Research Paper 3

GUIDE TO THE DEMOGRAPHIC MODULE FOR POVERTY ANALYSIS AND PROJECTION (DMPAP): AN EXCEL WORKBOOK WITH APPLICATIONS TO VENEZUELA, BRAZIL, AND JAMAICA

by Ralph Hakkert

Chief Technical Advisor, Project RLA5P201

GUIDE TO THE DEMOGRAPHIC MODULE FOR POVERTY ANALYSIS AND PROJECTION (DMPAP): AN EXCEL WORKBOOK WITH APPLICATIONS TO VENEZUELA, BRAZIL, AND JAMAICA

Ralph Hakkert

Chief Technical Advisor, Project RLA5P201

Research Paper 3

IPEA/UNFPA Project RLA5P201: Regional support to Population and Development in the implementation of the MDGs in the LAC Region

Brasília DF, August 2007

An abbreviated version of this paper in Spanish, without the example of Jamaica, can be found in *Notas de Población* no. 84, Oct. 2007, pages 149-173, under the title "Población y pobreza: un modelo a escala de hogar y ejemplo de su aplicación en la república boliviariana de Venezuela y Brasil".

Research Papers of Project RLA5P201 - Regional support to Population and Development in the implementation of the MDGs in the LAC Region:

1. Sergei Soares – Relative and absolute demographic bonus in schooling
2. André Junqueira Caetano; Durval Magalhães Fernandes & José Irineu Rangel Rigotti - Migration and the Millennium Development Goals: Latin America and the Caribbean
3. Ralph Hakkert - Guide to the Demographic Module for Poverty Analysis and Projection (DMPAP): an EXCEL workbook with applications to Venezuela, Brazil, and Jamaica
4. Project RLA5P201 – Potential contributions to the MDG agenda from the perspective of ICPD: summary and programme implications
5. Proyecto RLA5P201 – Contribuciones potenciales a la agenda de los ODMs a partir de la perspectiva de la CIPD: resumen e implicancias programáticas
6. Ralph Hakkert: Demographic Module for Poverty Analysis and Projection (DMPAP): an application to Suriname
7. Ralph Hakkert: The demographic bonus and population in active ages
8. Ralph Hakkert: Un análisis del efecto de la fecundidad no deseada sobre la pobreza a nivel de los Departamentos y zonas de residencia de Honduras, 2006
9. Ralph Hakkert: Demographic Module for Poverty Analysis and Projection (DMPAP): an application to Bolivia

Introduction

Traditionally, economists project the incidence of poverty as a direct function of economic growth, using the so-called income-poverty elasticity, an empirically measured index which quantifies how much poverty reduction occurs for each 1% increase in the per capita income of a country. If the elasticity is high, poverty responds strongly to economic growth, possibly because most of the poor are living just below the poverty line. If it is low, even large increases of the per capita income will be relatively ineffective in reducing poverty. Economic growth may not be neutral in terms of its distributional effects, and it may actually exacerbate poverty by worsening the income distribution, but even if it does not alter the income distribution, the reduction of poverty resulting from 1% of growth is usually less than 1% (i.e. the elasticity is smaller than 1), especially if the initial distribution is very unequal. In Latin America, in particular, the income-poverty elasticities tend to be relatively low, exactly because of the enormous distributional inequalities found in this region.

In December of 2002, ECLAC, IPEA, and UNDP issued a joint publication on poverty reduction in Latin America under the title *Meeting the Millennium Poverty Reduction Targets in Latin America and the Caribbean*, where they introduced a methodological innovation which makes it possible to better visualise the role of inequality in poverty reduction and in the determination of the income-poverty elasticity. The ECLAC/IPEA/UNDP model uses two, instead of one, parameters to transform the income distribution for the purpose of poverty projections. The parameter β, as they call it, is the more conventional one: it stands for economic growth, representing a proportional increase of all incomes, i.e. if $\beta = 0.4$ all incomes are multiplied by a factor 1.4, as illustrated in Figure 1.B. In the absence of distributional changes, β may be thought of simply as the growth of the per capita national income. The innovation is in the parameter α, which expresses the income distribution effect: it represents a linear contraction of the entire income distribution in the direction of the over-all mean, proportional to the initial distance with respect to the mean. For example, if $\alpha = 0.3$, all incomes move 30% toward the over-all mean, as illustrated in Figure 1.A. In the extreme case where $\alpha = 1$, all incomes collapse at the mean and income inequality ceases to exist. Usually, an α parameter larger than 0 will be associated with poverty reduction, but not necessarily so. In very poor countries, where the poverty line is higher than the mean income, the contraction of the income distribution in the direction of the mean will actually exacerbate the incidence of poverty, although it will diminish the depth of poverty of the very poorest.

Figure 1.A.: Effect of a change in the α parameter on the income distribution and poverty

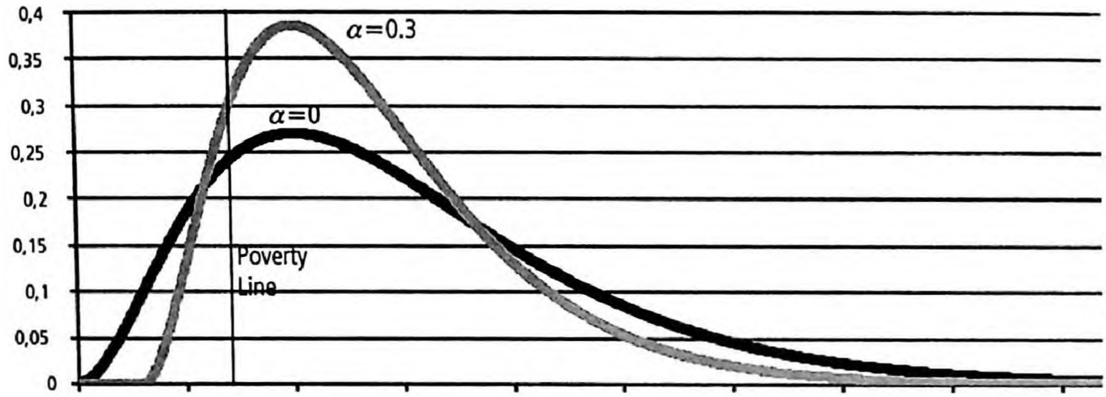
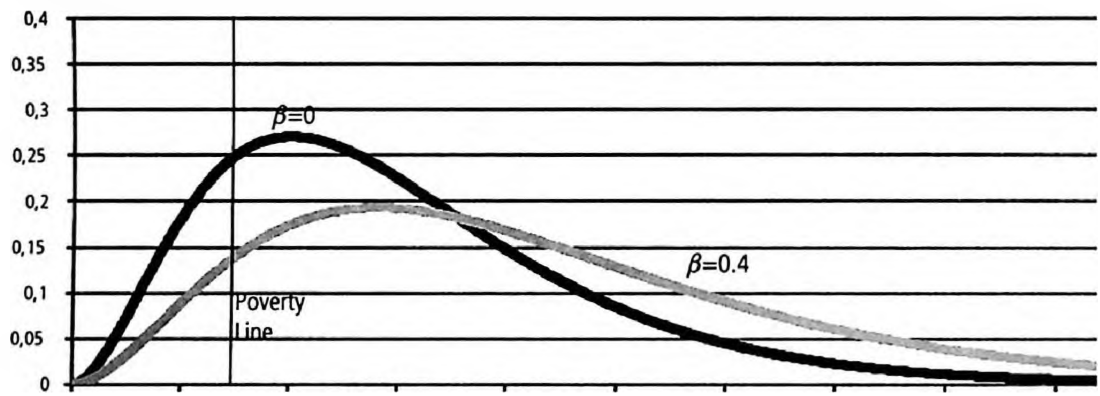


Figure 1.B.: Effect of a change in the β parameter on the income distribution and poverty



Analytically, it can be demonstrated that the Gini index is reduced proportionally to α, which is quite convenient. Thus, by varying α and β, it is possible to demonstrate how different combinations of economic growth and reduction of inequality will affect poverty. A slightly controversial aspect of the model, which affects its ability to actually predict future poverty, is exactly the specific mechanism by which the Gini index is linked to the transformation of the income distribution. A particular reduction of the Gini index can, in principle, be caused by different changes in the distribution. For example, the transfer of incomes can take place between the extremes of the distribution or rather between the middle income strata slightly below or slightly above the mean. In the first case, the impact on poverty – and particularly extreme poverty – will be much greater than in the second. The transformation parametrised by α which is used in the ECLAC/IPEA/UNDP model implies a much greater change in the tails of the income distribution than in the intermediate range. Consequently, the model implicitly postulates a strong relationship between changes of the Gini index and their poverty effects. If this relationship is indeed as strong as the model suggests is a matter for empirical verification.

Following the logic of the model and using different combinations of α and β, one can define *iso-poverty curves*, to visualise the combinations of income increase and reduction of inequality which lead to a given result in terms of poverty reduction. ECLAC, IPEA, and UNDP used these curves to demonstrate that relatively modest reductions of income inequality in Latin America may result in poverty levels that are equal or smaller than the ones brought about by substantial proportional increases of all incomes. Hence the conclusion for the purpose of policy formulation is that social policies aimed at reducing income inequality may be more effective in the struggle against poverty than stimulating economic growth, even if the latter can be achieved without deteriorating the income distribution. As was mentioned above, whether this finding actually represents the reality of the economic and social policy trajectories in the countries of the region or whether it is simply a built-in characteristic of the model is something to be verified empirically.

Neither the traditional methodology of projecting poverty by means of income-poverty elasticities nor the more recent methodological proposal of ECLAC/IPEA/UNDP give any explicit consideration to the demographic transformations that are occurring in Latin America and other parts of the developing world as a result of the demographic transition, which is reflected especially in the decline of fertility. More often than not, the only consideration that economists have given to demographic trends in connection with poverty reduction is that demographic *growth* may reduce the per capita economic growth rate and thus reduce β. Of course, the income-poverty elasticities and the α and β parameters, to the extent that they are determined empirically, incorporate this and other demographic effects implicitly, but usually no elements are provided to evaluate how these parameters may be affected in the future by changes in demographic trends and particularly in the population *composition*.

In this context, it is important to consider that, for the purpose of poverty analysis, populations do not consist of isolated individuals, but of households which pool their resources.[1] Poverty, therefore, is not an individual characteristic, but a household characteristic. Consequently, the composition of these households, in terms of the sex and age of their heads as well as the other household members, is an important determinant of poverty. In a 1999 article, Hausmann and Székely

[1] To what extent it is actually true that households pool their resources and to what extent the internal distribution is equitable is a matter of some dispute that is currently being researched, particularly with respect to the possible gender biases that may affect the distribution. Some discussion of the recent literature on this issue can be found in the article *Poverty among women in Latin America: feminization or over-representation ?* by Marcelo Medeiros and Joana Costa, International Poverty Centre, IPEA/UNDP, Working Paper 20, 2006. Conventional poverty analysis, however, does not consider such internal distributional biases and neither does the present paper.

clearly show, for instance, that poverty in Latin America varies considerably by the number of dependent children in the household.[2] As the LAC region advances in its demographic transition, the relations between the different age groups will become more favourable, at least during the next few decades. At the macro level, this phenomenon is known as the demographic "bonus" or "dividend" or "window of opportunity", a transitory opportunity related to the present increase of the proportion of the population in the active age groups. This macro phenomenon has a counterpart at the micro level, in the composition of families and households. In upcoming years, household sizes in most countries of the LAC region will decrease and their dependency ratios in 2015 will be smaller than at present. All of this has important implications for the reduction of poverty which so far have not been sufficiently accounted for.

Although the ECLAC/IPEA/UNDP methodology does not address population composition as an explanatory factor, one of its authors, in a separate publication, has investigated the issue from a historical perspective. Paes the Barros et al.[3] analysed data from the Brazilian Household Survey (PNAD) over the 1976-1996 period and concluded that the effect of compositional factors on poverty reduction in households with heads aged 36-40 years born between 1910 and 1960 was equivalent to the one that would have been produced by an additional average economic growth rate of 0.4% p.a. during this period. Over this period, the head count index of poverty came out 7 percentage points below what it would have been in the context of a constant age composition of the population, particularly due to the decline of the population share below the age of 22 and even more particularly the population share under age 15. In the Northeast, the least developed region of Brazil, the reduction was about twice as large. The other important conclusion they reached was that, while aggregate changes in household composition were an important determinant of poverty reduction over time, particularly in the Northeast, the differences in

[2] Ricardo Hausmann & Miguel Székely. 1999. *Inequality and the family in Latin America.* Washington DC, BID, Office of the Chief Economist, Working Paper 393. Although they recognise the importance of macro determinants, like the structure of the economies of the region and geographical, cultural, and ethnic factors, they call attention to three personal characteristics which explain an important part of income variations between households: fertility, female labour force participation, and education. The same conclusion is also is also reached by Ricardo Paes de Barros; S. Duryea & Miguel Székely. *What's behind the Latin American inequality?* 1999. Washington DC, BID, Office of the Chief Economist.

[3] Paes de Barros, Ricardo *et al.* 2001. "Demographic changes and poverty in Brazil." In: Nancy Birdsall, Alan C. Kelley & Steve W. Sinding (eds.). *Population matters: demographic change, economic growth, and poverty in the developing world.* Oxford, Oxford University Press: Ch. 11. Despite the recognition of the method's shortcomings in this regard, these authors declare their approach, which is similar to the one used here, almost ideal for the estimation of the direct effects of demographic change on poverty. The alternatives that they identify are regression analysis on cross-national country data and the Computable General Equilibrium (CGE) model. However, the first cannot separate direct from indirect effects. In addition, it either has to assume that demographic change was exogenous or has to rely on debatable choices of instrumental variables. The second alternative can, in theory, provide estimates of both the direct and the indirect effects, but it is affected by uncertainty regarding the correct specification of the model and its parameters, which may exert a substantial effect on the results.

household composition between poor and non-poor at any given moment in time did not explain a significant proportion of the income differential.

It should be emphasized that the poverty reduction effects that are the object of this paper are most significant in the long term and therefore contain a great deal of inertia. The important population related reduction of poverty expected in the three countries analysed here during the 2004-2015 period is, to a large extent, the result of population trends that have occurred during the past decades and that are now yielding economic returns. This, of course, is the same issue as the one regarding the interpretation of the demographic bonus. While it is certainly important to make the point that the collective investments in fertility reduction during the past decades generate important economic returns that will become evident in the years ahead, it does leave planners and decision makers with the question about the relevance of the model with respect to the policy decisions they are facing now. As will become evident in this paper, there are certainly policy issues to be decided on now that will affect the poverty outcomes over the next decade or two, but undoubtedly much of what will happen during this period is already pre-determined by the inertia of past demographic trends.

The general approach

The methodology proposed in this paper follows the basic principles of the ECLAC/IPEA/UNDP projection model and of the historical poverty analysis by Paes de Barros et al. referred to in the previous section, but it extends the analysis by applying it in a projection framework and proposes some analytical alternatives. Like the original ECLAC/IPEA/UNDP model, DMPAP consists of a decomposition of trends which makes it possible to extrapolate in more meaningful ways.

The differences between the present model and the approach chosen by Paes de Barros et al. have to do with the following aspects:

1. The present model uses a more complex division of categories of households and household members. Paes de Barros et al. focused on households with heads in the 36-40 year age range and divided their members in four age categories: 0-14, 15-21, 22-64, and 65+. The present document uses a slightly finer age division: 0-9, 10-14, 15-24, 25-34, 35-49, 50-64, 65-74, and 75+. The 75+ age category was added, despite its small size, to allow the assessment of the impact of ageing in an age category that is predominantly dependent on transfers for its income. In addition, the present model works with heads of households in all age categories, differentiates them by sex, and relates the income generating capacity of household members to their relationship with the head, in addition to their age and sex.

2. Paes de Barros et al. assumed proportional changes in the number of household members in each age category, in accordance with historical trends of the average number of household members in that age category. Because the present model deals with projections, it does not have access to that information, but instead has to rely on age-sex specific population projections at the level of the entire population, which are not broken down by household membership. To assign these individuals to households, a number of different procedures may be followed, including the proportional adjustment of all numbers of household members by age and sex, in order to achieve consistency with the population projections, which would be the closest equivalent to the procedure followed by Paes de Barros et al. However, the model also allows more elaborate hypotheses in which the adjustment is differentiated by income levels or by age and sex of the head of household. This lends additional realism to the model and may cause substantial changes in the poverty projections.
3. Maybe most importantly, the procedure for determining the income generating potential of household members in different categories is distinct from the one used by Paes de Barros et al., who related the income generating potential of individuals directly to personal income data from household surveys. But this does not take into account how the interactions and mutual support between household members may leverage their income generating capacity, e.g. by relieving some of them of certain non-economic household tasks, so that they can participate in the labour market. This paper will explore some alternatives, depending on the specific data available in each of the countries.

