*Paper prepared for the 26th General Conference of*
*The International Association for Research in Income and Wealth*
*Cracow, Poland, 27 August to 2 September 2000*

# Making Cross-Country Comparisons of Income Distributions

Timothy M. Smeeding (Syracuse University, USA)
Michael Ward (The World Bank, USA)
Ian Castles (Academy of the Social Sciences, Australia)
Haeduck Lee (The World Bank, USA)

*Canberra Version*
*August 3, 2000*

For additional information please contact:

Timothy Smeeding
Center for Policy Research
The Maxwell School
Syracuse University
426 Eggers Hall
Syracuse, NY  13244-1020  USA
E-mail: tmsmeed@maxwell.syr.edu
Fax: + 315/ 443-1081
Phone: + 315/ 443-9042

This paper is placed on the following websites:        www.stat.gov.pl
                                                       www.econ.nyu.edu/dept/iariw.

# Table of Contents

## Abstract

One of the empirical hallmarks of the Canberra group project will be a set of guidelines to help produce estimates of income distribution that are more or less directly comparable when released by Central Statistical (CSO's) offices and other micro-data producers than at present. The paper is intended to provide guidance for comparability issues in rich OECD countries and to a lesser extent, in developing nations as well. This paper suggests principles and provides guidelines for how to construct comparable measures, and advice on avoiding the "land mines" one encounters when making these calculations from a micro-household survey perspective. Specifically, several issues of sampling and nonsampling error; data definitions (top and bottom coding); weighting for persons or units (families and households); the use of equivalence scales, and related definitional issues are addressed.

The paper also discusses the frontier issues which we feel are the next steps on the list of making broader definitions of income more comparable across countries. In particular, better measures of annual capital income (capital gains in particular); imputed rent for owner-occupiers and subsidized renters; and social transfers in-kind are discussed. These developments are somewhat daunting on a national scale and even more challenging in a comparative cross-national context. Finally, the issue of "real" microdata-based income comparisons using purchasing power parities (PPPs) is addressed. Here where principle and practice diverge greatly is perhaps a promising area for making some progress.

# I.    Introduction

From the beginning, the Canberra Group on Household Income Statistics decided to consider setting a standard for an expanded, uniform income definition which goes beyond money disposable personal income (DPI).  The group expressed a desire for a definition which comes closer to the macroeconomic notion of personal income as suggested by Harrison (2000) and other National Income Accountants (e.g., Franz, Walton, Ramprakash 1998), but one which is also motivated from the bottom-up microdata perspective (Smeeding 1997; Smeeding and Weinberg 1998).  There is also the need to help microdata producers to directly create more comparable estimates of DPI inequality from their existing micro-datasets.  This paper presents a *guidebook* for producing comparable DPI as well as a *framework* for the next steps in developing a more comprehensive household income definition as seen from the microdata point of view. The Canberra Group was organized to coordinate expert opinion on exactly these issues and we hope this chapter facilitates that discussion.

This paper should be seen as complimentary to the more historical and theoretical papers in the Canberra set (van der Laan 2000; Church et al. 2000), and the macro-oriented comparability chapter (Harrison 2000).  Its purpose is, in part, to spell out the "bottom-up" guidelines for reaching the current level of comparable household income distribution data, which we call here "LIS disposable income" (LIS DPI).  It is also designed to discuss how we move beyond this level on several fronts.  The chapter is also designed to complement Atkinson et al. (2000) on producing comparable trends in inequality.  And it makes use of the work in other Canberra papers (e.g., Sherman 2000; Harris 2000; Epland and Jansson 2000) as well.  The paper also draws heavily on previous contributions by the authors to Canberra reports of earlier

vintage[1] and discusses a recent Eurostat (2000) document which suggests that they have converged independently on a similar set of frontier issues for extending and improving practical income inequality measurement.

To be sure, individual national household income surveys probably could not, at present, measure all the components necessary to create the most comprehensive measure of income discussed in other papers. Nevertheless, we intend to summarize the methods for producing international comparisons by suggesting how national survey organizations might approach them. National Statistical Offices may follow the recommendations of the Canberra Group on how to measure household income and thereby produce more comparable estimates. Moving international database resources like the Luxembourg Income Study (LIS) forward further, will further ensure that participation by researchers can make consistent cross-national comparisons—that is, present appropriate statistics on an apples-to-apples basis rather than an apples-to-oranges basis.

