	45.0%	37.7%	40.2%	38.4%	34.2%
Ibagué*	38.2%	40.0%	43.0%	39.5%	32.6%	28.7%	26.6%
Montería	48.7%	50.8%	50.6%	46.9%	40.5%	38.4%	39.5%

Source: MESEP calculations based on DANE's Household surveys (ECH for the 2002-2005 period spliced by MESEP and GEIH for the 2008 - 2010 period)

*The differences between the parallel survey that was conducted in 2008 with ECH specifications (ECHP) and the 2008 GEIH for the Barranquilla, Manizales and Ibagué geographic domains register a significance level that is greater than or equal to 90%.

Therefore, the gaps in the series between 2005 and 2008 can be explained in part by the limitations of any methodology for splicing series.

Coef.: Value of the indicator. Std. Err.: Standard Error

Note: It is worth mentioning that for some departments and cities, significant changes occur between the estimates levels of the 2002-2005 series (spliced ECH) and the year 2008 (GEIH), in spite of the fact that the differences between the 2008 ECHP and the 2008 GEIH are not significant. This is due to the characteristics of the splicing procedure whose accuracy reduces as the level of disaggregation for the analysis of the figures increases.

Table J2. Incidence of Extreme poverty by metropolitan area 2002-2010

City	2002	2003	2004	2005	2008	2009	2010
Barranquilla*	8.5%	10.7%	8.5%	7.8%	10.0%	8.2%	7.4%
Bucaramanga	6.1%	5.2%	4.7%	4.8%	2.6%	1.7%	1.2%
Bogotá	7.0%	6.9%	6.0%	4.7%	3.4%	3.2%	2.6%
Manizales*	7.3%	8.3%	9.0%	7.6%	6.8%	6.7%	4.7%
Medellín	7.9%	6.7%	5.6%	5.0%	6.1%	6.2%	5.6%
Cali	6.1%	5.3%	5.3%	5.0%	7.4%	7.4%	6.4%
Pasto	10.8%	11.3%	10.4%	11.7%	10.0%	10.6%	11.3%
Villavicencio	7.4%	6.8%	5.3%	6.2%	5.2%	5.3%	4.8%
Pereira	4.0%	3.0%	3.3%	3.7%	4.7%	4.6%	3.8%
Cúcuta	11.1%	13.2%	13.2%	11.3%	7.9%	6.9%	8.4%
Cartagena	9.5%	7.5%	6.9%	4.7%	6.9%	6.8%	6.1%
Ibagué*	8.4%	9.6%	10.4%	8.5%	7.3%	5.0%	4.3%
Montería	9.5%	11.2%	9.9%	9.3%	6.7%	5.9%	6.7%

Source: MESEP calculations based on DANE's Household surveys (ECH for the 2002-2005 period spliced by MESEP and GEIH for the 2008 - 2010 period)

* The differences between the parallel survey that was conducted in 2008 with ECH specifications (ECHP) and the 2008 GEIH for the Barranquilla, Manizales and Ibagué geographic domains register a significance level that is greater than or equal to 90%. Therefore, the gaps in the series between 2005 and 2008 can be explained in part by the limitations of any methodology for splicing series.

Coef.: Value of the indicator. Std. Err.: Standard Error

Note: It is worth mentioning that for some departments and cities, significant changes occur between the estimates levels of the 2002-2005 series (spliced ECH) and the year 2008 (GEIH), in spite of the fact that the differences between the 2008 ECHP and the 2008 GEIH are not significant. This is due to the characteristics of the splicing procedure whose accuracy reduces as the level of disaggregation for the analysis of the figures increases.

Annex K. Poverty and Extreme Poverty Rates by Department 2002-2010

Table K1. Incidence of Poverty by Department 2002 - 2010

Department	2002	2003	2004	2005	2008	2009	2010
Antioquia	47.5%	46.1%	44.0%	41.7%	38.4%	35.2%	31.3%
Atlántico	50.2%	52.3%	49.9%	48.7%	47.9%	47.5%	43.8%
Bogotá	31.3%	31.8%	28.7%	26.6%	19.8%	18.3%	15.6%
Bolívar	65.3%	53.7%	54.8%	51.5%	58.6%	56.9%	49.2%
Boyacá	67.2%	64.3%	63.4%	58.0%	57.0%	46.6%	46.6%
Caldas	45.8%	46.9%	47.7%	43.7%	43.0%	42.8%	39.6%
Caquetá	52.3%	52.3%	54.5%	53.8%	48.0%	52.2%	44.3%
Cauca	58.2%	60.7%	58.6%	55.0%	66.0%	66.1%	64.3%
Cesar*	61.1%	58.7%	59.1%	56.8%	63.4%	58.5%	53.6%
Córdoba	66.7%	64.1%	68.0%	64.2%	62.1%	61.8%	63.7%
Cundinamarca*	49.9%	48.7%	49.7%	44.0%	30.0%	26.8%	25.3%
Chocó	67.3%	69.1%	71.7%	74.0%	72.7%	67.4%	64.9%
Huila*	69.9%	67.3%	64.1%	55.6%	58.0%	57.4%	53.3%
La Guajira	65.5%	58.3%	57.5%	57.1%	69.8%	65.9%	64.3%
Magdalena	65.3%	59.7%	55.1%	58.7%	64.6%	58.1%	58.1%
Meta	41.6%	40.5%	38.0%	38.4%	32.7%	35.8%	32.3%
Nariño*	65.7%	69.5%	65.2%	58.4%	56.2%	54.6%	56.1%
Norte de Santander	56.1%	58.8%	61.0%	59.4%	50.2%	47.6%	43.1%
Quindío	46.5%	37.8%	46.0%	45.5%	43.3%	49.8%	43.3%
Risaralda	37.4%	34.6%	35.4%	35.5%	34.5%	32.2%	33.1%
Santander	44.5%	44.3%	42.1%	41.3%	30.8%	27.4%	21.5%
Sucre	69.1%	58.2%	64.9%	65.3%	67.0%	66.6%	63.7%
Tolima	55.3%	52.0%	56.2%	51.0%	47.0%	48.2%	45.1%
Valle del Cauca*	39.1%	39.3%	38.3%	37.0%	33.0%	33.0%	30.6%