The over-all effect of these differences (particularly the third and sometimes the second) tends to be a larger impact of demographic factors on poverty reduction than in the previous model. Whereas Paes de Barros et al., in their Brazilian study, estimated that the demographic effect might be equivalent to 0.4% of additional economic growth, the present example, with data from Venezuela, suggests that over the 2004-2015 period this figure may actually be in excess of 1.0%. Apart from the differences in model specification mentioned above, this also owes to the fact that the period analysed by Paes de Barros et al. was rather large and characterised by distinct rhythms of demographic change.[4] Furthermore, as the authors themselves indicate, the effect in the less developed regions of Brazil, like the Northeast, was substantially larger than the national average.

[4] Fertility in Brazil only started declining appreciably by the late 1960s, more than 20 years into the 50 year period covered by the analysis.

Like its precursor, the present model does have some limitations – both from a demographic and an economic viewpoint – which should be made explicit from the start. The first limitation is that, even though the demographic projections it uses are based on a traditional cohort-component approach (the preferred method of demographers), the household compositions are projected "horizontally" (in terms of the Lexis diagram), i.e. by age-sex category of the head of household, and not by cohort. The reasons have to do with the technical difficulties of making a cohort projection work if the unit of analysis is the household and not the individual. They will be explained in some more detail in the next section.

The other major limitation is that, like the models previously referred to, DMPAP provides a statistical decomposition of trends and not a representation of economic theory as it does not consider, for example, how the relative size of the production factors might affect their prices. The implicit ceteris paribus hypothesis of the model is that population sizes and compositions vary, but leave the income generating capacities of household members with particular age, sex, and household status characteristics unaffected, unless the economic environment changes through the α and/or β parameters. There is no intrinsic mechanism in the model that suggests how α and/or β might change as a direct result of demographic trends, e.g. capital dilution because of labour force growth beyond the growth of investments. Similarly, it does not consider how changing age structures affect savings behaviour and thereby investments. These effects, which Paes de Barros et al. identify as *indirect* effects of demographic change on poverty, have to be incorporated exogenously, through hypotheses on the evolution of α and β.

In addition to structural limitations, there are certain conceptual limitations of models of this nature, that have to do not so much with their internal structure as with the interpretation of their results. First of all, the poverty concept that underlies the analysis presented here is the monetary poverty concept implied by the first Target of the Millennium Development agenda. This means that certain broader dimensions of poverty, as expressed by the capability approach, among others, are not considered here. Services received by households that may enhance their options are considered only to the extent that they generate income, e.g. productivity increases resulting from access to health care or child care services, or change the household composition, e.g. births avoided because of access to reproductive health services, but not the intrinsic benefits of having access to, for instance, cultural activities. Similarly, the model (like the monetary poverty concept itself) does not consider patterns of expenditure, e.g. whether or not households have to pay for their education, health care, basic sanitation, or domestic service (in case the woman works outside the home).

Finally, it should be pointed out that, although models of this kind reproduce a major or maybe even the largest part of the population effects relevant to poverty reduction, they do not exhaust all the causal pathways by which policy actions in the population area impact on poverty. Several other mechanisms do not operate through the age-sex structure of the population, which is the intermediate variable emphasized here, but rather affect α and/or β directly or change the relative contribution of household members of particular age and sex characteristics to the household budget. The issue of adolescent fertility is a case in point. Births to very young mothers, particularly if they are unwanted, may cause a number of effects that increase poverty, not all of which are adequately represented by the model. The fact that the families of the young mothers have another mouth to fill is evidently reflected in the model, as is the fact that this situation may limit the labour participation of the mother. But the fact that her schooling may be truncated as a consequence of this event, thus curtailing her economic opportunities later in life, is not something the model takes into account.

Some of these mechanisms may be much more amenable to policy actions than the ones outlined in this paper, even though their poverty impact may be less dramatic. Actions to promote continued schooling for adolescent mothers may increase the productivity of women in the 15-24 age group. The promotion of institutionalised child care may reduce the negative impact of young children on the household income.[5] The prevention of household violence may reduce the number of days of sick leave and thereby increase household income (β). And the generalisation of social health insurance may prevent catastrophic disease spells, thus stabilising income levels over time. and increasing α. It is important, therefore, to remember that, to obtain a comprehensive picture of the entire network of population-poverty interactions, the population effects modelled in this paper should be complemented by others.

Particularities of country applications

One of the difficulties that arise in the standardised application of DMPAP is that the national surveys underlying the application of the model differ in several ways that need to be accommodated in each particular application. This accommodation affects the format of the DMPAP Work Book as such, but it requires even greater adjustments in the SPSS or STATA programmes needed to set up the basic data in the Data1 and Data2 spreadsheets (see the next section).

[5] In illustrative application of this kind is shown in the final section, on Venezuela.

Annex 1 of this paper contains an illustrative SPSS syntax file that was used to generate the Data2 spreadsheet from the 2005 National Household Survey (PNAD) of Brazil. This syntax file contains a number of particular operations that were necessary in Brazil, but may not be required in other countries. Some of these are fairly trivial, e.g. the need to use a DATA LIST command because the original data came in ASCII format, rather than as an SPSS file. But others are more profound. The Brazilian household survey is huge in comparison to similar surveys in other countries, with 116,452 households. This is far too much to accommodate in the Data2 spreadsheet, which has only 7,199 lines, each corresponding to a household. The solution, therefore, was to apply a clustering process in which households with similar compositions and per capita income levels were merged, summing their household weights and using their average composition data, as if, in fact, they were one single household. In countries like Bolivia, Jamaica, or Suriname there was no need for this, so that the SPSS syntax could be abbreviated.

In the standard application, the household information in the Data1 and Data2 spreadsheets is organised into blocks of up to 600 lines, each specific for an age and sex category of the heads of household. In countries like Brazil, where clustering was used, the clustering process is structured in such a way that the number of sub-categories within each of these main age-sex categories will be smaller than 600. However, in countries where the original households were used as units, there may be age-sex categories which require more than 600 lines. This happened, for example, in Bolivia and Suriname. Consequently, the block structure in these cases had to be adapted.

The clustering operation mentioned in the previous paragraph is primarily relevant for the SPSS syntax file. Once it has been implemented, the resulting Data1 and Data2 files are almost identical to the ones for countries without household clustering. A more complex issue is that of single or multiple poverty lines. Some countries, like Brazil, Venezuela and in practice Suriname, have a single poverty line for the entire country. In the illustrative SPSS syntax file in Annex 1, this is the line of R$ 150 in the COMPUTE pline=150 command. Other countries have more than one poverty line. Jamaica, for instance, has one for Metropolitan Kingston, one for other urban areas, and one for the rural areas. Bolivia has at total of 9 different poverty lines, differentiated by Departments and rural-urban areas of residence. These complications require adaptations not only of the SPSS syntax file, but also of the DMPAP Work Book itself. Below, this will be illustrated with the case of Jamaica.

Most of the countries where the DMPAP model has been applied so far use the per capita household income or consumption criterion to define poverty. Suriname, however, uses an adult equivalence scale that distinguishes between individuals under and over age 15 and also contains an economy-of-scale parameter. This does not affect the SPSS syntax file, but it does require some adaptations of the DMPAP Work Book.

The most complex issue, however, is the determination of the income generating capacity coefficients (see below), which indicate the relative income generating capacity of household members of different age, sex, and relationship-to-the-head-of-household categories. In countries that have individual income data, such as Brazil, the easiest procedure is to estimate these coefficients based on this information, even though there is still room for some variation in the precise way to do this. In countries like Jamaica and Suriname, however, this information is not available. In Jamaica this is because the 2004 Living Standards Measurement Survey strictly followed the consumption, rather than the income criterion. In Suriname, it is because the MICS asked for both income and consumption, but only at the household level. In these cases, there are basically two strategies which one can follow:

1. Use the coefficients of another country; or
2. Determine the coefficients indirectly, through a procedure which will be illustrated below.

The first alternative does not appreciably alter the DMPAP Work Book, but the second is more difficult to apply and needs some additional resources. It is also possible to use a combination of both approaches, as will be illustrated with the case of Jamaica.

The detailed structure of the model

The spreadsheets

DMPAP has been implemented as a set of eight interlinked EXCEL spreadsheets:

Main (Principal) is a relatively small spreadsheet with the main input parameters and results for the two base years and a projected year, where the latter may be chosen anywhere between the second base year and the projection horizon;

Population (Población) is also a small spreadsheet, which contains the projected populations by age and sex, as provided by the user, for all the years between the second base year and the projection horizon, plus a limited number of manipulations of these data, to allow the user to modify the implicit fertility rates, to assess their impact on the poverty projection;

Param (Parám) is also a small spreadsheet containing a number of less frequently modified parameters relating to the relative contribution of different kinds of household members to the household income and some calibration parameters for the projections, all of which are subject to optimisation;

Data1 (Datos1) is a large spreadsheet with the household data for the first base year that can be up to 7200 lines long, preferably divided in 12 blocks of up to 600 lines each, although in some countries (e.g. Bolivia and Suriname) a slightly different division had to be used;

Data2 (Datos2) is the same as Data1, but applies to the second base year;

Gini is a small spreadsheet that computes a number of Gini indices that are reported in Main;

Trends (Tendencias) is a summary of the Data1 and Data2 spreadsheets that computes time trends for the purpose of projections;

Projection (Proyección) has the same structure and the same number of lines as Data2, but contains the projected results for a hypothetical set of households based on Data2 and the parameters and trends specified in Main and Trends.

Apart from the initial setup of the data, the only spreadsheets that the user needs to manipulate are Main, Param, and Trends, and within these three spreadsheets only the fields in yellow or green. All other fields contain either data or results computed based on the data and should not be accessed by the user. The difference between yellow and green fields is the following. The yellow fields can be substituted freely, for example, to change the projection year from 2015 to 2018 in the Main spreadsheet, or to restate the projected change in a particular headship ratio from 0.05 to 0.1 in the Trends spreadsheet. The green fields can also be changed by the user, but these are usually subject to optimisation in order to obtain the best results. The typical example is that of the Parameters column in the Param spreadsheet, which should be optimised in order to obtain zeros in the adjacent Criterion column. In addition to the yellow and green fields, there are cells with underlined numerical values. These are normally not manipulated by the user. The reason why they are underlined is to signal that they contain matrix or array formulas, which are used because these provide shortcuts to a lot of computations involving large data ranges such as the columns of Data1 and Data2. Matrix formulas should be enclosed between brackets { ... }; if they are not, they will produce erroneous results. In case they are accidentally lost, braces must be restored by selecting the cell and then simultaneously pressing <Ctrl><Shift><Return>.

The data spreadsheets Data1 and Data2 contain the following information:
categor – A summary indicator containing information on the age and sex of the household head, the presence of a spouse, the number of children under 10, and the income category (only in countries where households were clustered into categories, but not in countries like Jamaica, where the number of cases was small enough to use the original households as units);

totincome (ingrtot) – Total household income;
weighthd (pesojefe) – Sample weight of the household or the head of household;
num09 – Number of children under age 10 in the household;
num1014m – Number of males aged 10-14 in the household;
num1014f – Number of females aged 10-14 in the household;
num1524m – Number of males aged 15-24 in the household;
num1524f – Number of females aged 15-24 in the household;
num2534m – Number of males aged 25-34 in the household;
num2534f – Number of females aged 25-34 in the household;
num3549m – Number of males aged 35-49 in the household;
num3549f – Number of females aged 35-49 in the household;
num5064m – Number of males aged 50-64 in the household;
num5064f – Number of females aged 50-64 in the household;
num6574m – Number of males aged 65-74 in the household;
num6574f – Number of females aged 65-74 in the household;
num75m – Number of males aged 75 or more in the household;
num75f – Number of females aged 75 or more in the household;
numc1524 – Number of spouses aged 15-24 in the household;
numc2534 – Number of spouses aged 25-34 in the household;
numc3549 – Number of spouses aged 35-49 in the household;
numc5064 – Number of spouses aged 50-64 in the household;
numc6574 – Number of spouses aged 65-74 in the household;
numc75 – Number of spouses aged 75 or more in the household;
members (miembros) – Number of household members, including children;
percap – Per capita household income, usually not weighted for age and sex (but in Suriname an adult equivalent weighting system was used).

In the case of Jamaica and other countries with multiple poverty lines, the spreadsheet also contains the following additional columns:
perline – The poverty line applicable to each specific household due to its location;
poor – 1 if the household is poor, 0 if it is not;
gap – Difference between the per capita income and the specific poverty line.

All of these except **poor** and **gap** are base data directly obtained from the household surveys (Living Standard Measurement Surveys or similar instruments) in the two base years. In countries that use adult equivalence weights, like Suriname, **percap** is also computed, to incorporate the weighting information. In the illustrative cases of Venezuela and Jamaica, Data1 and Data2 are based on the household surveys taken in the first semesters of 1999 and 2004.[6] In Brazil, the 1999 and 2005 National

[6] The 5 year interval used here is recommended, although not strictly necessary.

Household Surveys (PNADs) were used and in Jamaica the 1999 and 2004 Living Standard Measurement Surveys (LSMS). In addition, Data1 and Data2 have four final variables (**autonomy, p-factor, pc-gr(o)up(o)**, and **stratum (estrato)**), which are computed from the basic data and whose meaning will be explained below.

The age and sex of the head of household is not stated explicitly, but is part of the information contained in the **categor** variable or, in the absence of a **categor** variable, it is implicit in the ordering of the blocks of households in the Data1 and Data2 spreadsheets. Lines 2-599 contain information of households with male heads aged 15-24. Lines 600-1199 are reserved for households with male heads aged 25-34, and so on, up to lines 3000-3599, which are reserved for households with male heads over age 75. For the female headed households, the same pattern is repeated in lines 3600-7199. When replacing the base line data of the Data1 and Data2 spreadsheets with data for another country or time period, care should be taken to maintain this same pattern, i.e. lines 2-599, 600-1199, 1200-1799, etc. should be blanked out and replaced by the new data for households with heads aged 15-24, 25-34, 35-49, etc. Again, in some countries, such as Bolivia and Suriname, it was necessary to deviate slightly from this pattern because there were more than 600 households in some categories.

As mentioned before, the Brazilian and Venezuelan data involve some clustering in order to limit the size of the data bases Data1 and Data2.[7] The way this clustering process works is as follows. Households are merged, summing their household weights and averaging their other variables, if:

- Their heads belong to the same sex and age (15-24, 25-34, 35-49, 50-64, 65-74, 75+) category;
- They have the same number of spouses (none or at least one);
- They have the same number of children under age 10 (to a maximum of 5); and
- They belong to the same per capita income category.

In the case of Venezuela, the per capita income categories were defined as follows:
Category 0: Less than 0.3 times the official poverty line;[8]
Categories 1-10: Ten equally spaced income brackets between 0.3 and 0.5 times the official poverty line;
Categories 11-26: Sixteen equally spaced income brackets between 0.5 and 1.0 times the official poverty line;

[7] Even with this clustering, the total size of the Venezuelan Work Book is still as large as 13.4 Mb.

[8] In the case of Venezuela, the official poverty line was 48,628.8 Bolivares in the first semester of 1999 and 123,880 Bolivares in the second semester of 2004. If there is one than more poverty line, as in Jamaica, the criterion should be applied to the poverty line of the appropriate territorial unit.

Categories 27-36: Ten equally spaced income brackets between 1.0 and 1.5 times the official poverty line;
Categories 37-41: Five equally spaced income brackets between 1.5 and 2.0 times the official poverty line;
Categories 42-46: Five equally spaced income brackets between 2.0 and 3.0 times the official poverty line;
Category 47: Between 3.0 and 4.0 times the official poverty line;
Category 48: Between 4.0 and 7.5 times the official poverty line;
Category 49: More than 7.5 times the official poverty line.

This division, however, may vary by country. In Brazil, where the operational poverty line of half a minimum wage per capita is lower compared to average incomes than in Venezuela, the categories were adapted to reflect this difference. The 50 categories define the last two digits of the **categor** variable (modulo 50), whereas the other merging criteria define the remainder of **categor**. Because households with the same characteristics are merged, each value of **categor** is listed at the most once, and since there are only 600 possible values for **categor** in each age/sex category of heads of households, this is also the maximum number of lines for each block.