As we will see, there is at least one current income measure which presents an apples-to-apples comparison. This measure, cash and near cash (disposable personal income, DPI), excludes other sources of personal income which would allow us to include pears, bananas, and other types of income to present comparable types of "fruit salads" across countries. To continue the metaphor, cash-to-cash—or apples-to-apples—comparisons may be inadequate because national "fruit salads" may contain different mixes of apples, pears, and bananas, so that apples alone may therefore be an incomplete index of true economic well-being. Here we attempt to provide a framework which includes those elements to measure for comparable DPI, but also to work toward improvements from a bottom-up perspective.

In fact, we start with a particular perspective from which alternatives can develop. This perspective is to include in our definition of more ideal income measure, all components that

2

contribute to improving *current* economic well-being, whether in cash or in-kind (as in Weinberg et al. 2000). The contrasting position would be to also include components that contribute to *potential* economic well-being, such as employer contributions to pensions, interest earned on retirement-based assets (e.g., Individual Retirement Accounts in the United States), unrealized capital gains, etc. Here we focus on the former, rather than on the latter broader concept. Indeed, congruence between national income accounting and microdata-based survey measures will require some explanation for these benefits in any case and so these guidelines might be seen as practically approaching the issues presented in other Canberra papers from a "CSO" microdata survey taker and inequality data producer perspective.

By choosing current economic well-being as our organizing principal, the choice of which income components to include can be made by addressing the simple question: "Is the income receiving unit (e.g., the household) better off today?" For example, along the dimensions of regular versus irregular income (or cash versus noncash) one would include both, if they are received in a form that can be spent today. Thus, for example, life insurance proceeds are income to the beneficiaries that year, whether spent or saved. "Permanent income" or regular income streams plus some annuitized value of irregular streams, a concept akin to "lifetime earnings" might be an appropriate measure to use to make comparisons within one country's income distribution, but it seems to us less relevant for year-to-year international comparisons.

Along the dimension "income versus wealth," net additions or subtractions to the latter are not considered to affect income, unless they result from diversions of current income rather than from the vagaries of the capital market. Along the dimension of "income versus expenditures," we adopt the view that the latter can exceed the former only through a reduction in net worth (by drawing down assets or from an increase in debt). This paper generally ignores microdata-based efforts on producing comparable measures of consumption, expenditure or

wealth. In principle, a paper similar to this one could address either missing components in net worth (or change in net worth), or expenditure (consumption). Few, if any, microdata sets collect survey data on all three concepts. A few expenditure surveys collect data on incomes as well, but they are normally not held to any standards of formal internal consistency. In other words, "income" and "expenditure" are not consistent enough such that when aggregated up by component, the difference between them can be counted as "change in net worth." Such surveys are sure to be in the "sweet dreams" of researchers and income accountants, but would in turn be in the nightmares of survey takers.

## II. Principles and Guidelines for Making Data More Comparable from a Bottom-Up Perspective

The task of harmonization of microdata can be much easily accomplished if the setting of standards produces an *ex-ante* acceptable set of useful guidelines that permit easier cross-national comparability of microdata across nations and continents. From our perspective, greater international comparability of income statistics (and eventually wealth and consumption statistics) should be the foremost aim of the Canberra Group.

There are two different methods of counting income sources. One begins with GDP and SNA and tries to work "top down" to the household level. However, survey takers are daunted by such aggregates because important microeconomic distributional issues, such as taxes and transfers, as netted out. Moreover, the SNA's offer does not cover such activities as the shadow economy, nonprofit institutions, and the like. Finally, no good measure of "income" comes from this approach other than the broadest aggregates.

Rather, micro-economists and survey takers seek international comparability from the "bottom up," adding individual components of income to household totals and then allowing those to be aggregated from sub-national region (e.g., "states") to country; country to "greater

4

region" (e.g., "Scandinavia" or "Europe"), or country to economic development category, or country to sub-global aggregates (e.g., "G-7" or "OECD"), and finally to the global aggregate level (e.g., "World Income Distribution" studies). In principle then, the micro approach can serve a wide range of needs and aggregation principles if comparable measures of income can be realized across countries.