Source: MESEP calculations based on DANE's Household surveys (ECH for the 2002-2005 period spliced by MESEP and GEIH for the 2008 - 2010 period)

* The differences between the 2008 ECHP and the 2008 GEIH for the Cesar, Cundinamarca, Huila, Nariño and Valle del Cauca geographic domains register a significance level that is greater than or equal to 90%.

Therefore, the gaps in the series between 2005 and 2008 can be explained in part by the limitations of any methodology for splicing series.

Coef.: Value of the indicator. Std. Err.: Standard Error

Note: It is worth mentioning that for some departments and cities, significant changes occur between the estimates levels of the 2002-2005 series (spliced ECH) and the year 2008 (GEIH), in spite of the fact that the differences between the 2008 ECHP and the 2008 GEIH are not significant. This is due to the characteristics of the splicing procedure whose accuracy reduces as the level of disaggregation for the analysis of the figures increases.

Table K2. Incidence of Extreme poverty by Department 2002-2010

Department	2002	2003	2004	2005	2008	2009	2010
Antioquia	17.1%	15.1%	14.4%	13.7%	15.4%	13.0%	10.4%
Atlántico	12.7%	11.4%	9.5%	9.5%	12.3%	11.5%	9.4%
Bogotá	7.0%	6.9%	6.0%	4.7%	3.5%	3.2%	2.6%
Bolívar	30.4%	13.7%	14.6%	12.1%	26.0%	22.2%	14.7%
Boyacá	39.2%	35.7%	33.0%	26.8%	30.3%	19.1%	19.1%
Caldas	12.7%	12.9%	13.4%	12.0%	13.7%	12.8%	11.1%
Caquetá	21.5%	24.6%	19.0%	21.4%	17.0%	16.6%	10.0%
Cauca	24.5%	26.7%	25.6%	24.4%	41.4%	38.8%	35.9%
Cesar*	19.3%	13.0%	18.7%	15.4%	30.2%	23.8%	18.5%
Córdoba	33.0%	32.7%	32.4%	28.2%	23.9%	25.2%	25.8%
Cundinamarca*	18.2%	16.9%	15.9%	14.8%	9.6%	8.4%	7.8%
Chocó	32.2%	36.3%	39.3%	43.3%	45.4%	39.2%	33.6%
Huila*	35.4%	33.5%	27.1%	20.4%	29.7%	28.9%	25.9%
La Guajira	27.9%	20.3%	18.3%	21.7%	43.7%	34.8%	37.4%
Magdalena	23.5%	14.8%	14.3%	15.1%	31.7%	23.8%	23.5%
Meta	13.7%	13.3%	9.3%	10.6%	11.1%	10.5%	9.7%
Nariño*	33.4%	32.8%	29.1%	23.9%	20.9%	21.0%	18.7%
Norte de Santander	16.3%	18.3%	20.7%	19.4%	17.2%	14.1%	11.0%
Quindío	12.7%	10.3%	13.4%	10.4%	15.2%	17.2%	12.3%
Risaralda	7.8%	6.1%	7.1%	8.4%	10.1%	7.6%	7.6%
Santander	16.0%	15.1%	12.4%	11.8%	9.8%	8.1%	4.7%
Sucre	29.0%	19.0%	22.9%	23.7%	33.4%	29.1%	28.5%
Tolima	20.9%	21.0%	21.0%	15.9%	17.6%	17.7%	17.3%
Valle del Cauca*	9.5%	9.2%	8.5%	8.5%	10.1%	9.6%	8.4%

Source: MESEP calculations based on DANE's Household surveys (ECH for the 2002-2005 period spliced by MESEP and GEIH for the 2008 - 2010 period)

* The differences between the 2008 ECHP and the 2008 GEIH for the Cesar, Cundinamarca, Huila, Nariño and Valle del Cauca geographic domains register a significance level that is greater than or equal to 90%.

Therefore, the gaps in the series between 2005 and 2008 can be explained in part by the limitations of any methodology for splicing series.

Coef.: Value of the indicator. Std. Err.: Standard Error

Note: It is worth mentioning that for some departments and cities, significant changes occur between the estimates levels of the 2002-2005 series (spliced ECH) and the year 2008 (GEIH), in spite of the fact that the differences between the 2008 ECHP and the 2008 GEIH are not significant. This is due to the characteristics of the splicing procedure whose accuracy reduces as the level of disaggregation for the analysis of the figures increases.