One of the consequences of the clustering process is that some information regarding per capita income levels is lost in the process. Thus, depending on the particular poverty line used in the analysis, the results may be slightly different from those that would be obtained from an analysis of the full sample. The income categories were chosen in such a way that the poverty levels estimated for the two base years will not be affected if the poverty line is fixed at the official level or at 0.3, 0.5, 1.5, 2.0, or 3.0 times the official level.[9] However, a poverty analysis in terms of a hypothetical poverty line chosen at 0.75 times the official level may yield results that are up to 1% larger or smaller than the results that would be obtained from an analysis of the original survey sample. Analyses with poverty lines lower than 0.3 or higher than 3.0 times the official level are not recommended because in these ranges the clustering of households has particularly strong effects. In some circumstances, this may be considered inconvenient, but the alternative, of carrying over the entire data base of the base line surveys to Data1 and Data2, in several countries would challenge the operational limitations of EXCEL spreadsheets.

The Main spreadsheet is mostly based on Data1, Data2, and Projection, including summary results for the projections, except for a few parameters that the user can define and that are marked in yellow. In cells C3, C4, and C5, the user can specify the first base year, second base year, and the year of the projection. In D3,

[9] The same holds for the intermediate transitions between intervals, i.e. 0.32, 0.34, 0.36, etc. times the official poverty line.

D4, and D5, one may specify one of two years of reference (either C3 or C4) for the income generating coefficients capacity c_{ijk} which will be explained below. In E3 and E4, the poverty lines for both base years are specified. If there is more than one poverty line, this specification extends to the next cells, e.g. E3 to G4 in the case of Jamaica, where there are three poverty lines. To the right of these, a similar line, identified as the *reference line*, is included. Typically, this reference line will be the same as the poverty line, but in some cases it is convenient to work with different values. This is particularly the case if one wishes to experiment with the impact of a different choice of poverty line, rather different from the official value. In this case, it is convenient to maintain the *reference line* close to the official poverty line, while varying the actual poverty line. What this does is to redefine the criterion for the computation of poverty levels, while maintaining some other divisions of the sample that depend on income levels, such as the distribution of the population by income brackets further down on the Main spreadsheet, and the income brackets used in the Trends spreadsheet. The reason for this double criterion is that these brackets were designed to be used with the official poverty line and using a very different criterion may give rise to distortions. In any case, it is recommended to use a poverty line between 0.3 and 3 times the official value and to choose the reference line as a multiple of the poverty line, not too far removed from the official value.

Below the first yellow field, there is a column in yellow in cells C8 to C12 that identifies the upper limits of the **pc-group** or **pc-grupo** intervals. The need to specify this variable derives from the fact that the projection parameters vary by per capita income strata. There are a total of six such strata and their limits can be set by the user, primarily to make sure that there is some balance in the distribution of households between the different strata. Usually (but not necessarily), the first stratum corresponds to households in extreme poverty and the second to poor households above the extreme poverty line. The choice of the third and higher intervals is more arbitrary. In the application for Venezuela, shown below, the upper limits were set at 1.00, 2.00, 3.00, 5.00, and 7.50 times the reference lines, and in Brazil at 0.50, 1.00, 2.00, 3.00, and 5.00 that value. In the case of Jamaica, they were set at 0.75, 1.00, 1.50, 2.00, and 3.00 times the reference lines.

The Population spreadsheet contains the externally provided population projections by age (conventional 5-year age groups) and sex for all the years between the second base year and the projection horizon and 15 years back from the second base year. The latter is necessary for the computation of the expected number of children aged 0-9 and 10-14 under alternative projection hypotheses. The actual projection year is chosen in cell C5 of the Main spreadsheet, taking any year within this interval. Usually, the data in Population are chosen according to some officially published population projection, such as those of the National Statistical Institute, in the case of Venezuela, but it is possible to introduce fertility variations with respect to this recommended projection. To this end, one may introduce deviations in the column marked as "Deviations from Projected Fertility" (Desvíos de Fecundidad Proyectada) in the E-column of the Main spreadsheet. If these cells are set to 0, the implicit fertility and mortality rates are those of the official projection. A positive number indicates a positive proportional variation in the respective age specific fertility rate with respect to the standard projection by the end of the projection interval, e.g. 0.1 for 10% higher fertility by the year 2015 if cell C5 is equal to 2015. A negative number denotes a lower fertility than the one used in the standard projection. In the intermediate years leading up to the projection year, a linearly interpolated fertility is used. The modified population figures are displayed in lines 43-81 of the Population spreadsheet. In addition, the modified fertility rates affect the proposed average numbers of children under age 10 and between 10 and 14 in the Trends spreadsheet. Some more detail about this adjustment will be provided later. Theoretically, it would be possible also to incorporate mortality variations, but since these have only a marginal effect on the poverty projections, such an extension was not implemented. Actually, the fertility functions used in the computations of the Population spreadsheet are combined fertility-mortality functions, which reflect the joint effect of both on the number of surviving children aged 0-4 years in each projection year.

The income decomposition

The most complex and probably most important aspect of the model has to do with the way it computes the income generating capacity of each household as a result of its composition. This computation is subject to a number of variants, all of which are based on the following simple decomposition formula for the household income Y_i in year t:

$$Y_{it} = Y_{it}(\text{p-factor}) \sum_j n_{ji} * c_{iji}$$

where the sum is taken over each category j of household members, classified by their age, sex, and relationship to the head of the household. The first factor, which is identified as Y_i(p-factor), has to do with the global *economic productivity* of the household, due to determinants such as the average educational level of the productive household members, urban or rural residence, the sector of economic activity, access to land, the minimum wage and other labour legislation, etc. This factor actually explains most of the variation in per capita incomes between households, and it is this factor which is modulated by the α and β parameters of the ECLAC/IPEA/UNDP model, independently of household composition. However, from the viewpoint of the present model, the factor of greatest interest is the one associated with the *composition* and more in particular with the number of economically dependent and independent household members.

Depending on his/her age, sex, and relationship with the head of household, each member of the household contributes a number of income generating units. The average number of income generating units (c_{ijt}) per household member is defined in the model as **autonomy**; in a way, this is the opposite of the more conventional concept of dependency used in demography. Multiplying **p-factor** by **autonomy** yields the per capita income. It should be pointed out that the interpretation of these coefficients is an average which considers not only income differences between categories of individuals, but also differences in their degree of economic participation.

The differences between the variants refer to the way in which they estimate the c_{ijt} coefficients. These coefficients have not received the same attention in the literature as the more common (adult equivalent) consumption weights, most likely because there is a straightforward and apparently natural way to compute them, here referred to as the *direct* method, namely by taking the income of different kinds of household members in a reference survey[10] and determining their relative size with respect to the reference category, which (in this case) is constituted by male heads of household aged 35-49. Thus, in the case of Venezuela, a typical female spouse aged 15-24 makes 13.7% of the contribution made by a typical male head of household aged 35-49 and her coefficient is 0.136953, whereas that of the head of household aged 35-49 is, by definition, 1.000000.

Although simple to apply, the direct method for the determination of the c_{ijt} coefficients is subject to a number of refinements, which will be discussed below. The issues involved in these refinements are basically three:

1. The coefficients may vary by socioeconomic levels, particularly between the poor and the non-poor, with evident implications for the projections based on them;
2. The contributions made by household members may show up indirectly, not in their own incomes, but in the incomes of other household members; and
3. Sometimes it is not possible to derive direct estimates, due to the absence of individual income data, as in the case of Jamaica, where the LSMS is entirely based on the consumption criterion and no good alternative data are readily available.

[10] In the particular case of Venezuela, the reference survey chosen is neither the 1999 nor the 2004 round, but the household survey for the first semester of 2000. The reasons for this choice have to do entirely with the ease of access to data.

There are different ways to address the first issue. In the original methodology of Paes de Barros et al. c_{ijt} varies with i (the household identifier). This variant is here referred to as *heterogeneous direct*. However, despite this advantage, the method is difficult to apply when projections, rather than counter-factual historical scenarios, are the object of analysis. In addition, in the methodology of Paes de Barros et al., c_{ijt} does not vary with t, and this may be problematic in some circumstances. In the case of Brazil between 1999 and 2004, for example, there is a clear tendency for the c_{ijt} coefficients to become more homogeneous between male and female heads of households. This convergence may affect the projection results.

The Brazilian and Jamaican applications shown in this paper addressed the heterogeneity issue by defining four different strata, marked by the **stratum** or **estrato** variable in the Data1, Data2, and Projection spreadsheets, each of which has a different set of c_{ijt} values.[11] The Venezuelan application, which was the first to be developed, does not differentiate between strata and defines c_{ijt} purely in terms of age, sex, and relationship to the head of household. One of the variant used in the case of Venezuela, which is identified as the *homogeneous direct* method, uses weights that are fixed in terms of i, but variable in terms of t.

Both the Brazilian and the Venezuelan spreadsheets contain options for the c_{ijt} alternatives to be used in particular scenarios. In the case of Venezuela, for example, one may select the homogeneous direct method by marking "3" in the F23 cell of the Main spreadsheet. This causes the selection of the coefficients contained in cells G4 to G45 or H4 to H45 (depending on the reference years chosen in D3, D4, and D5 of the Main spreadsheet) of the Param spreadsheet for the purposes of computing the number of income generating units. Because the methodology is applied to two sets of base data (e.g. the Venezuelan household surveys of 1999 and 2004), it is possible to estimate a trend which is then extrapolated.

The disadvantage of both variants of the direct method is that they ignore the interaction between household members that may result in indirect benefits generated by one member to the extent that he or she leverages the productivity of another member. It is likely, for instance, that the contribution of children under age 10 is not 0, but negative, since they require the presence of at least one adult household member at home. The elderly mother or mother-in-law of the head of household, who does not have any income of her own, may take care of the children, so that the spouse may go out to work, thus making a positive contribution to

[11] It may not be obvious why there are two stratification variables: **pc-group** and **stratum**. The answer is that these two variables came up in different phases of the development of DMPAP: one to stratify the projection parameters (with 6 levels) and one to stratify the productivity parameters (with 4 levels). Actually, there is little reason to maintain these separate stratifications, and one might just as well use one variable, with 4 or 5 levels. In a future phase, of simplification of the model, this specification will probably be changed.

household income despite her lack of earnings. Girls aged 10-14 may fulfil a similar role in some cases. On the other hand, sickly or disabled adult members of the household may reduce the potential of other household members to earn income, as they require their constant assistance.

The (uncorrected) indirect method constitutes an attempt to incorporate some of the interactions described above, at least in part. In the case of Venezuela, it is selected by typing "1" (uncorrected) or "2" (corrected) in the F23 cell of the Main spreadsheet, thus causing the selection of the coefficients contained in cells C4 to C45 or D4 to D45, rather than G4 to G45 or H4 to H45, of the Param spreadsheet. As in the case of the direct method described above, **autonomy** is computed by taking the average of each household member's relative contribution. The method for computing C4 to C45 and D4 to D45 is rather different from the computation of G4 to G45 and H4 to H45. It is based on a regression-like procedure, with the difference that the relationship with the error term is multiplicative, rather than additive. The coefficients c_{it} (like in the homogeneous direct method, the coefficient are not specific by i) are initially unknown and have to be approximated with trial values. The criterion to determine their final value is to choose those values that minimise the variation of Y_{*} (p-factor) between households, either in terms of the coefficient of variation or the Gini index.[12] In the case of children under age 10, the formula is modified, making it non-additive. The reason for the latter is that it is plausible that the effect of the number of children on the productivity of other household members, and particularly their mother's, may not be linear. After all, the income opportunity costs to a mother represented, for example, by the care for two children under the age of 10 are not likely to increase if she has a third child.[13] For this reason, cells C27 to C31 and D27 to D31 define a function to express the relative contribution of different numbers of children. Finally, all of the coefficients C4 to C45 or D4 to D45 are determined by minimising the variation of **p-factor** (i.e. of the per capita income of the household divided by **autonomy**) between individuals, using the Solver module of EXCEL.[14]

The results obtained by using the direct or the indirect procedures are usually similar, as are the coefficients themselves. In the case of Venezuela, using the minimisation of the naturally weighted Gini indexes (with the adaptation explained in the next paragraph) as the criterion for defining C4 to C45, the correlation

[12] This is analogous to minimising the variation of the error term in conventional regression.

[13] The coefficients associated with the number of children under age 10 tend to be negative, meaning that they take away income earning potential from one of the other household members. In general, this person will be the mother. Therefore, an alternative procedure would consist in defining the relative income contribution of women with different numbers of children. In theory, it is possible for the entire **autonomy** variable to be negative in some households, but this will occur only rarely.

[14] Solver is an Add-in module of EXCEL that resides under the Tools heading. If it has not been activated, one may do so by selecting it on the Add-in sub-menu of Tools. Due to the size of the Work Book, execution may be very slow. Things are helped somewhat if one temporarily disables the Projection spreadsheet while running Solver.

between C4 to C45 and G4 to G45 is 0.814, as shown in cell C49 of the Param spreadsheet. The correlation between D4 to D45 and H4 to H45 is 0.755. As was mentioned earlier, one of the main differences between the two methods is that, in the direct approach, the coefficients of children under age 10 are, by definition, 0, whereas in the indirect approach they will generally be negative. This is *not* because young children represent net expenditures with health, education, and other needs: the criterion for the determination of c_{ik} is income earned, independently of how it is spent. Rather, the reason for the negative coefficients is that the presence of young children in the household may create obstacles for the economic participation of other household members, especially their mothers.

While the indirect method has certain conceptual advantages over the direct method, it also has some potential drawbacks. If household compositions were randomly distributed among households of different income levels (as measured by **p-factor**), the procedure described above would yield unbiased estimates of the mean number of income generating units per category of household members. However, in practice this is not the case: certain kinds of household structures are more common among the poor, others are more common among the rich.[15] Due to this correlation, it is possible, that some (inverse) causal influences of per capita income on household composition will show up as productivity effects in the coefficients C4 to C45 and D4 to D45. For example, one may argue that since poorer families tend to have more children, the child productivity coefficients C27 to C31 and D27 to D31 might come out more negative than they actually are.

It is possible, albeit laborious, to minimise this bias through a correction procedure which results in the coefficients listed in cells E4 to E45 and F4 to F45 of the Param spreadsheet. These are selected by marking "2" in cell F23 of the Main spreadsheet. The computation of these coefficients proceeds in the same manner as C4 to C45 and D4 to D45, with the difference that the household weights in Data1 and Data2 are deliberately biased to eliminate any excess number of children under age 10 and in the 10-14 age bracket among low income households, compared to higher income households. The appropriate amount of bias to this end must be estimated iteratively in cells E51 and E52 or F51 and F52. These cells must be chosen initially so that E59 or F59 become 0. After that, a set of E4 to E45 or F4 to F45 must be estimated so as to minimise E47, E48, F47, or F48, whatever the case may be. This, however, will change the value of E59 or F59, which again has to be calibrated to 0, and so on, until the process converges. Unfortunately, this is quite time consuming, but eventually an unbiased set of estimates E4 to E45 or F4 to F45 will be obtained.

[15] In the econometric literature, problems of this kind are referred to as *endogeneity biases*, for which there are a number of different solutions, such as the use of instrumental variables and two-stage least square estimation.

In some cases, it may be necessary to impose restrictions on the c_{ijt} coefficients. The most common problem is that some households may end up having negative **autonomy**. Although this will not cause the program to crash, it is certainly inconvenient. One may therefore want to impose restrictions on cells E50 and F50 of the Param spreadsheet, which report the minimum values of the **autonomies** in Data1 and Data2, to prevent this situation. But there may be other problems as well. When the indirect coefficients of different categories of household members were first estimated with the Venezuelan data, the coefficient for female spouses aged 15-24 came out negative. This negative value persisted after the correction procedure of the previous paragraph had been applied. This seemed to make little sense. Why would young female spouses represent a net cost to the household ? Upon closer inspection, the variance of the estimate turned out to be rather large and it seemed likely that female spouses in this age group were predominantly young mothers from families with below average incomes. Apart from the problem of inverse causation referred to above, the negative contribution seemed to be attributable to the number of children, rather than to the age category of the mothers per se. In order to avoid an unrealistic assessment, an additional condition was imposed on all spouses under age 65 by not allowing their coefficients to be negative. Finally, there may be situations in which the c_{ijt} coefficients associated with the number of children under age 10 progress in erratic ways. In these cases, it may be reasonable to impose some restriction in order to guarantee a more regular behaviour. The important thing to remember about the c_{ijt} coefficients is that they are important not so much because of their individual values, but rather because of their contribution to the **autonomy** variable at the level of the household as a whole.