## General Lessons

Our experiences over the Canberra meetings and other similar meetings in previous years have taught us several general lessons:

- What is ideally "most meaningful" may not always be what is most useful or most practical. In fact, reasonable "common denominators," such as cash disposable income, provide a sound and useful basis which can be used to help build better and more comprehensive measures of economic well-being. Our first priority is the implementation of such measures, with room for additional building blocks and "best practice" criteria to continue to update and expand them. (See sections III and IV below.)

- Sources and concepts of household income data vary tremendously across and even within nations presenting important issues for harmonization and comparability more generally. Times and surveys continue to change, hence, the establishment of any set of guidelines, these included, need to be implemented, and then regularly and periodically updated in the same way that the SNAs are periodically updated.

- Widespread availability of household income microdata is the only safe, practical, and sensible way to ensure long-term comparability and to build towards more comprehensive and complete measures of economic well-being. Nations should be prepared to share their data in a way that facilitates international comparability but does not compromise the privacy or confidentiality of survey respondents.

- Steps taken to expand measures of income used for inequality comparison across nations via imputation and simulation are very difficult to coordinate on a cross-national basis, e.g., social transfers in-kind. Hence, great care needs to be taken in developing identical choices of concepts, data, and imputation techniques before we can implement these measures on a cross-national basis. (See section V below, and Sutherland 1997.)

**Goals**

The development of international income distribution guidelines (IDG) must be shaped by a set of goals as determined by the ultimate users. The still "preliminary" United Nations IDG from 1977 did not really help microdata users because they were only of the macro to micro "top down" variety (United Nations 1977). They yield little in the way of practical advice for microdata producers or microdata users. And so they provided an example of what happens when too much is attempted on an abstract level by too few—nothing is implemented and so not much results. In order to improve on these previous efforts, every constituency and stakeholder which seeks to promote international comparability should have a chance to express the goals and principles upon which the IDG should be built. This has been the procedure and credo of the Canberra group.

The following is a partial list of priorities and criteria that have emerged from these meetings and papers:

- The cross-national Income Distribution Guidelines (IDG) must be *instructive*. They must say what is important to measure and how to measure it in order to be in compliance with international guidelines. They must, therefore, be of practical use.

- In order to be constructive, the IDG must *set priorities*. That is, there must be a set of basic minimum acceptable standards which need to be met in order to judge international comparability and compliance. Datasets not meeting these standards may not be useful for many types of international comparisons.

- The user's *ability to choose* must be stressed by the IDG. The acceptable degree of data quality will vary by use and purpose of the comparison at hand. What will be necessary then, are a set of quality standards which allow the users to decide whether the data being analyzed meet their standards.

- These IDG must be *flexible* in that over time, quality/quantity of measurement will also change. Widespread availability of household income microdata prepared according to a common plan will provide both a basis of comparison and facilitate data harmonization. It will also provide the new material to move from current standards to more complete measures of incomes, built from the bottom-up approach. Finally, it will provide the opportunity to expand the range of nations and economies who will find the IDG relevant and useful.

- Above all, the IDG must be ***widely accepted and practiced*** by the data-producing community.  Adoption and endorsement by international bodies is useful mainly to the extent that it facilitates greater actual usage by survey takers, because the IDG are only useful if they are practiced and implemented.  Top down approaches, which satisfy macrodata producers only may do very little at the end of the day to improve microdata comparability and satisfy microdata producers and users.  There must be a common understanding from both sides on this issue.

- Finally, and perhaps most importantly, the methods for producing new and broader estimates of income distribution must produce ***greater comparability*** of data.  Not all techniques can be implemented in a comparable way.  Often the tradeoff between the "best" and "second best" technique results in accepting a lesser quality measure, or a less inclusive measure, because no other set of countries can produce the "best" measure with existing resources.  Here the best measure can serve as an example for others to follow.  But in the end, comparability with other nations is the key value added by these guidelines.

With these goals and principles in mind, we address the issues at hand.

## III.  Components of LIS DPI and More Ideal Income Measures

We begin with the logical contents of a bottom-up broad full household income measure compared to the currently practiced DPI measure.  Table 1 presents the framework for one possible definition in schematic form based on the LIS database and on earlier Canberra papers (Smeeding and Weinberg 1998; Smeeding 1997).  It is very close to Table 1 in both Church et al. (2000) and Harrison (2000).  This definition does not attempt to be exhaustive in that it is not intended to address all sources of income, placing each of them into a particular component (more exhaustive definitions are found in these other papers).  Rather this definition and its structure reflects our prejudices as to the current state of affairs in survey-based income measurement and our choices for next steps in moving forward.  The next sections discuss the difficulties of collecting such data in a household survey environment and making them more comparable.  Then we focus on implementation and measurement issues, particularly on those items which can be collected from surveys versus those elements which must be estimated from

other data (simulation and imputation techniques or administrative records), and those elements which may require both types of estimations.

Any changes to "official" national definitions of income must be discussed extensively in formal settings by the relevant government officials in consultation with the private sector before being adopted. Yet countries such as Australia, Sweden, the United Kingdom, and the United States already present expanded experimental measures of income for public discussion, and a group consensus definition, or a set of definitions, could lead to development of such experimental measures and ultimately to their adoption on a widespread basis.

**Beyond LIS Disposable Personal Income: Toward More Ideal Income Measures**

The section headings that follow correspond to the categories of income in Table 1. This series of building blocks begins with the boxed measure of household net disposable income (including currently received cash and nearcash income net of direct tax) on the left side of Table 1. This is the income concept which is most often found in published inequality estimates (Weinberg et al. 2000) and it is the one most used in cross-national comparisons, as we shall see below. We begin here by describing the items outside the box labeled "Existing LIS DPI" commenting on how it could be expanded to include the "more ideal" income components on the right-hand side of Table 1. Because the left-hand side components are discussed for the most part in other papers (e.g., Church et al. 2000), only a cursory overview of the boxed items are presented here.

Employer contributions to social insurance, for "fringe benefits" and other forms of earnings in-kind, are not counted as disposable income. Thus, the employer cost of labor is not well estimated by current employee remuneration from work (on the left hand side of item A). Earnings-in-kind are very important sources of income for developing economies; they are less important in rich nations (Szekely and Hilgert 1999). An appropriate and comparable measure

of household income in developing countries necessitates such measures. Without estimates of the value of production for our consumption or barter, income in rural areas is seriously underestimated. Broadening the definition to include total factor income includes three important issues; one both a conceptual and a measurement issue (imputed rent); and two which require new, sensitive, and difficult measurement issues alone (realized capital gains and losses, and interest income paid). These are discussed in section V below.

Also included on the right-hand side of Figure 1 are Social Transfers In-Kind (STIK). These may be of a social insurance or of a social assistance type. They provide current goods and services in noncash form. These include food, housing, medical (health care), education, transportation, and child care benefits provided free or below cost to citizens (see also Walton 1999). They are by most accounts the most notable exclusions from DPI as compared to a "more ideal" income definition. Hence, they are discussed more clearly and completely below.

Current practice counts only amounts received at regular private interhousehold transfers in gross cash income, and so that income would be double-counted (in the aggregate) since they are also counted in the gross cash income of the people making the transfer. If they are to be accounted for, they should be added to recipients' incomes *and* subtracted from donors' incomes. One distinction that some find useful is whether these transfers are mandatory (e.g., as the result of legally binding agreements) or voluntary, though determining such a distinction accurately in a survey context is unlikely.

**Gross Cash Income, Direct Taxes, and Net Income**

The sum total of all of the amounts on the left-hand side equals gross cash income (G.), a concept used by several statistical offices, e.g., the United Kingdom and the United States as part of their regular income series. Subtraction of direct taxes paid produces the measure of disposable income used by LIS, OECD (1999), and others (e.g., Canada). On the left, direct

taxes include personal income taxes (negative if they contain refundable tax credits) and payroll taxes on employees and the self-employed. Most taxes are simulated (estimated) by applying tax rules to the income collected in the survey. These methods produce good estimates of average taxes, but are not very good at capturing the variance in taxes paid at higher income levels. To extend these estimates to include indirect taxes, such as VAT, or employer payroll taxes greatly extends the amount of information needed to make such estimates. (See the right side of Table 1.) For instance, one needs to include taxable consumption (for VAT), and incidence assumptions regarding employer taxes on labor (payroll tax) and on capital (corporate income tax). Moreover, treatment of property or wealth taxes on owner occupied homes may be redundant if imputed rent is also estimated to include these payments.