In the case of Brazil, a different approach was used to estimating the cijt coefficients. As was mentioned earlier, one of the differences consists in the disaggregation of the coefficients by four socioeconomic levels, as defined by the variable **estrato**. The other difference is that, due to the operational difficulties of applying the indirect method (especially the corrected variant) to this more complex stratified design, a simpler solution was attempted to the problem of indirect contributions of household members to the household budget which are not reflected in their own incomes. The modified direct coefficients of the Brazilian variant were determined in a way similar to the homogeneous direct weights above, including disaggregation into four per capita income categories, but with the difference that only households without children under age 15 and dependent elderly persons (i.e. who were not heads of households or spouses) were taken into consideration. It was then assumed that the presence of children under age 10 takes away a certain amount of income generating capacity from spouses aged 15-34, whereas the presence of elderly persons, particularly elderly women, restores part of this capacity. Based on earlier empirical observations, it was also assumed that the largest negative effect is reached with 3 children; when there are more than 3, their hampering effect on income generation diminishes somewhat as the older ones start to take care of the younger ones. The user has to specify (in the F23 cell of the Main spreadsheet) whether the scenario should be prepared with this option or with the simpler homogeneous direct method explained above.

In the case of the Jamaican application, as it was implemented so far, the user has no option in the determination of the c_{iit} coefficients which are estimated through an ad hoc procedure which combines elements of the indirect procedure used in Venezuela and the modified direct approach used in Brazil. The details will be discussed in the actual application at the end of this document.

The household projections

One of the more difficult aspects of the projection methodology to decide on was its time-age structure; by period or by cohort. From a demographic viewpoint, there are major advantages to a cohort projection, which makes it possible to compare the same individuals at different points in time and to obtain estimates of the transformations occurring during the interval in between. However, the object of analysis of this model consists of households, identified in terms of the age and sex of their heads, not individuals. The age and sex of heads of households has a much more erratic behaviour than that of individuals. Individuals can die or migrate, but otherwise all of them will still be present 10 years after, having aged 10 years. In the case of households, however, new units are constantly formed, old ones are dissolved, and changes in headship occur frequently. Young households (with heads under age 35) are particularly unstable. In addition, any cohort methodology would have to be based not on true cohorts, but on pseudo-cohorts because the data were obtained from distinct samples. In practice, this makes it very difficult to track household changes over time. Due to these difficulties, it was concluded that the possibility of comparing the same household cohorts at different moments in time was more hypothetical than real. Instead, the *population* was projected in conventional cohort terms, whereas the *household structure* was projected "horizontally", i.e. maintaining the same households with adapted headship rates.

In practice, the projection process works as follows. The households in the Projection spreadsheet are the same as those in the Data2 spreadsheet, but transformed in terms of the following characteristics:

- The household weights change proportionally to the number of individuals in the respective age and sex categories and due to changes in headship rates. For example, if a given household has a male head aged 25-34 and the size of the population in the latter age-sex category increases by 15%, so does the household weight. If, in addition, the headship rate in this category falls by 5%, this reduction is also applied to the household weight.

- The number of other household members in each category of age and sex changes in ways which can partly be specified by the user, under the constraint that these changes have to be consistent with the over-all age-sex structure of the population projected in the Population spreadsheet. Similarly, the number of adults of a given age and sex who are spouses of the household head can also be specified by the user, subject to the condition that the household head can have a maximum of one spouse, of the opposite sex. If not, the programme automatically makes the necessary adjustments in the Projection spreadsheet.
- The variable **p-factor** changes in accordance with the α and β parameters specified in the Main spreadsheet. This result, combined with the changes in the numbers of different kinds of household members, who are characterised by different contributions (c_{ij}) to the household budget, determines the total household income and per capita income. The projections of monetary values are made in real terms, using the monetary valuations of the second base year. Hence the poverty and reference lines are assumed to be the same as in the second base year.

In order to prepare actual projections, the following elements have to be considered:

1. The projection year, as specified in cell C5 of the Main spreadsheet, anywhere between the second base year and the projection horizon.
2. Any fertility variations with respect to the recommended population projections that one might want to introduce by choosing the Deviations from Projected Fertility in the Main spreadsheet different from 0.
3. The corrected projection values for the rates and averages in the yellow lines of the Trends spreadsheet, which can be set with the help of a number of auxiliary parameters in the top lines of the same spreadsheet.
4. The calibration procedure, to enforce consistency with the population projections.

Points 1 and 2 have already been discussed previously. In the remainder of this section, points 3 and 4 will be elaborated on, starting with 3. Because household composition is the key to poverty reduction in this model, the choice of the household composition parameters is treated in great detail, in the Trends spreadsheet, which provides the necessary inputs for the Projection module. It contains preliminary projections of these parameters, subject to correction by the user, for the following quantities:

headship (jefatura) – Headship rate;
num09 – Average number of children under age 10 in the household;
num1014m – Average number of males aged 10-14 in the household;

num1014f – Average number of females aged 10-14 in the household;
num1524m – Average number of males aged 15-24 in the household;
num1524f – Average number of females aged 15-24 in the household;
num2534m – Average number of males aged 25-34 in the household;
num2534f – Average number of females aged 25-34 in the household;
num3549m – Average number of males aged 35-49 in the household;
num3549f – Average number of females aged 35-49 in the household;
num5064m – Average number of males aged 50-64 in the household;
num5064f – Average number of females aged 50-64 in the household;
num6574m – Average number of males aged 65-74 in the household;
num6574f – Average number of females aged 65-74 in the household;
num75m – Average number of males aged 75 or more in the household;
num75f – Average number of females aged 75 or more in the household;
numc1524 – Spouseship rate in the 15-24 age category;
numc2534 – Spouseship rate in the 25-34 age category;
numc3549 – Spouseship rate in the 35-49 age category;
numc5064 – Spouseship rate in the 50-64 age category;
numc6574 – Spouseship rate in the 65-74 age category;
numc75 – Spouseship rate in the 75+ age category.

These rates and averages are specified by age and sex categories of the heads of households and by household income class in terms of **percap**. For each of the six categories identified in the Main spreadsheet, the Trends spreadsheet first provides the values found in the two base years, based on Data1 and Data2, then a preliminary Proposal (Propuesta) for the projection year, which the user can correct in the yellow line marked as Correction (Corrección), and finally the value actually used in the projection. The easiest way to complete the spreadsheet is simply to copy all proposed values to the Correction line immediately below it. However, in some cases the proposed values may be implausible and it is good to be alert to this situation. The spreadsheet itself already signals some values that may be in need of correction by marking in red all of the cells in the Proposal line that differ from those of the second base year by more than 0.5 absolute units or that are less than 0.5 or more than 2.0 times its size.[16]

To allow a lot of flexibility, e.g. to make it possible to model different trends for different income strata, the entire projection process of the variables above is fairly elaborate. The formula that governs the proposed projection values of the Trends spreadsheet is the following:

[16] These maximum deviation parameters can also be adjusted, if desired.

$$r_s^{(P)} = A\,(B\,R^{(2)} + (1-B)\,r_s^{(2)})\,(r_s^{(2)} / r_s^{(1)})^C$$

where $r_s^{(P)}$ is the proposed projection of the rate or average for the stratum s, $r_s^{(1)}$ and $r_s^{(2)}$ are the corresponding rates or averages for base year 1 and 2, respectively, and $R^{(2)}$ is the mean of $r_s^{(2)}$ for a given age-sex category of heads of household, i.e. taking the average over the income classes. A, B, and C are the three relevant parameters. A is the calibration parameter which is adjusted in the Param spreadsheet to make sure that the number of persons in each age and sex category is consistent with the population projection in the Population spreadsheet. In the simplest application, B and C are simply zero, so that all averages of numbers of household members are adjusted proportionally. However, one has the option of modulating this procedure through the B and C parameters. B is the homogeneity parameter. By choosing it greater than 0 and smaller than 1, one introduces a degree of convergence between income classes into the projection. If B = 1, all income classes will have the same projected rate or average, which will continue to vary only by the age and sex of the head of the household. C is the trend sensitivity parameter. It expresses to what extent the projection will follow past trends. In theory, it may take on any value. Large positive values imply an exacerbation of past trends, whereas a value of 0 implies that past trends will not be considered. Because the extrapolation of past trends may yield some rather implausible numbers if the ratio $r_s^{(2)} / r_s^{(1)}$ is very different from 1, the Maximum Deviance parameter in line 7 of the Trends spreadsheet allows the user to impose a limit on this ratio, e.g. if the parameter is chosen equal to 2, all ratios will be truncated at 2 or ½. All four parameters (A, B, C, and the Maximum Deviance) are specific to the particular rate or average being projected.

In some cases the user may want to fix a projected rate or average in the Trends spreadsheet, without having it adjusted afterwards by the calibration procedure. In the case of averages (the rates are not subject to calibration), one may do this by typing in a counter-proposal on the yellow line of the Trends spreadsheet as a negative value, thus signalling to the calibration procedure to leave this particular value alone.[17] Otherwise, all initial parameters are subject to adjustment. The calibration parameters A are specified not in the Trends spreadsheet, but in the Param spreadsheet. As was mentioned before, these parameters are the ones that ultimately guarantee consistency of the projected averages with the population projections in

[17] This cannot be done for *all* rates and values, for it would make the model inherently inconsistent. The negative sign only signals that the particular value is fixed; in the actual projection, it will be omitted. Note, however, that fixing an average in the Trends spreadsheet only exempts this value from the calibration procedure, but not from other kinds of necessary adjustments. It is still possible that the final values used in the projection will be slightly different from those specified by the user due to the truncation that the projection procedure has to apply in order to avoid inconsistencies within households, like negative numbers of household members or numbers inconsistent with the age and sex of the head of household. The projection module also adjusts the number of spouses if the sum (over all age groups) exceeds 1 and it adjusts the number of children to be consistent with the maximum completed family sizes specified in the H-column of the Main spreadsheet.

the Population spreadsheet. After the model has been calibrated, the values in the Criterion column of the Param spreadsheet, which are repeated in cells F8 to T8 of the Trends spreadsheet, should all be 0. When this is the case, the last line of each age-sex-income block displays the values actually used in the projection.

Another set of values that influences the actual projection and that may interfere with the specifications for the 0-9 and 10-14 age groups made in the previous paragraphs is the Maximum Final Descendancy (Completed Family Size) in the H-column of the Main spreadsheet. What these parameters do is to specify a maximum target for completed family size by household income level (identified in terms of **percap**). According to this specification, women in age groups 15-24, 25-34, 35-49, and 50-64 should not have more than a certain number of children under age 15, if their projected completed family size is to be consistent with this target. For example, a woman aged 25-34 who already has 3 children is unlikely to comply with the target of 3.25 surviving children by the end of her reproductive life. In this case, the projection module truncates her number of children at the maximum number (2.5, for instance) which would be compatible with the stated target. If no truncation is wanted, the Maximum Final Descendancy cells in the H-column of the Main spreadsheet should be set at a high value, e.g. 25 or 100. In a somewhat crude fashion, these parameters may be used to model the impact of compliance with desired fertility levels. The main limitation of the procedure in this case is that it has to assume that the desired completed family size is the same for all women belonging to a particular income class. Since actual fertility preferences vary, this means that some large family sizes, which may be the result of individual preferences, will be unduly truncated. The other limitation has to do with children who do not live with their mothers. If a household has no female member in the 15-64 age category, the procedure is applied to the male members of the household.

The α and β parameters also exert an indirect influence on the resulting rates and averages because they may shift population segments from one income bracket to another. This may confound the interpretation of the projected rates and averages read from the last line of each block. In order to make the procedure more transparent, it is therefore recommended to define the rates and averages in the Trends spreadsheet with α and β both set to 0. After the results have been calibrated, the projected values of α and β may be introduced without affecting the calibration.

Finally, it should be clarified that not all of the rates and averages are treated in the same manner. The rates in the column marked as **headship (jefatura)** are conventional headship rates, specific by age, sex, and income stratum, which can be extrapolated without any need for consistency in terms of the total population. Line

4 of the Trends spreadsheet does allow for the specification of Calibration parameters for this and other rates, but these can be chosen arbitrarily by the user, without need for adjustment, so while the formula by which the projection proposal is obtained is basically the same as in the case of the averages, the interpretation of the parameter is slightly different. The denominator of the headship rate is the number of persons of the correct age and sex that belong to the specified income stratum. The extrapolation procedure whose result is displayed in Proposal automatically truncates any possible negative values or values larger than 1.

The spouseship rates in the columns marked as **numc1524** to **numc75** indicate what percentage of individuals in the eligible age and sex categories that live in the household are spouses. Thus, if there are two women aged 25-34 in a given household and one of them is the spouse of the (male) head of household, the spouseship rate for the 25-34 age group is 0.5. In the same manner as the headship rates, these rates have to be located in the interval between 0 and 1, and they do not require consistency with the population projection, although they do require consistency of projections at the household level: if the projection suggests more than one spouse of the head of household, the Projection module will correct this situation.

The remaining columns refer to average numbers of household members in specified age and sex groups. These have to be non-negative and – in the case of the column corresponding to the age and sex of the head of household – at least 1.0. The Proposal line automatically truncates any values that do not comply. However, after the specification of the corrected value in Correction, these columns still have to go through a calibration process to guarantee their consistency with the corresponding populations by age and sex in the Population spreadsheet.

Calibration

As was mentioned previously, the calibration is carried out in the Param spreadsheet. Note also that the calibration of the 0-9 and 10-14 age categories is slightly different from the others because these age groups are automatically adjusted for changes in fertility declared in the Population spreadsheet. Therefore, the calibration in this case only covers the residual adjustment that has not been adequately covered by the latter procedure. After all the rates and averages in the Trends spreadsheet have been defined, one uses the Solver procedure of EXCEL to choose the parameters in cells L31 to L45 so that the entire Criterion column next to it becomes 0.[18] This causes the Projection spreadsheet to modify the chosen averages of the Trends spreadsheet in such a way that the number of individuals in each age

[18] When performing this procedure, one should apply Solver from L45 upward to L31, rather than starting at L31. This is because the adjustment of the adult population, particularly the female population, will affect the adjustment of the population under age 15.

and sex category sums to the correct total. Any cells that the user has marked with a negative sign in the Correction lines of the Trends spreadsheet will be unaffected by this procedure. Each time changes are made in either the Population spreadsheet or the Trends spreadsheet, this procedure has to be repeated. Unfortunately, due to the size of the Work Book, this can be time consuming. If the projection is not properly calibrated this will show up as an inconsistency between the cells H11 and H12 of the Main spreadsheet; in this case, H12 will be displayed in red.

An example with Venezuelan data

Venezuela was the first country to which the DMPAP methodology was applied and consequently this application was subject to some experimentation with aspects of the model which were later resolved more permanently and more satisfactorily. The changes have to do especially with the definition of the income generating units (c_{ij}). In the case of country surveys that contain individual income data, the most satisfactory solution so far was the one implemented in the case of Brazil, i.e. the second application. In surveys that do not have individual income data, such as the one used in Jamaica, alternative solutions have to be found which, in this particular case, came down to the adoption of the c_{ij} values used in Brazil, adjusted by an algorithm of the kind used in the present application to Venezuela.

The data for the Venezuelan application come from the national household surveys for the first semesters of 1999 and 2004. In order to be compatible with the historical time series and population projections of the National Statistical Institute, adjustments were applied to the age and sex structures of the populations surveyed in both years. This procedure is not exclusive to the case of Venezuela; it was applied in the same way to Brazil and Jamaica. As a consequence of these adjustments, the base line statistics for 1999 and 2004 are slightly different from the official figures. This is not a major drawback as long as the objective of the analysis is to demonstrate the relative impact of population factors on poverty reduction, rather than measuring poverty as such.

In cells C3 and C4 of the Main spreadsheet, the two base years are entered: 1999 and 2004. In cells E3 and E4, one enters the chosen poverty lines. In this case, the official poverty lines for each year were used: B$ 48,628.8 for 1999 and B$ 123,880 for 2004. The reference lines in F3 and F4 were set at the same values as E3 and E4. With this information and the contents of the spreadsheets Data1, Data2, and Projection, the programme produces a series of summary indicators in the first lines of the Main spreadsheet, in the same way as in the later spreadsheet versions of Brazil and Jamaica.