We return to these issues in Section V, but first in Section IV we discuss the current state of comparability, i.e., the construction of elements in the left-hand side of Table 1 in comparable survey form, and their use in producing inequality statistics.

## IV. Producing Comparable Estimates: Experiences from the Luxembourg Income Study (LIS) and from Other Comparable Projects

The establishment of useful and practical micro-micro international guidelines for income distribution (IDG for short) is a task to which the LIS is well-suited to contribute. LIS is uniquely situated to help out because of its experience, history, and goals. Founded in 1983, LIS is committed to the open sharing of harmonized household income survey microdata at zero user cost while still preserving the confidentiality and privacy of survey respondents. At this moment, the LIS project provides more than 100 datasets, covering 28 nations over the 1970-1997 period, including the transition economies of Central and Eastern Europe and soon the rapidly growing countries of the Pacific Rim. It does not include several EC nations data due to

severe user restrictions imposed by the European Statistical Office for use of its European Community Household Panel (ECHP) datasets.

The LIS *modus operandi* is to obtain existing national household income survey data and to do the best it can to harmonize and make these data comparable. Data harmonization improves comparability and therefore, the ratio of signal (true values) to noise (statistical or other differences) in datasets. The harmonized LIS data are made available to users via electronic mail access using STATA, SPSS and SAS software. Over its lifetime, LIS has continued to add to and expand the richness of its data and its usefulness. New variable definitions for the IVth wave of LIS (1994-1997 datasets) include separate categories for new forms of public transfer income, e.g., guaranteed child support, child care subsidies, allowances for care of invalids, and greater detail among original LIS income categories (e.g., a finer breakdown of pension income sources).

The LIS income categorization scheme can also be unfolded so as to enlarge the scope and definition of household income to include greater detail and breadth (see below). We also continue to update our technical and institutional documentation (RAR's) so that survey quality can be ascertained and so that the numerical values which LIS contains can be put into a social, legal, and political context. As part of our technical documentation, we collect micro-macro comparisons of the quality of income data collected whenever possible, though we cannot guarantee the accuracy of these comparisons.

More than anything else, it is our belief that our most recent analyses of income distribution for the OECD (Atkinson, Rainwater and Smeeding 1995) (or ARS), can help lead the way to improved income distribution estimates. It is a foundation from which future efforts can push the measurement envelope outward in building block fashion.

**Household Income Definitions**

Since the publication of the ARS volume, numerous national and international studies have begun to use this "LIS-DPI" definition. For instance, OECD (1997; 1999) has produced a series of studies on income and poverty which were compiled by national CSO's based on the ARS disposable income definition using the definitions discussed below. Similarly, the InterAmerican Development Bank (IDB) had begun to harmonize Latin American datasets based on the LIS model (e.g., Szekely and Hilgert 1999), but with considerable attention to production for home (nonmarket) consumption.

Both the Harrison (2000) and Weinberg et al. (2000) papers provide a very useful overall income accounting framework for review. Church et al. (2000) further defines each of the components in the left-hand-side box of Figure 1. Harrison (2000) links micro and macro, but do not provide guidelines for ways to achieve these measures in practice. One might begin by asking if tax and transfer household disposable (spendable) income (i.e., left column and tax in Table 1), is the most appropriate measure for international income comparisons? In our opinion, the most appropriate measure for comparison depends on *both* the ideal income measure or "framework" *and* what is achievable in practice. Table 1 presented this contrast by deconstructing LIS disposable income, the definition recently used by ARS and others (e.g., Smeeding 2000) to compare incomes in 20 or more OECD nations.

We note that almost all of the components of more ideal income *could* be estimated by LIS were appropriate data available. That is, the LIS data template provides for a wide range of income items that, *if available* could allow the researcher to estimate a broader measure of full income. Only a few items in Table 1 are not currently counted in the LIS template, and many of these are to be added to the Wave IV LIS data structure. On the other hand, LIS disposable income as defined in the left-hand column (in the box), is now available on a consistent basis for

over 20 OECD and other nations.  The main difference between a full income measure (the more

ideal" definition) and the LIS practical definition are noncash transfers or STIK (and other

noncash income such as imputed rent on owned homes) and realized capital gains.  Each of these

is taken up in the next section of the paper. Indirect taxes (VAT, sales taxes, property and wealth

taxes, corporate taxes, and employer contributions for social security) are not discussed here.

See ABS (1999) and Harriss (1999) for Australia and British approaches to this issue, based on

their expenditure surveys.[2]

## Microdata Availability and Choices

Within the policies and practices of survey constraints, the full availability of household

income microdata should be stressed.  Cross-national comparisons of inequality and income

distribution vary enormously according to the definitions and choices made by the data analyst

and the data collector.  The unit of observation, unit of weighing, time period, income definition,

adjustments for differences in needs, classification of households, etc., are all open to choices

made by data analysts and should be subject to sensitivity tests.  This can only be achieved by

the *same* set of analysts being allowed to apply the *same* set of choices (same computer

program!) to the *same* datasets.  Once collected, the marginal cost of additional data use is zero.

Remote access systems such as LIS are being improved to the degree where privacy and

flexibility can be achieved together at very low costs.

Moreover, the availability of microdata to researchers allow further experimentation with

full income concepts by providing the basic survey material on which to build these measures in

a cooperative and consistent way.[3]  It also allows data users to begin to further refine the

classificatory variables:  occupation, work, industry, education, ethnicity, etc., which are also

important to policy relevant comparisons.   And finally, the availability of well-defined

techniques and microdata-based inequality measures for as wide a set of nations as meet

international guidelines can immediately improve the data comparability situation. These meso-data can simply be made available on a website where users can obtain a set of accurate cross-national comparisons given the definitions imposed by the creators which are completely described in the footnotes and appendices to the table.[4] These definitions, and even the *actual software program written* to extract the data and create the comparisons, should also be included as a footnote to the data series.

**Basic Standards and RARS**

The availability of microdata must be balanced by a set of specific survey information which allows the users to judge the relative quality of datasets being compared, i.e., the Robustness Assessment Report (Harris 2000) (RAR). Full cross-national comparability is impossible. What is important is the ratio of signal (true economic differences) to noise (survey sampling and nonsampling differences) and we can only judge this ratio if we have proper information of which to compare data sources. The RAR provides a beginning toward such a set of standards. Because survey design, sampling and nonsampling errors, differences in imputations for nonreporting of incomes, survey income data editing (e.g., top and bottom coding), and other features all vary by national practice, it is important to know how each survey has been constructed. Here, more detail is preferred to less, but some *minimal* level of information *must* be made available by the dataset creators.

Issues such as "grossing up" survey incomes (also known as "corrections for income underreporting") to register and administrative record amounts or can only be sensibly pursued on a cross-national basis with full information on data quality. Because current practices provide widely different levels of income quality: raw survey data; edited surveys with imputed data; surveys based on tax files; surveys based on national income registers, etc., full survey

descriptors are *crucial* to those who wish to make international comparisons.  For more on this topic, see ARS (1995, Chapters 2 and 3) and Smeeding and Weinberg (1997).

**Standardized Descriptors, Measures, and Presentations**

Along with income measures, tabular presentation of data requires that we use standard socioeconomic descriptors and socioeconomic classifications.  This is because measures of economic well-being are only important so far as they can help describe the situations of different policy relevant groups in a society.  These concepts currently used in practice are in great need of standardization.  Definitions of "single-parent families," the "unemployed," as well as occupation, industry, education, and work status are also needed.  Here, availability of microdata allows flexibility to judge the sensitivity of current practices to measurement techniques and classifications.  What needs to be laid down are a set of agreed upon classifiers, particularly those related to household structure, which all surveys should adopt and make available.  The Canberra paper by Sheridan (2000) on units of classification is a good first step.  Also needed are standard measures of inequality and standardized measurement techniques.  Top and bottom coding of data, use of equivalence scales, and related measurement issues need be carefully dealt with when presenting data summaries.  References to standard textbooks on income distribution are not enough.  Examples from such works as ARS (1995), Gottschalk and Smeeding (1997, 2000), and Jenkins (1991) are very useful here.