The number of cases in Data1 is 2930 and in Data2 3710. This is substantially less than the number of households interviewed in the two surveys, due to the fact that household information in Venezuela and Brazil was clustered. The nominal average income at the household level increased from B$ 322,109 in 1999 to B$ 594,870 in 2004. However, due to the inflation, which is reflected in the increase of the poverty lines, the incidence of poverty at the household level actually went up, from 43.26% to 52.74%. At the level of individuals, shown below the household data, the average per capita income increased from B$ 69,760 to B$ 141,152 per capita, and poverty from 51.58% to 60.49%. The Gini index, also at the individual level, was reduced slightly, from 0.4672 to 0.4527.[19] The data are also shown separately for male and female headed households. This reveals a higher incidence of poverty among individuals living in female headed households: 57.00% in 1999 and 65.75% in 2004, compared to 49.55% and 58.14% for male headed households.

In addition to the incidence of poverty at the individual level (the head count index or P_0, in the terminology of Foster, Greer, and Thorbecke)[20], the first lines of the Main spreadsheet also display the poverty gap index (brecha de pobreza, or P_1) and the severity index (P_2). The poverty gap index, which was 22.28% in 1999 and 28.35% in 2004, is interesting in that, multiplied by the poverty line and the total population size, it yields the total amount of resources that would have to be transferred from the non-poor to the poor in order to eliminate poverty. In the case of Venezuela in 2004, this would have been a total of B$ 917,487 million, or 24.88% of the total personal income earned.

The summary of population data indicates that the number of households increased from 5,168,310 in 1999 to 6,199,556 in 2004, while the number of household members increased from 23,864,237 to 26,127,351. There was also an increase in the proportion of households headed by women. The number of income generating units (using 2004 income generating coefficients), here shown in cells P24 and P25, increased from 7,094,264 to 8,349,887, i.e. a 17.7% increase, whereas the number of individuals increased by only 9.5%. All else being equal, this suggests an increase of productivity per person of about 7.5%. Lines 14-20 of the Venezuelan spreadsheet[21] display the distribution of individuals between age and sex groups and categories of relationships with the household head. The main

[19] The Gini index may be biased slightly downwards due to the clustering of households. In the case of Venezuela, the actual Gini indices for per capita incomes based on the entire household sample were 0.4631 in 1999 and 0.4613 in 2004, compared to the values of 0.4624 and 0.4603 found in the clustered sample. These differences are small enough to be ignored.

[20] See J. Foster; E. Greer & E. Thorbecke. 1984. "A class of decomposable poverty measures." *Econometrica* 52.

[21] Because some elements of the analysis (e.g. multiple poverty lines) were implemented in some versions of the model and not in others, the exact location of the lines may vary between country applications. In a later version of the model it would be desirable to unify all the formats, but for the time being it is more important to incorporate the adjustments necessary to accommodate the particularities of each country application.

differences refer to the increase of the percentages of heads of households (from 21.66% to 23.73%) and the decline of the percentage of children under age 10 (from 22.97% to 21.31%). The percentage of spouses also increased in most age categories, as did the percentage of elderly.

In order to appreciate the impact of these changes on household incomes, one should inspect the columns C through H of the Param spreadsheet. In the case of Venezuela, the differences between the indirectly determined (corrected) coefficients c_{jt} of household heads of different ages and sexes are relatively minor, with the exception of women over age 65, who are characterised by lower coefficients: 0.834 for the 65-74 age group, and 0.655 for the group over the age of 75 in 2004. Very young household heads of both sexes also have a low coefficient: 0.789 for males and 0.718 for females in the 15-24 age group. The directly determined coefficients (columns G and H) are more varied, as are the coefficients that characterise spouses. As was explained in the previous section, the 0 values for female spouses aged 15-24 were actually imposed in the minimisation procedure, in order to avoid negative coefficients for adults under the age of 65. Not too much significance should be attributed to this value and it certainly should not be interpreted as an endorsement of any ideological position with respect to the intrinsic value of young wives or any other categories of household members for the social reproduction process. All it means is that – statistically speaking – households with female spouses in this age category tend to be poorer than households with similar compositions in which the spouses are older. As was explained in the previous section, there is some collinearity between the coefficients c_{jt} and their individual variances are relatively large. More important than their individual values is the question whether, on the whole, the **autonomy** of the households, estimated through the c_{jt} set, provides an adequate representation of their income generating capacities.

As expected, the coefficients for children under age 10 are negative and increase with the number of children. The exception is the case of families with 5 or more children, where the coefficient is less negative than in the case of 4 children and even becomes positive in 2004. It would seem that some sort of economy of scale is at work here. The coefficients for children in the 10-14 age group are also negative, with only a very small difference between girls and boys. This last finding does not give support to the idea, advanced earlier, that girls in this age group contribute significantly to the household economy by looking after their younger siblings. The largest contribution of income generating units among the other household members over age 15 who are not heads or spouses of the heads is found in the case of women aged 25-49 years. The contributions of other household members over the age of 65 are mostly negative, but some are not. Again, the hypothesis raised earlier, about the indirect economic contribution of elderly women, is not borne out by the estimates. The directly determined coefficients for other adult household members are generally somewhat larger than the indirectly determined ones, with a maximum of 0.641 in the case of men aged 35-49.

If the projection year in cell C5 of the Main spreadsheet is set to the same year as the second base survey (i.e. 2004, in this case), and the α and β parameters in N37 and N38 are both 0, the projected values should, in principle, be equal to those of 2004. If they are not, this is either because the parameters in cells H31 to H45 of the Param spreadsheet are not zero, because the corrected projection values in the yellow lines of the Trends spreadsheet are not equal to the observed values in 2004, or because the Maximum Completed Family Size parameters in column H have been set too low. Any of these will change the household composition and therefore the projected base indicators. The way to correct the latter is to set all the projection values in the yellow lines equal to the proposed values immediately above them and to set the Homogeneity and Trend Sensitivity parameters in cells H5 to V6 of the Trends spreadsheet equal to zero. The Consistency Verification cell (H12), which should be compared with the projected population (H11), allows the user to verify if these conditions are satisfied.

Comparing the 1999 and 2004 results in terms of the α and β parameters of the ECLAC/IPEA/UNDP model for per capita income, without considering household composition effects, one finds a value of 0.0309 for α and 1.0234 for β (cells J37 and J38). The first value suggests some reduction of inequality, whereas the second does not mean much, because it is based on nominal income values, without inflation correction. However, for the purposes of the analysis which follows this is not of any major importance. What is relevant is that the α and β for the p-factor of individuals, which eliminates the effect of household composition on household income, are 0.0323 and 0.8812 (cells K37 and K38) if corrected indirectly determined weights for the economic contributions of different types of household members are used.[22] This shows that, if household composition is controlled for, and that about 15% of the apparent increase in average nominal incomes per capita can be accounted for by this very change in the composition of households. If it weren't for this effect, the percentage of poor individuals in 2004 would have been 64.50% instead of 60.49%. This can be verified by taking the two parameters listed under "Difference" (cells L37 and L38), substituting them in the yellow cells N37 and N38, and reading the resulting poverty projection from cell M11 or from the per capita income projection

[22] By typing a "3" in cell F23, one obtains the estimates for the case of weights directly obtained from the income data of the two household surveys, which are 0.0211 and 0.9431, respectively. In this case, the number of income generating units would increase less, from 7,716,334 to 8,879,528, a 15.1% rise.

column M41 to M46.[23] The latter can be compared to column J41 to J46, showing, for example, that, in relative terms, the effect of household composition on the percentage of extremely poor (less than half the poverty line) would be slightly greater than the overall effect: 27.46% (with change in household composition) instead of 30.52% (without change).

In the Venezuelan case, the ECLAC/IPEA/UNDP model reproduces the incidence of extreme poverty (under 0.5 times the reference line) in 2004 quite well: 27.46% predicted on the basis of the values of α and β, against an observed 28.08%. At the other extreme (incomes larger than 3.0 times the reference line), the percentages are also quite similar: 5.11% predicted, compared to an observed 5.62%. In the case of incomes under 0.3 times the reference line the fit is less satisfactory (10.75% predicted, against an observed 12.61%), whereas the general poverty index is also predicted less precisely: 62.15%, against an observed 60.49%. Apparently, there is no tendency for the model to overstate the poverty reduction effect.

A conventional projection, without effects of the population structure

This concludes the visual inspection of trends between the two base years and sets the stage for the actual projections. For this purpose, it is assumed that from 2004 until 2015 Venezuela will maintain an economic growth rate of 3.0% per year, at first without any reduction of income inequality. Discounting expected population growth during the period, as implied by the population projections, this implies that $\alpha = 0$ and $\beta = 0.1660$, values which are substituted in the yellow field of cells N37 and N38. The projection year in cell C5 is maintained at 2004, meaning that population changes are not considered at first. With these classical settings of the ECLAC/IPEA/UNDP model, the predicted reduction of individual poverty is from 60.49% to 52.52%, as can be read from cell M11. Assuming that, in addition, there would be a 7.5% reduction in income inequality ($\alpha = 0.075$), poverty would go down slightly further, to 51.94%.[24] The reason for this small impact of α in the case of Venezuela is that the (official) poverty line specified in cell E4 is very close to the mean income in cell K10, so that a contraction of the distribution in the direction of the mean shifts few people across the poverty line. Extreme poverty (under half

[23] For this to work, it is necessary that the projection year in cell C5 be equal to the second base year in cell C4 and that the other conditions mentioned earlier be satisfied, so that the populations in cells H11 and H12 coincide.

[24] A difference between the "standard" application of the ECLAC/IPEA/UNDP model and the implementation in DMPAP is that in DMPAP the transformations governed by the α and β parameters refer to p-factor and not directly to per capita incomes. Therefore, the combination of α and β values that brings about a certain combination of per capita income growth and reduction of per capita inequality does not correspond directly to the amounts by which one wants to increase per capita incomes and reduce inequality. To formulate hypotheses in terms of per capita incomes, it is necessary to first encounter the correct α and β values that will bring about the desired changes. In the case of Brazil and Jamaica, the hypotheses were actually formulated this way, but in the Venezuelan application the hypotheses simply refer to transformations of p-factor, without any systematic consideration of their impact on per capita incomes.

the poverty line) would fall from 28.08% (or 27.46%, according to the projected distribution) to 19.81%.

The Main spreadsheet provides some additional statistics which may be interesting in this context, such as β values for the Poor and the Non-Poor, i.e. the relative increase of incomes in these two groups. If $\alpha = 0$, this value should be equal to the overall β in both groups, but if $\alpha > 0$, incomes among the Poor should grow more than among the Non-Poor. If $\alpha = 0.075$, one finds a value of 0.2259 for the Poor and 0.1341 for the Non-Poor. Another interesting statistic is the proportion of the *growth* of total incomes that accrues to the Poor and the Non-Poor, in this case 0.3982 to the Poor and 0.6018 to the Non-Poor. That the latter value is greater than the former is because β is applied to a larger initial income in the case of the Non-Poor than in the case of the Poor. Sometimes it is preferable to formulate hypotheses in terms of these parameters than in terms of α. For instance, suppose that one would want to elaborate a scenario based on the premise that 50% of all income growth should benefit the Poor. In that case, one should choose α equal to 0.1365, at least for the purpose of the present scenario, without changes in the population structure. This implies that the incomes of the Poor would grow by 27.51% and those of the Non-Poor by 10.81%. If β is changed or if changes in the population structure are introduced, as below, this value of α may have to be adjusted.

Now it is time to introduce changes in the population structure, setting the projection year in cell C5 to 2015. At first, this will cause an imbalance in the values of cells H11 and H12 (causing the latter to show up in red) which has to be adjusted through calibration. The Homogeneity and Trend Sensitivity parameters of the Trends spreadsheet are initially set to 0. This means that the initial projected values of all rates and averages in the Trends spreadsheet are proportional to their 2004 observations. Here the c_{ijt} are the corrected indirect coefficients for 2004. While the model is being recalibrated, it is recommended to keep α and β set at 0 at first. With these settings and after calibration, the following results can be observed:

- Individual poverty falls from 60.49% to 53.37%;
- The poverty gap (P_1) falls from 28.35% to 23.63%;
- The Gini index at the individual level falls from 0.4527 to 0.4407.

The base projection

When the previously chosen values of α and β are reintroduced,[25] one obtains what will be called the base projection, with the following results:

[25] With α and β set at 0, the headship and spouseship rates, which are not affected by the calibration procedure, will remain fixed at the values assigned by the user. However, once different values of α and β are introduced, the values of these rates between income categories will change, although the total for each age-sex category of the household head will remain the same. The reason is that the change in α and β moves households between income categories, which will affect all rates and averages defined in terms of these categories, including the headship and spouseship rates.

- Individual poverty falls from 60.49% to 44.62%;
- The poverty gap (P_1) falls from 28.35% to 17.31%;
- The Gini index at the individual level falls from 0.4527 to 0.4175.

This means as much as half of the entire poverty reduction expected between 2004 and 2015 is attributable to the population effect. Much of the population effect is related to the increase of the percentage of household heads, from 23.73% to 26.86%, and the reduction of the percentage of children under age 10, from 21.31% to 18.73%. The population over age 65 also increases substantially, from 4.91% to 6.70%.

If the uncorrected, rather than the corrected c_{ijt} coefficients are chosen in cell F23, it accelerates the poverty reduction effect. In this case

- Individual poverty falls from 60.49% to 43.62%, rather than 44.62%;
- The poverty gap (P_1) falls from 28.35% to 16.89%, rather than 17.31%;
- The Gini index at the individual level falls from 0.4527 to 0.4164, rather than 0.4175.

If, on the other hand, the direct, rather than the indirect specification is chosen in cell F23, it dampens the poverty reduction effect considerably. In this case

- Individual poverty falls from 60.49% to 46.99%, rather than 44.62%;
- The poverty gap (P_1) falls from 28.35% to 18.55%, rather than 17.31%;
- The Gini index at the individual level falls from 0.4527 to 0.4212, rather than 0.4175.

This shows that the difference between both approaches in terms of their poverty reduction effect may be substantial. The difference introduced by using corrected, rather than uncorrected coefficients, is less important, but points in the expected direction. A final variation that was tried is one that uses the 1999, rather than the 2004 corrected indirect coefficients. This results in a 2015 poverty estimate of 43.95%. The relevance of this finding is that the difference between this finding and the 44.62% found with the 2004 coefficients may express a trend in the evolution of the c_{ijt} coefficients over time. If this were the case, one might expect, based on the extrapolation of this trend, an incidence of poverty in the order of 46.09% in 2015. Apparently, this would bring the estimate close to the direct variant, but remember that the direct variant may be affected by a similar trend, which would move it upward as well. The latter trend has not been explored here.

Translating the differences between approaches into equivalent rates of additional economic growth, the results would be as follows. If population factors are not considered (i.e. if the projection year is kept at 2004, and α still at 0.075), the β needed to bring about the same poverty reduction as under the direct specification of income generating units would be 0.2870, whereas under the corrected indirect specification with 2004 coefficients it would be 0.3515. This corresponds to annual growth rates of 3.93% and 4.39%, respectively, compared to the 3.0% assumed at the beginning. Of course, as was alerted to in the Introduction, to the extent that this substantial increase of implicit economic growth rates involves an expansion of the economically active population, there is some question as to whether the base rate of economic growth, of 3.0%, will actually be sustainable under these circumstances.

A projection with convergent rates and averages

In their study on Brazil, Paes de Barros et al. found that differences in household composition between families at different income levels did not explain much of their poverty differential. To check whether this is also true in Venezuela, a second projection scenario can be prepared in which the initial rates and averages are not equal to their 2004 observed values within each income class, but to the overall averages for households with heads of the same age and sex, so that by 2015 household compositions would be uniform by income categories, varying only by age and sex of the household head. Headship and spouseship rates are also made uniform. This is done by choosing all the Homogeneity parameters G5 to AB5 in the Trends spreadsheet equal to 1. The new parameters require some recalibration. After recalibration, with cell F23 marked with a "2", the following changes with respect to the previous scenario emerge:

- Individual poverty falls further, from 44.62% to 42.07%;
- The poverty gap (P_1) falls further, from 17.31% to 14.80%;
- The Gini index at the individual level falls further, from 0.4175 to 0.3826.

The changes in the incidence of poverty are substantial, particularly the change in the Gini index. The latter finding is important in that it expresses the impact of demographic equality on economic equality, which must be summed with the impact of fertility reduction already implied by the base projection. The projection may be somewhat unrealistic, however, in that the total convergence of all household structures is not easily achieved by any public policy intervention and contains a substantial effect of inertia due to processes inherited from the past. A more limited convergence, which falls within the domain of conceivable policy interventions, is the one that only homogenises the numbers of children under age 10 in each household by 2015, while all other rates and averages are projected as earlier. Under this scenario, poverty is reduced from 44.62% to 44.20% (rather than 42.07%), with the poverty gap falling to 16.34% and the Gini index to 0.4075. Note that in this case the projected average numbers of children aged 0-9 in different kinds of households will no longer be uniform due to differences in the numbers of women of reproductive age. What is made homogeneous in this case is not the average number of children, but a child-woman ratio specific by age of the woman. If, in addition to the number of children under age 10, headships rates converge between income strata, poverty falls to 43.70%.

A projection with faster fertility decline

Another alternative scenario might be one in which fertility falls more than under the standard projections, e.g. another 10% by 2015. This scenario is obtained by typing -0.1 in cells E28 to E34. It also requires a revision of β, since the lower population growth implies a larger growth of income per capita. The revised value of β becomes 0.1782, instead of 0.1660. Because the modification of the fertility scenario also affects the population in the 0-14 age group, it is necessary to recalibrate the first three entries of the L column of the Param spreadsheet. With these new parameters, the reduction of poverty with respect to the base scenario is characterised by the following indicators:

- Individual poverty falls further, from 44.62% to 43.28%;
- The poverty gap (P_1) falls further, from 17.31% to 16.57%;
- The Gini index at the individual level falls slightly further, from 0.4175 to 0.4157.

A slightly different scenario is obtained if the fertility reduction, instead of being homogeneous in all age groups, is concentrated in the 15-19 age group. To this end, one should choose the first fertility modification cell equal to - 0.52 and the remainder equal to 0. This results in the same population growth between 2004 and 2015 as under the previous variant, so the value of β (0.1782) can be maintained. Under this scenario, the following results are obtained:

- Individual poverty falls from 44.62% to 43.27%, instead of 43.28%;
- The poverty gap (P_1) falls from 17.31% to 16.53%, instead of 16.57%;
- The Gini index at the individual level falls from 0.4175 to 0.4152, instead of 0.4157.

These differences are surprisingly small. However, before concluding that poverty does not respond to the reduction of adolescent fertility any more than it does to the reduction of general fertility, one should remember the observation made in the last paragraph of the introduction. The DMPAP model analyses primarily how demographic change brings about changes in the level of poverty through its impact on household structure. In the case of adolescent fertility, however, it is reasonable to suppose that there are other, more immediate pathways through which poverty might be reduced and which are not adequately accounted for in the model.

A projection with constant fertility

Another alternative scenario might be one in which fertility remains constant at its 2004 level. This scenario is obtained by typing 0.145 in cells E28 to E34. The required revision of β yields a value of 0.1488, instead of 0.1660. Again, it is necessary to recalibrate the first three entries of the L column of the Param spreadsheet. With these new parameters, the reduction of poverty with respect to the original scenario is characterised by the following indicators:

- Individual poverty falls less, i.e. not to 44.62% but to 46.60%;
- The poverty gap (P_1) falls less, i.e. not to 17.31% but to 18.39%;
- The Gini index at the individual level falls slightly less, i.e. not to 0.4152 but to 0.4202.

Table 1 provides an overview of the different projections that were carried out in this chapter. The main conclusions of this overview may be summarised as follows:

1. The potential contribution of demographic trends in Venezuela to the reduction of poverty up to 2015 is very substantial, corresponding to an equivalent additional growth rate of the per capita GDP in the order of 1-2%, with an incidence of poverty in 2015 that could be up to 10% lower than the projected figures in the absence of population composition effects.
2. Although the results vary depending on the specifics of the methodology used, all methodologies point at a substantial demographic effect.
3. Most of this contribution is already implicit in demographic changes that took place previous to 2004, as the inertial effect of the fall of fertility in the past.
4. Nevertheless, not everything is determined by the past. Depending on the characteristics of the changes in fertility and household composition from 2004 until 2015, poverty in 2015 may be as low as 42% or as high as 49%.
5. The speed of the fertility decline between 2004 and 2015 is a less important factor than the distribution of different household compositions between social strata.

Table 1: Overview of the different projection scenarios and their results for Venezuela

	Poverty (P_0)	Poverty gap (P_1)	Gini index
Situation in 2004	60.49%	28.35%	0.4527
Projections to 2015 based on 3.0% growth			
Without considering population composition	51.94%	21.81%	0.4303
Base projection with direct coefficients	46.99%	18.55%	0.4212
Base projection with corrected indirect coefficients (2004)	44.62%	17.31%	0.4175
Base projection with uncorrected indirect coefficients (2004)	43.62%	16.89%	0.4164
Base projection with corrected indirect coefficients (1999)	43.95%	16.92%	0.4164
Convergent rates and averages (with 2004 corrected indirect coefficients)	42.07%	14.80%	0.3826
Only convergence in 0-9 age group	44.20%	16.34%	0.4075
Convergence of 0-9 age group and headship	43.70%	16.07%	0.4039
Additional 10% of fertility decline	43.28%	16.57%	0.4157
Same decline concentrated in 15-19 age group	43.21%	16.90%	0.4159
Constant fertility	46.60%	18.39%	0.4202

An example with Brazilian data

The main difference between the Venezuelan and the Brazilian application is that the c_{ijt} coefficients in Brazil were disaggregated by per capita income levels and determined in a more direct way than in the case of Venezuela. The two options provided to the user in cell F23 of the Main spreadsheet are to use either modified direct coefficient (option 1) or simply the standard direct coefficients (option 2), that were marked as the third option in the Venezuelan application. This procedure for the specification of the c_{ijt} values is now the preferred option in countries where individual income data are available. In the case of Jamaica, where this information is not available, a combination of this methodology and the indirect estimation method used in the Venezuelan application were used.

The 2005 poverty line in Brazil was fixed at half a minimum salary (R$ 150) and the 1999 poverty line at its deflated equivalent of R$ 91.18.[26] With this specification, the incidence of poverty at the household level fell from 30.3% in 1999 to 23.3% in 2005 and at the individual level from 37.4% to 30.9%,[27] as the nominal per capita income increased from R$253.14 to R$ 434.64. Unlike the Venezuelan case, poverty in Brazil was found to be slightly higher among individuals living in male headed households: 37.4% in 1999 and 31.4% in 2005, compared to 37.3% and 29.4%, respectively, in female headed households. The Brazilian poverty gap index (P_1) was 17.1% in 1999 and 13.7% in 2005. This implies that the amount of income that would need to be transferred from the non-poor to the poor in 2005 in order to eliminate poverty was R$3.693 billion, or 4.7% of the total personal income earned. The main differences in the distribution of individuals between age and sex groups and categories of relationships with the household head changed in that there was an increase of the percentages of heads of households (from 26.47% to 28.37%) and the decline of the percentage of children under age 10 (from 19.63% to 18.77%). The percentage of spouses also increased in most age categories, as did the percentage of elderly. Finally, there was a substantial increase in the proportion of households headed by women, from 22.56% to 27.25%.

Comparing the 1999 and 2005 results for Brazil in terms of the α and β parameters of the ECLAC/IPEA/UNDP model for per capita income, without considering household composition effects, one finds a value of 0.0431 for α and 0.7170 for β. The first suggests some reduction of inequality, whereas the second does not mean much, because it is based on nominal income values, without inflation correction. What is

[26] Unlike most Latin American countries, Brazil does not have an official poverty line, but the criterion adopted here is common among Brazilian poverty specialists. It would have been possible to choose the US$ 1 per day (PPP) criterion, but this leads to very low poverty estimates and is not used frequently in the country. The R$ 150 limit corresponds to US$ 2.22 per day (nominal, not PPP), according to the exchange rate of Sept. 2005.

[27] These percentages are slightly different from the ones cited earlier, due to the need to reweight.

relevant is that the α_m and β_m for the p-factor of individuals,[28] which eliminates the effect of household composition on household income, are 0.1420 and 0.5019, if modified weights for the economic contributions of different types of household members are used. This shows that a substantial portion of the apparent increase in average nominal incomes per capita can be accounted for by changes in the composition of households. In fact, if it weren't for this contribution, the percentage of poor individuals in 2005 would have been 32.8% instead of 30.9%.

A conventional projection, without effects of the population structure

For projection purposes, it is assumed that from 2005 until 2015 the strictly economic component of per capita incomes in Brazil will maintain an economic growth rate of 2.5% per year, at first without any reduction of income inequality. Discounting expected population growth during the period, as implied by standard population projections, this implies that α = 0 and β = 0.1303. These values may appear rather low, considering recent historical growth trends of the GDP per capita, but it must be borne in mind that demographic trends account for part of this historical growth, so once population change is introduced the 2.5% mentioned above will increase by another percent or so. With these classical settings of the ECLAC/IPEA/UNDP model, the predicted reduction is as follows:

- Individual poverty in Brazil falls from 30.9% to 28.4%, i.e. by 2.5 percentage points.

Assuming that, in addition, there would be a 7.5% reduction in income inequality (α = 0.075), poverty would go down further, but under this specification there is a difference between α and β, on the one hand, and α_m and β_m, on the other, and the results depend somewhat on the choice of coefficients for the determination of the income generating capacity. With direct coefficients, one needs α_m = 0.1004 and β_m = 0.1421 to obtain α = 0.075 and β = 0.1303, leading to the following poverty reduction:

- Individual poverty falls from 30.9% to 22.1%, i.e. by 8.8 percentage points.

With modified coefficients, the values needed are α_m = 0.1176 and β_m = 0.1354, leading to the following poverty reduction:

- Individual poverty falls from 30.9% to 23.1%, i.e. by 7.8 percentage points.

No great relevance should be attributed to the difference between these percentages; they merely provide a benchmark against which the population effects described below should be measured.

[28] As was observed in footnote 24, the use of the α and β parameters in DMPAP is not exactly the same as in the original ECLAC/ IPEA/ UNDP model. In the latter, these parameters the overall per capita income, whereas in DMPAP they affect only the strictly economic part called p-factor, which does not include the household composition effect. To distinguish the two, the the α and β parameters in DMPAP are marked as the α_m and β_m.

A projection with proportional adjustments of household compositions

Changing the 2005 population composition for the 2015 composition, while making proportional adjustments to the number of individuals in each age-sex category for consistency with the population projections and using modified coefficients, one obtains the following results:

- Individual poverty in Brazil falls from 23.1% to 17.9%, i.e. by 5.2 percentage points;
- The poverty gap (P_1) falls from 8.7% to 6.8%;
- The Gini index at the individual level falls from 0.5225 to 0.5157.

This means that as much as 40% of the entire poverty reduction expected between 2005 and 2015 is attributable to the population effect.

If the direct, rather than the modified specification is chosen, it dampens the poverty reduction effect somewhat. In this case

- Individual poverty falls from 22.1% to 18.2%, i.e. by 3.9 percentage points;
- The poverty gap (P_1) falls from 7.3% to 5.8%;
- The Gini index at the individual level falls from 0.5225 to 0.5197.

This smaller effect is to be expected, exactly because of the way direct and modified coefficients were defined.

As was mentioned above, age and sex composition effects increase the overall economic growth rate. Using modified coefficients, one finds an actual β of 1.2472, instead of 0.1303, which corresponds to an annual growth rate of 3.49%, rather than 2.5%. But the poverty reduction impact of population change is not limited to this effect. If population factors are not considered, the β needed to bring about the same poverty reduction as under the direct specification of income generating units would be 0.3267. This corresponds to an annual growth rate of 4.16% and 4.39%.

The previous figures were projected based on the 2005 values of the income generating capacity of individuals by age, sex, and relationship to the head of household. The coefficients describing these relationships may themselves change over time. In order to check what effect this might have on poverty, an analogous projection was elaborated with coefficients extrapolated to 2015. This does not change the results to any major extent. The head count index of poverty falls by about another percentage point, whereas the poverty gap and the Gini index may move up or down a bit, depending on whether direct or modified coefficients are used.

A projection with convergence of the number of children under age 10

Table 2 shows a number of alternative scenarios, such as the one where household compositions converge. In this second projection scenario, the initial rates and averages are not equal to their 2005 observed values within each income class, but to the overall averages for households with heads of the same age and sex, so that by 2015 household compositions would be uniform by income categories, varying only by age and sex of the household head. Because it may be somewhat unrealistic to expect the total convergence of all categories of household members, which is not easily achieved by any public policy intervention and contains a substantial effect of inertia due to processes inherited from the past, the scenario is limited to the convergence of numbers of children under age 10. As is evident from Table 1.5, the impact of this change is dramatic:

- Individual poverty falls from 17.9% to 10.5%, i.e. by another 7.4 percentage points;
- The poverty gap (P_1) falls from 6.8% to 3.5%;
- The Gini index at the individual level falls from 0.5157 to 0.4739.

If, in addition, age specific headship rates were to converge between income strata, the incidence of poverty would fall by another 1.2%.

Table 2: Overview of the different projection scenarios and their results for Brazil

	Poverty (P_0)	Poverty gap (P_1)	Gini index
Situation in 2005	30.92%	13.65%	0.5649
Projections to 2015 based on α = 0.075 and β = 0.1303			
Without considering the population composition			
With direct coefficients, α_m =0.1004 and β_m =0.1421	22.06%	7.25%	0.5225
With modified coefficients, α_m =0.1176 and β_m =0.1354	23.11%	8.66%	0.5225
Percentage points of difference with respect to the corresponding previous projections			
Projections with proportional adjustments of household composition			
According to methodology of Paes de Barros et al.	-3.95%		
With direct coefficients (of 2005)	-3.86%	-1.42%	-0.0028
With modified coefficients (of 2005)	-5.22%	-1.90%	-0.0068
With direct coefficients (extrapolated to 2015)	-4.88%	-1.83%	-0.0045
With modified coefficients (extrapolated to 2015)	-6.34%	-1.83%	-0.0035
Projections with convergence of rates and averages in the 0-9 year age group			
With direct coefficients (of 2005)	-9.59%	-3.84%	-0.0329
With modified coefficients (of 2005)	-12.65%	-5.16%	-0.0486
Also of the headship rates, direct coefficients (2005)	-10.75%	-4.24%	-0.0383
Idem with modified coefficients (2005)	-13.82%	-5.58%	-0.0541
Projections with additional fertility decline of 20%			
With direct coefficients (of 2005)	-5.76%	-2.10%	-0.0060
With modified coefficients (of 2005)	-7.40%	-2.74%	-0.0126

Source: Hakkert (2007)

Another alternative scenario is one in which fertility falls more than under the standard projections, e.g. another 20% by 2015. As shown in Table 2, this results in a reduction of poverty to 15.7%, rather than 17.9%. While this reduction is certainly significant, it is much less dramatic than the one found under the previous scenario. One should conclude, therefore that the distribution of fertility decline among social strata is a more important determinant of poverty reduction than the overall decline of fertility.

Although there are some methodological differences between this analysis of Brazil and the earlier one on Venezuela (Hakkert, 2006), the main results are similar. Using corrected indirect coefficients (see Technical Note 2 for an explanation) and proportional adjustments in household composition, poverty was projected to fall by 7.3% more than in a standard projection without population composition effects based on a 3.0% annual growth rate in Venezuela between 2004 and 2015, compared to 5.2% in Brazil. Convergence in the 0-9 age group added only 0.4 percentage points to this difference, as opposed to the rather sweeping effect found in Brazil. This may be due to the methodological differences between both analyses. A further decline of 10% in fertility between 2004 and 2015 would lower poverty by another 1.3 percentage points, whereas a stagnation of fertility decline at its 2004 level would increase poverty by 2.0 percentage points.

The main conclusions of Table 2 may be summarised as follows:

1. The potential contribution of demographic trends in both Venezuela and Brazil to the reduction of poverty up to 2015 is very substantial, corresponding to an equivalent additional growth rate of the per capita GDP in the order of 1-2%, with an incidence of poverty in 2015 that could be 4-11% lower than the projected figures in the absence of population composition effects.
2. The effect of population dynamics on *inequality* generally points in the same direction as the effect on the head-count index of poverty. More demographic inequality generally implies more economic inequality. On the whole, the range of possible Gini indexes in Venezuela in 2015 varies from about 0.38 to 0.42, compared to a value of 0.4303 if demographic effects are ignored. In Brazil, the range is from 0.50 to 0.54, compared to a value of 0.5487 if demographic effects are ignored.
3. Most of this contribution is already implicit in demographic changes that took place previous to 2005, as the inertial effect of the fall of fertility in the past. But not everything is pre-ordained by the past. Depending on the characteristics of the fertility change from 2005 until 2015, poverty in 2015 may be as low as 9.4% or as high as 17.9% in Brazil.

4. The importance of demographic factors during this upcoming period refers more to the distribution of trends among different social strata than to the aggregate rhythm of fertility decline – at least in countries like Brazil, where the demographic transition is already fairly advanced. A more vigorous decline of fertility would increase the rhythm of poverty reduction, but a convergence between social strata of the levels of fertility already projected would have a considerably greater impact in Brazil and possibly in other countries as well.
5. Apart from this last acknowledgement, the importance of the preceding analysis resides primarily in its application to target setting for poverty reduction. The fact that a reduction in the order of 4-6 percentage points by 2015 is already implicit in the current demographic evolution of Brazil and other countries should stimulate the setting of more ambitious political targets, whose attainment will require actual public policy intervention, rather than simply riding the demographic tide.

An example with Jamaican data

The data for the Jamaican application come from the Living Standards Measurement Surveys (LSMS) held in 1999 and 2004. In order to be compatible with the historical time series and population projections of CELADE, adjustments were applied to the age and sex structures of the populations surveyed in both years, as in the previous illustrations of Venezuela and Brazil. In cells C3 and C4 of the Main spreadsheet, the two base years are entered: 1999 and 2004. In cells E3 to G4, one enters the chosen poverty lines. Again, the official poverty lines for each year were used. In 1999, these were J$ 27,771.4 for rural areas, J$ 29,824.4 for non-metropolitan urban areas, and J$ 31,294 for metropolitan Kingston. In 2004, the corresponding limits were J$ 41,987.25, J$ 45,091.13, and J$ 47,312.94, respectively. The reference lines in H3 to J4 were set at the same values. Again, with this information and the contents of the spreadsheets Data1, Data2, and Projection, the programme produces a series of summary indicators in the first lines of the Main spreadsheet.

The number of cases in Data1 is 1875 and in Data2 1981. Here these cases refer to the original sample, without clustering. The nominal average income at the household level increased from J$ 241,665 in 1999 to J$ 327,468 in 2004. However, as in Venezuela, due to the inflation, the incidence of poverty at the household level went up, from 11.15% to 13.11%. At the level of individuals, shown below the household data, the average per capita income increased from J$ 68,153 to J$ 97,592 per capita, and poverty from 17.39% to 20.21%. The Gini index, also at the individual level, was reduced slightly, from 0.3890 to 0.3880. The data are also shown separately for male and female headed households. This reveals a higher incidence of poverty among individuals living in female headed households: 20.19% in 1999 and 21.63% in 2004, against 14.91% and 18.78% for male headed households. The poverty gap index was 4.95% in 1999 and 5.05% in 2004. Multiplied by the poverty line and the total population size, this yields the total amount of resources that would have to be transferred from the non-poor to the poor in order to eliminate poverty, but the existence of multiple poverty lines complicates the computation, so it has not been carried out here.

The summary of population data indicates that the number of households increased from 720,937 in 1999 to 797,703 in 2004, while the number of household members increased from 2,556,384 to 2,676,681. There was also an increase in the proportion of households headed by women, from 47.0% to 50.3%. The number of income generating units (using 2004 income generating coefficients) increased from 915,950 to 919,804, i.e. a 0.4% increase, whereas the number of individuals increased by 4.7%. All else being equal, this suggests an decrease of productivity per person of about 4%. Lines 20-26 display the distribution of individuals between age and sex groups and categories of relationships with the household head. The main differences refer to the increase of the percentages of heads of households (from 28.20% to 29.80%) and the decline of the percentage of children under age 10 (from 21.24% to 19.73%). The percentage of spouses also increased in the age categories under 50, but curiously enough the percentage of elderly did not.

Jamaica presented the most difficult situation of the three countries with respect to the income generating capacity coefficients c_{ijt}. In principle, the goal was to generate separate coefficients by economic strata, in the same way as in Brazil. However, unlike the Brazilian PNAD, the Jamaican LSMS does not contain individual income data, nor can these easily be constructed from alternative data sources. This is because the LSMS in Jamaica is based on the consumption, rather than the income criterion. The census does contain individual income information, but this is generally considered to be of poor quality. The absence of these data, therefore, made it necessary to choose between three alternatives, none of which are entirely satisfactory:

1. Use the coefficients computed in the previous application, of Brazil, even though significant differences are to be expected between the structure of incomes by age and sex in Brazil and in Jamaica;
2. Collapse the four social strata into one and apply the algorithm for the minimisation of the variation p-factor that underlies the estimation of the coefficients in the case of Venezuela, even though there are good reasons to expect differences between economic strata;

3. Try to devise a computational algorithm similar to the one used in the case of Venezuela, but that will allow separate estimates for each economic stratum.

After experimenting with different options, in the end it was decided to prepare two sets of estimates: one based on the first and one based on the third criterion. Clearly, both of these are likely to be deficient, but at least the comparison of the two provides some idea about the likely margins of error. In the case of the last option, applying the algorithm to minimise the variation in p-factor directly in order to obtain all 4x42 c_{ijt} coefficients in each of the two base years turned out to be impractical and led to implausible results in some cases. As the next best thing, the algorithm was applied using the coefficients of the highest of the four strata as the only ones that were allowed to be adjusted, whereas the relationship of these values with the coefficients of the other three strata was maintained the same as in the Brazilian case. This produced a more or less acceptable set of c_{ijt} values, even though it would be highly desirable to validate the set of coefficients obtained in this manner with actual individual income data.

A conventional projection, without effects of the population structure

As in the previous applications, it is assumed that from 2004 until 2015 Jamaica will maintain an economic growth rate of 3.0% per year, at first without any reduction of income inequality. Discounting expected population growth during the period, as implied by the population projections, this implies that $\alpha = 0$ and $\beta = 0.2440$, values which are substituted in the yellow field of cells N43 and N44. The projection year in cell C5 is maintained at 2004, meaning that population changes are not considered at first. With these classical settings of the ECLAC/IPEA/UNDP model, the predicted reduction of individual poverty is from 20.21% to 16.21%, as can be read from cell M11. Assuming that, in addition, that $\alpha = 0.075$, poverty would go down further, to 13.64%. Extreme poverty (under half the poverty line) would fall from 2.14% (or 2.57%, according to the projected distribution) to 0.24%. As was explained earlier, these α and β values do not correspond exactly to a 24.4% increase of average incomes and a 7.5% decline of the Gini index, due to the fact that they modify **p-factor** and not **percap** directly. In order to achieve that, slightly different values have to be chosen, namely $\alpha = 0.1539$ and $\beta = 0.2216$. Under this specification, poverty falls to 12.40% and extreme poverty to 0.15%.

The base projection

When the population structure of 2015 is introduced without further parameter changes, one obtains what has previously been called the base projection, with the following results:

- Individual poverty falls from 12.40% to 9.59%;

- The poverty gap (P_1) actually increases slightly, from 1.83% to 2.20%;
- The Gini index at the individual level also increases marginally, from 0.3589 to 0.3611.

In this scenario, the numbers of members of all age-sex categories in the households are adjusted proportionally, to obtain the correct aggregate figures. As is clear from the results, this reduces the head-count measure of poverty by just under 3%, whereas the other poverty measures actually deteriorate slightly. More substantial reductions are obtained under the following scenario, in which the household structures converge, at least in terms of their headship rates and numbers of children under age 10.

A projection with convergent rates and averages

As in the previous applications, a second projection scenario was prepared in which the initial rates and averages are not equal to their 2004 observed values within each income class, but to the over-all averages for households with heads of the same age and sex, so that by 2015 the number of children under age 10 would be uniform by income categories, varying only by age and sex of the household head. Headship and spouseship rates are also made uniform. After the usual recalibrations, the following results are obtained:

- Individual poverty falls much further, from 9.59% to 4.53%;
- The poverty gap (P_1) falls from 2.20% to 0.97%;
- The Gini index at the individual level also falls much further, from 0.3611 to 0.3047.

The changes in the incidence of poverty are substantial, particularly the change in the Gini index. Again, the latter is important in that it expresses the impact of demographic equality on economic equality, which must be summed with the impact of fertility reduction already implied by the base projection.

Two inertial scenarios, based entirely on population change

At present, the Planning Institute of Jamaica (PIOJ) is working on a macro-model for the country based on the Threshold 21 paradigm. Because this is a macro-model, rather than a micro-simulation model for household data, like DMPAP, the two are not particularly easy to integrate. Nevertheless, it would be useful, for the purposes of Threshold 21, to clarify what the inertial scenario for poverty reduction up to 2015 would be, in the absence of any economic growth or income redistribution (α=0 and β=0). It is often implicitly assumed that under this "business-as-usual" scenario, the incidence of poverty will remain stationary, but as should be clear by now from the preceding scenarios, this assumption is not correct as demographic change will continue to push poverty rates down.

Under the homogeneous scenario, where the number of members of particular age-sex categories changes proportionally in all households, in accordance with the aggregate age-sex specific population projections, poverty is expected to decline from 20.21% in 2004 to 18.75%, i.e. a 1.5% decline over a period of 11 years. However, if the demographic scenario were to be characterised by a convergence of demographic behaviour among social strata, so that headships rates and numbers of children under age 10 were to become equal among all strata by 2015, the effect would be considerably stronger and poverty rates might fall as low as 13.39%, i.e. by more than 7%, without change in the underlying economic parameters.

Scenarios based on the Brazilian income generating capacity coefficients

In order to obtain some idea about the sensitivity of the results to a different choice of the c_{ik} coefficients, the analyses above were repeated using the coefficients estimated for the case of Brazil. In the case of the simple projection without account of changes in the age-sex structure of households, the main changes have to do with the α and β parameters needed to achieve a 24.4% increase of all incomes coupled with a 7.5% reduction of the individual Gini index. With the Brazilian c_{ij} coefficients, these values now become α=0.2009 and β=0.1631, bringing poverty down to 12.52%, rather than the 12.40% found above. Extreme poverty falls to 0.84%, rather than 0.15%.

The two scenarios with changes in the age-sex structure also yield different results. Under the scenario of homogeneous change in all households, poverty falls to 8.30%, rather than 9.59%, whereas the poverty gap falls to 2.14%, rather than 2.20%, and the Gini index to 0.3517, rather than 0.3611. Overall, this indicates a slightly stronger poverty reduction effect than under the previous specification, with the c_{ij} coefficients estimated by minimisation of the variation in p-factor. The other projection with changing household compositions, based on the convergence of headship rates and numbers of children under age 10, yields a poverty rate of 3.09%, rather than 4.53%, whereas the poverty gap falls to 0.84%, rather than 0.97%, and the Gini index to 0.2969, rather than 0.3047. Again, this indicates that the poverty reduction effect of a changing population structure is stronger if the Brazilian coefficients are used.

With regards to the inertial scenarios, the results for 2015 are 16.98% and 13.04%, respectively, instead of the 18.75% and 13.39% found above. Again, this confirms that the population effects derived in the projections above are relatively conservative and that the use of the Brazilian coefficients leads to lower poverty projections.

ANNEX 1: Illustrative SPSS syntax file for the generation of the EXCEL file Data2 from data of the 2005 PNAD of Brazil

```
DATA LIST FILE='c:\bases\brasil\2005\dados\pes2005.txt'
  /uf 5-6 controle 5-12 serie 13-15 ordem 16-17 sexo 18-18 edad 27-29
parentco 30-30 rendtot 721-732 rendom 733-744 area 776-776 peso 777-
781 pessoas 792-793 percap 794-805.
EXECUTE.

SELECT IF ((uf>16) OR (area<4)).

COMPUTE pline=150.
COMPUTE alfa=0.075.
COMPUTE beta=0.1866.
IF (sexo=2) sexo=1.
IF (sexo=4) sexo=2.
IF (rendtot=-1) rendtot=0.
EXECUTE.

* Crear clave de hogar hhkey.
IF (parentco=1) #t=#t+1.
IF (parentco=1) #e=edad.
COMPUTE hhkey=#t.
COMPUTE edadjef=#e.
EXECUTE.

* Crear grupos etarios para toda la población.
COMPUTE grupo=TRUNC(edad/5)+1.
IF (grupo>16) grupo=16.
IF (grupo<4 & parentco=1) grupo=4.
COMPUTE grupo2=1.
IF (edad>9) grupo2=2.
IF (edad>14) grupo2=3.
IF (edad>24) grupo2=4.
IF (edad>34) grupo2=5.
IF (edad>49) grupo2=6.
IF (edad>64) grupo2=7.
IF (edad>74) grupo2=8.
IF (grupo2<3 & parentco=1) grupo2=3.
EXECUTE.

* Tabular población por edad y sexo, usando el factor de expansión de
los individuos.
WEIGHT BY peso.
VALUE LABELS sexo 1 'M' 2 'F'.
VALUE LABELS grupo 1 '0-4' 2 '5-9' 3 '10-14' 4 '15-19' 5 '20-24' 6
'25-29' 7 '30-34' 8 '35-39' 9 '40-44' 10 '45-49' 11 '50-54' 12 '55-
59' 13 '60-64' 14 '65-69' 15 '70-74' 16 '75+'.

* Eliminar pensionistas, empleados domésticos etc.
SELECT IF (parentco<5 & percap<9999999 & edad<999 & edadjef<999).
EXECUTE.
```

```
* Tabular población por edad y sexo, usando el factor de expansión de
los individuos.
WEIGHT BY peso.
CROSSTABS
  /TABLES=grupo  BY sexo
  /FORMAT= AVALUE TABLES
  /CELLS= COUNT
  /COUNT ROUND CELL.

* Corrección de pesos según proyecciones de población.
COMPUTE pesop=peso.
COMPUTE uno=1.
IF (grupo=1 & sexo=1) pesop=peso*1.248871.
IF (grupo=2 & sexo=1) pesop=peso*0.999813.
IF (grupo=3 & sexo=1) pesop=peso*0.961683.
IF (grupo=4 & sexo=1) pesop=peso*0.996700.
IF (grupo=5 & sexo=1) pesop=peso*1.058678.
IF (grupo=6 & sexo=1) pesop=peso*1.071798.
IF (grupo=7 & sexo=1) pesop=peso*1.066893.
IF (grupo=8 & sexo=1) pesop=peso*1.082690.
IF (grupo=9 & sexo=1) pesop=peso*1.036316.
IF (grupo=10 & sexo=1) pesop=peso*0.991135.
IF (grupo=11 & sexo=1) pesop=peso*0.929714.
IF (grupo=12 & sexo=1) pesop=peso*0.924684.
IF (grupo=13 & sexo=1) pesop=peso*0.923614.
IF (grupo=14 & sexo=1) pesop=peso*0.958888.
IF (grupo=15 & sexo=1) pesop=peso*0.926772.
IF (grupo=16 & sexo=1) pesop=peso*0.938513.
IF (grupo=1 & sexo=2) pesop=peso*1.230460.
IF (grupo=2 & sexo=2) pesop=peso*1.020141.
IF (grupo=3 & sexo=2) pesop=peso*0.983730.
IF (grupo=4 & sexo=2) pesop=peso*1.008551.
IF (grupo=5 & sexo=2) pesop=peso*1.059016.
IF (grupo=6 & sexo=2) pesop=peso*1.024611.
IF (grupo=7 & sexo=2) pesop=peso*1.026635.
IF (grupo=8 & sexo=2) pesop=peso*1.022655.
IF (grupo=9 & sexo=2) pesop=peso*1.044285.
IF (grupo=10 & sexo=2) pesop=peso*0.980973.
IF (grupo=11 & sexo=2) pesop=peso*0.939569.
IF (grupo=12 & sexo=2) pesop=peso*0.924599.
IF (grupo=13 & sexo=2) pesop=peso*0.929983.
IF (grupo=14 & sexo=2) pesop=peso*0.938688.
IF (grupo=15 & sexo=2) pesop=peso*0.882323.
IF (grupo=16 & sexo=2) pesop=peso*0.870072.
COMPUTE pesopro=peso.
IF (grupo=1 & sexo=1) pesopro=peso*1.223336.
IF (grupo=2 & sexo=1) pesopro=peso*1.065845.
IF (grupo=3 & sexo=1) pesopro=peso*1.045489.
IF (grupo=4 & sexo=1) pesopro=peso*0.974214.
IF (grupo=5 & sexo=1) pesopro=peso*0.973669.
IF (grupo=6 & sexo=1) pesopro=peso*1.168207.
```

```
IF (grupo=7 & sexo=1) pesopro=peso*1.319892.
IF (grupo=8 & sexo=1) pesopro=peso*1.248641.
IF (grupo=9 & sexo=1) pesopro=peso*1.157874.
IF (grupo=10 & sexo=1) pesopro=peso*1.246633.
IF (grupo=11 & sexo=1) pesopro=peso*1.331395.
IF (grupo=12 & sexo=1) pesopro=peso*1.365148.
IF (grupo=13 & sexo=1) pesopro=peso*1.358244.
IF (grupo=14 & sexo=1) pesopro=peso*1.266532.
IF (grupo=15 & sexo=1) pesopro=peso*1.239339.
IF (grupo=16 & sexo=1) pesopro=peso*1.347613.
IF (grupo=1 & sexo=2) pesopro=peso*1.203956.
IF (grupo=2 & sexo=2) pesopro=peso*1.085927.
IF (grupo=3 & sexo=2) pesopro=peso*1.065597.
IF (grupo=4 & sexo=2) pesopro=peso*0.979686.
IF (grupo=5 & sexo=2) pesopro=peso*0.968611.
IF (grupo=6 & sexo=2) pesopro=peso*1.111139.
IF (grupo=7 & sexo=2) pesopro=peso*1.266360.
IF (grupo=8 & sexo=2) pesopro=peso*1.158913.
IF (grupo=9 & sexo=2) pesopro=peso*1.112320.
IF (grupo=10 & sexo=2) pesopro=peso*1.197638.
IF (grupo=11 & sexo=2) pesopro=peso*1.350152.
IF (grupo=12 & sexo=2) pesopro=peso*1.398013.
IF (grupo=13 & sexo=2) pesopro=peso*1.395930.
IF (grupo=14 & sexo=2) pesopro=peso*1.283058.
IF (grupo=15 & sexo=2) pesopro=peso*1.219906.
IF (grupo=16 & sexo=2) pesopro=peso*1.312798.
EXECUTE.

* Verificar población corregida por edad y sexo, usando el factor de
expansión de los individuos.
WEIGHT BY pesop.
CROSSTABS
  /TABLES=grupo  BY sexo
  /FORMAT= AVALUE TABLES
  /CELLS= COUNT
  /COUNT ROUND CELL .
WEIGHT OFF.

* Calcular pesos de capacidad de generación de ingresos.
COMPUTE parco=parentco.
IF (parentco>2) parco=3.
COMPUTE pobre=(percap<pline).
COMPUTE estrato=1.
IF (percap>0.5*pline) estrato=2.
IF (percap>pline) estrato=3.
IF (percap>2*pline) estrato=4.
EXECUTE

WEIGHT BY pesop.
USE ALL.
COMPUTE filter_$=(rendtot<9999999).
```

```
VARIABLE LABEL filter_$ 'rendtot<9999999 (FILTER)'.
VALUE LABELS filter_$  0 'Not Selected' 1 'Selected'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
MEANS
  TABLES=rendtot  BY parco  BY sexo BY grupo2
  /CELLS MEAN COUNT STDDEV.
USE ALL.
MEANS
  TABLES=percap  BY sexo
  /CELLS MEAN COUNT STDDEV.
MEANS
  TABLES=pobre  BY sexo
  /CELLS MEAN COUNT STDDEV.

* Calcular y tabelar pobreza y calcular ingreso individual
proyectado.
COMPUTE rendpro=(1+beta)*((1-alfa)*rendtot+alfa*434.65).
COMPUTE pcap=0.1*trunc(10*percap/pline).
EXECUTE.
WEIGHT BY pesop.
FREQUENCIES
  VARIABLES=pcap
  /ORDER=ANALYSIS .
WEIGHT OFF.

* Crear peso, sólo para jefes.
IF (parentco=1) #p=pesop.
IF (parentco=1) #q=pesopro.
COMPUTE pesoj=#p.
COMPUTE pesoproj=#q.
EXECUTE.

*Crear dummies para los grupos etarios relevantes.
COMPUTE ed09=(edad<10)*pesop/pesoj.
COMPUTE ed1014m=(edad>9 & edad<15 & sexo=1)*pesop/pesoj.
COMPUTE ed1014f=(edad>9 & edad<15 & sexo=2)*pesop/pesoj.
COMPUTE ed1524m=(edad>14 & edad<25 & sexo=1)*pesop/pesoj.
COMPUTE ed1524f=(edad>14 & edad<25 & sexo=2)*pesop/pesoj.
COMPUTE ed2534m=(edad>24 & edad<35 & sexo=1)*pesop/pesoj.
COMPUTE ed2534f=(edad>24 & edad<35 & sexo=2)*pesop/pesoj.
COMPUTE ed3549m=(edad>34 & edad<50 & sexo=1)*pesop/pesoj.
COMPUTE ed3549f=(edad>34 & edad<50 & sexo=2)*pesop/pesoj.
COMPUTE ed5064m=(edad>49 & edad<65 & sexo=1)*pesop/pesoj.
COMPUTE ed5064f=(edad>49 & edad<65 & sexo=2)*pesop/pesoj.
COMPUTE ed6574m=(edad>64 & edad<75 & sexo=1)*pesop/pesoj.
COMPUTE ed6574f=(edad>64 & edad<75 & sexo=2)*pesop/pesoj.
COMPUTE ed75m=(edad>74 & sexo=1)*pesop/pesoj.
COMPUTE ed75f=(edad>74 & sexo=2)*pesop/pesoj.
COMPUTE con1524=(edad>14 & edad<25 & parentco=2)*pesop/pesoj.
```

```
COMPUTE con2534=(edad>24 & edad<35 & parentco=2)*pesop/pesoj.
COMPUTE con3549=(edad>34 & edad<50 & parentco=2)*pesop/pesoj.
COMPUTE con5064=(edad>49 & edad<65 & parentco=2)*pesop/pesoj.
COMPUTE con6574=(edad>64 & edad<75 & parentco=2)*pesop/pesoj.
COMPUTE con75=(edad>74 & parentco=2)*pesop/pesoj.
EXECUTE.

*Crear dummies para los grupos etarios proyectados relevantes.
COMPUTE edpro09=(edad<10)*pesopro/pesoproj.
COMPUTE edpro1014m=(edad>9 & edad<15 & sexo=1)*pesopro/pesoproj.
COMPUTE edpro1014f=(edad>9 & edad<15 & sexo=2)*pesopro/pesoproj.
COMPUTE edpro1524m=(edad>14 & edad<25 & sexo=1)*pesopro/pesoproj.
COMPUTE edpro1524f=(edad>14 & edad<25 & sexo=2)*pesopro/pesoproj.
COMPUTE edpro2534m=(edad>24 & edad<35 & sexo=1)*pesopro/pesoproj.
COMPUTE edpro2534f=(edad>24 & edad<35 & sexo=2)*pesopro/pesoproj.
COMPUTE edpro3549m=(edad>34 & edad<50 & sexo=1)*pesopro/pesoproj.
COMPUTE edpro3549f=(edad>34 & edad<50 & sexo=2)*pesopro/pesoproj.
COMPUTE edpro5064m=(edad>49 & edad<65 & sexo=1)*pesopro/pesoproj.
COMPUTE edpro5064f=(edad>49 & edad<65 & sexo=2)*pesopro/pesoproj.
COMPUTE edpro6574m=(edad>64 & edad<75 & sexo=1)*pesopro/pesoproj.
COMPUTE edpro6574f=(edad>64 & edad<75 & sexo=2)*pesopro/pesoproj.
COMPUTE edpro75m=(edad>74 & sexo=1)*pesopro/pesoproj.
COMPUTE edpro75f=(edad>74 & sexo=2)*pesopro/pesoproj.
EXECUTE.

SORT CASES BY hhkey (A).

* Agregar otras caracteristicas.
AGGREGATE
  /OUTFILE='c:\spss\agr3.sav'
  /BREAK=hhkey
  /pesodom=sum(pesop)
  /n09=sum(ed09)
  /n1014m=sum(ed1014m)
  /n1014f=sum(ed1014f)
  /n1524m=sum(ed1524m)
  /n1524f=sum(ed1524f)
  /n2534m=sum(ed2534m)
  /n2534f=sum(ed2534f)
  /n3549m=sum(ed3549m)
  /n3549f=sum(ed3549f)
  /n5064m=sum(ed5064m)
  /n5064f=sum(ed5064f)
  /n6574m=sum(ed6574m)
  /n6574f=sum(ed6574f)
  /n75m=sum(ed75m)
  /n75f=sum(ed75f)
  /np09=sum(edpro09)
  /np1014m=sum(edpro1014m)
  /np1014f=sum(edpro1014f)
  /np1524m=sum(edpro1524m)
```

```
  /np1524f=sum(edpro1524f)
  /np2534m=sum(edpro2534m)
  /np2534f=sum(edpro2534f)
  /np3549m=sum(edpro3549m)
  /np3549f=sum(edpro3549f)
  /np5064m=sum(edpro5064m)
  /np5064f=sum(edpro5064f)
  /np6574m=sum(edpro6574m)
  /np6574f=sum(edpro6574f)
  /np75m=sum(edpro75m)
  /np75f=sum(edpro75f)
  /nc1524=sum(con1524)
  /nc2534=sum(con2534)
  /nc3549=sum(con3549)
  /nc5064=sum(con5064)
  /nc6574=sum(con6574)
  /nc75=sum(con75)
  /rendhopr=sum(rendpro).

MATCH FILES /FILE=*
  /TABLE='c:\spss\agr3.sav'
  /BY hhkey.
EXECUTE.

COMPUTE indep=(n09+n1014m+n1014f+n6574m+n6574f+n75m+n75f=0).
IF (edadjef>64 & n09+n1014m+n1014f=0 & n6574m+n75m<=1 &
n6574f+n75f<=1) indep=1.
EXECUTE.

USE ALL.
COMPUTE filter_$=(rendtot<9999999).
VARIABLE LABEL filter_$ 'rendtot<9999999 (FILTER)'.
VALUE LABELS filter_$  0 'Not Selected' 1 'Selected'.
FORMAT filter_$ (f1.0).
FILTER BY filter_$.
EXECUTE.
WEIGHT BY pesop.
MEANS
  TABLES=rendtot  BY indep BY estrato BY parco  BY sexo BY grupo2
  /CELLS MEAN COUNT STDDEV.
USE ALL.
WEIGHT OFF.

COMPUTE miembros=n09+n1014m+n1014f+n1524m+n1524f+n2534m+n2534f+n3549m
+n3549f+n5064m+n5064f+n6574m+n6574f+n75m+n75f.
COMPUTE miempro=np09+np1014m+np1014f+np1524m+np1524f+np2534m+np2534f+
np3549m+np3549f+np5064m+np5064f+np6574m+np6574f+np75m+np75f.
COMPUTE ingtot=percap*miembros.
COMPUTE percapro=rendhopr/miempro.
COMPUTE tramo=trunc(100*percap/pline).
COMPUTE tamhog=pesodom/pesoj.
```

```
EXECUTE.
* Calcular y tabelar pobreza proyectada.
COMPUTE pobrepro=(percapro<pline).
EXECUTE.
WEIGHT BY pesopro.
MEANS
  TABLES=pobrepro  BY sexo
  /CELLS MEAN COUNT STDDEV.
WEIGHT OFF.

* Eliminar todos los no-jefes, que ya no son necesarios en el
análisis.
SELECT IF (parentco=1).
EXECUTE.

SORT CASES BY percap (A).

* Cálculo de los intervalos de ingreso.
COMPUTE intval=0.
IF (percap>0.3*pline) intval=trunc(20*(percap/pline-0.3)+1).
IF (percap>0.5*pline) intval=trunc(20*(percap/pline-0.5)+5).
IF (percap>pline) intval=trunc(16*(percap/pline-1)+15).
IF (percap>1.5*pline) intval=trunc(12*(percap/pline-1.5)+23).
IF (percap>2*pline) intval=trunc(8*(percap/pline-2)+29).
IF (percap>3*pline) intval=trunc(4*(percap/pline-3)+37).
IF (percap>4*pline) intval=41.
IF (percap>4.5*pline) inval=42.
IF (percap>5*pline) intval=43.
IF (percap>5.5*pline) intval=44.
IF (percap>6*pline) intval=45.
IF (percap>6.5*pline) intval=46.
IF (percap>7.5*pline) intval=47.
IF (percap>10*pline) intval=48.
IF (percap>15*pline) intval=49.
COMPUTE categor=intval+50*(nc1524+nc2534+nc3549+nc5064+nc6574+nc75>0)
+100*min(5,rnd(n09))+600*(grupo2-3)+3600*(sexo-1).
EXECUTE.

SORT CASES BY sexo (A) grupo2 (A) ingtot (A).

WEIGHT BY pesoj.
AGGREGATE
  /OUTFILE='c:\spss\agr5.sav'
  /BREAK=categor
  /ingrtot=mean(ingtot)
  /pesojefe=sum(uno)
  /pesohog=sum(tamhog)
  /num09=mean(n09)
  /num1014m=mean(n1014m)
  /num1014f=mean(n1014f)
  /num1524m=mean(n1524m)
```

```
   /num1524f=mean(n1524f)
   /num2534m=mean(n2534m)
   /num2534f=mean(n2534f)
   /num3549m=mean(n3549m)
   /num3549f=mean(n3549f)
   /num5064m=mean(n5064m)
   /num5064f=mean(n5064f)
   /num6574m=mean(n6574m)
   /num6574f=mean(n6574f)
   /num75m=mean(n75m)
   /num75f=mean(n75f)
   /numc1524=mean(nc1524)
   /numc2534=mean(nc2534)
   /numc3549=mean(nc3549)
   /numc5064=mean(nc5064)
   /numc6574=mean(nc6574)
   /numc75=mean(nc75).
EXECUTE.

WEIGHT OFF.
GET FILE='c:\spss\agr5.sav'.
IF (trunc(categor/600)=0 & num1524m<1) num1524m=1.
IF (trunc(categor/600)=6 & num1524f<1) num1524f=1.
IF (trunc(categor/600)=1 & num2534m<1) num2534m=1.
IF (trunc(categor/600)=7 & num2534f<1) num2534f=1.
IF (trunc(categor/600)=2 & num3549m<1) num3549m=1.
IF (trunc(categor/600)=8 & num3549f<1) num3549f=1.
IF (trunc(categor/600)=3 & num5064m<1) num5064m=1.
IF (trunc(categor/600)=9 & num5064f<1) num5064f=1.
IF (trunc(categor/600)=4 & num6574m<1) num6574m=1.
IF (trunc(categor/600)=10 & num6574f<1) num6574f=1.
IF (trunc(categor/600)=5 & num75m<1) num75m=1.
IF (trunc(categor/600)=11 & num75f<1) num75f=1.
IF (num09>10) num09=10.
COMPUTE miembros=num09+num1014m+num1014f+num1524m+num1524f+num2534m+n
um2534f+num3549m+num3549f
     +num5064m+num5064f+num6574m+num6574f+num75m+num75f.
IF (ingrtot=0) ingrtot=0.1.
COMPUTE percap=ingrtot/miembros.
EXECUTE.

* Calcular y tabelar pobreza y calcular ingreso individual
proyectado.
COMPUTE pobre=(percap<150).
COMPUTE pes=pesojefe*miembros.
COMPUTE sexo=1+(categor>3599).
COMPUTE sumnc=numc1524+numc2534+numc3549+numc5064+numc6574+numc75.
IF (numc1524>num1524f & sexo=1) numc1524=num1524f.
IF (numc1524>num1524m & sexo=2) numc1524=num1524m.
IF (sumnc>1) numc1524=numc1524/sumnc.
IF (numc2534>num2534f & sexo=1) numc2534=num2534f.
```

```
IF (numc2534>num2534m & sexo=2) numc2534=num2534m.
IF (sumnc>1) numc2534=numc2534/sumnc.
IF (numc3549>num3549f & sexo=1) numc3549=num3549f.
IF (numc3549>num3549m & sexo=2) numc3549=num3549m.
IF (sumnc>1) numc3549=numc3549/sumnc.
IF (numc5064>num5064f & sexo=1) numc5064=num5064f.
IF (numc5064>num5064m & sexo=2) numc5064=num5064m.
IF (sumnc>1) numc5064=numc5064/sumnc.
IF (numc6574>num6574f & sexo=1) numc6574=num6574f.
IF (numc6574>num6574m & sexo=2) numc6574=num6574m.
IF (sumnc>1) numc6574=numc6574/sumnc.
IF (numc75>num75f & sexo=1) numc75=num75f.
IF (numc75>num75m & sexo=2) numc75=num75m.
IF (sumnc>1) numc75=numc75/sumnc.
EXECUTE.

WEIGHT BY pes.
MEANS
  TABLES=pobre  BY sexo
  /CELLS MEAN COUNT STDDEV.
WEIGHT OFF.

FORMATS ALL (f8.2).
FORMATS categor ingrtot pesojefe percap (f8.0).

SAVE TRANSLATE OUTFILE='C:\spss\data2.xls'
  /TYPE=XLS /VERSION=8 /MAP /REPLACE /FIELDNAMES
  /CELLS=VALUES
  /DROP=pesohog,pobre,pes,sexo,sumnc.
```