**Summary of Choices for Inequality Measurement**

With LIS, and with any other set of household income data, measure, choices must be made when creating income distribution statistics.  The consistency of these choices will be absolutely essential to producing comparable outcomes across countries.  With LIS, the choices are laid out and naturally flow one to another.  If we start with other approaches to comparability studies (e.g., those found in the next section of the paper), these choices need to be fully laid out

*ex ante* to data producers.  This section of the paper offers a brief summary, other chapters or

volumes are much more complete in their treatment of these topics.  The list includes:

- *Income measure* (see here and Weinberg et al. 2000) and constraints imposed by data creator (e.g., top, bottom codes; imputations, etc.).  It also involves imputations, simulations (e.g., for taxes owed—or if net income is collection and gross income is desired to estimate income before taxes), or other statistical techniques used to go from survey income reported to the selected income concept (e.g., see box in Table 1).

- *Unit of account:*  household or other income sharing unit (see Sheridan 2000).

- *Unit of observation (or weighting of observations):*  person weights (counting each person's income as one observation) or household (other unit) weights (counting each unit as one observation).  Most analysts choose the person weight, but not all, e.g., U.S. Bureau of the Census (1998).  See also ARS (1995) on weights.

- *Time period:*  annual income (though this may need to be constructed in the case of some nations, e.g., from panel datasets or from surveys covering less than a month).

- *Measure of inequality*:  choice of summary measures, presentation techniques, etc. (see Epland and Jansson 2000).

- *Equivalence scale issue:*  adjustment for differences in household size and all the issues therein addressed are important (see Buhmann et al. 1998; ARS 1995; Burkhauser et al. 1996).  A simple adjustment for differences in need according to household size (S) is recommended for most international comparisons.  Hence, measuring adjusted household income (AI), as income (I) divided by the square root of household size (or $AI=I/S^e$ where e=.5) is a good starting place.  Moreover, choice of equivalence scale may vary according to the income concept.  If one has included with DPI a set of social benefits in-kind, e.g., education expenditures per pupil or health care benefits, the equivalence scale used to adjust this new income may be different than one which is applied to cash income alone (Smeeding et al. 1993). Finally, note that choice of "no equivalence" adjustment (or e=0 in the formula above) is in effect choosing a particular equivalence scale.  The e=0 choice means that the producer implicitly assumes complete economies of scale, so that a given cash income level produces the same level of utility if it is shared by 1, 2, or 6 different persons in the household.

- *Population included*:  most household surveys on which inequality estimates include the civilian noninstitutionalized population.  However, groups such as the military, homeless, those living in foster homes, and particularly legal (and illegal) immigrants (foreign-born) may or may not be included, according to the sampling frame (household address list or national register) and national practices.

- ***Zero DPI:*** in many datasets DPI is missing because one or more components of DPI are not measured completely. Best practice differentiates "true" zero incomes (or negative incomes) from missing incomes which are coded or treated as zeros. A choice must be made to include or exclude incomplete reporters. Counting "zero" incomes is not the same as omitting these cases altogether.

See ARS (1995) and the LIS website for more on these choices in the LIS context. We now turn to other methods of making cross-national comparisons which do not have the luxury of comparable microdata and which do not therefore allow for a full range of choices.

**Experiences of Other Comparability Studies**

The only microdata based "on line" comparative database is LIS. However, based in part on LIS methods and methodologies, several agencies have gathered various types of data summaries based on methodologies similar to those employed in this paper and employed by LIS. These are both strengths and weaknesses to such attempts (e.g., OECD, 1997, 1999). First, to the extent that one can "order" comparative output tailored to fit particular purposes, these methods can be useful and low cost ways to obtain comparable income distribution definition data.

However, much confusion over terms, choices (see bullet points above), terminology, computations of aggregates, etc., may produce a less than satisfactory result. While each nation which "participates" by submitting their data can be assured of *their* choices, the choices made by *other* nations are not known. Hence, the reliability and robustness of such comparisons across nations is open to criticism and critique. In fact, the hope for output of the Canberra group will be the enhanced ability to make such comparisons, following a specific set of guidelines and comparable outcomes, without the full-time and effort required by a LIS.[5]

However, we must offer a few words of caution. Even when the choices and issues are well known and discussed, errors will occur. In the three year production of the ARS (1995) OECD volume, one country provided data according to a specified plan and later, having joined

LIS, found its estimates taken directly from LIS different than the ones earlier estimated according to the specified "plan." Another nation found that differences in estimating taxes from one year to the next created incomparable (and nonsensical) results and thus changed its national estimates to improve comparability. Yet another nation, fearing its international comparison ranking, sent LIS a "better" dataset than the one originally obtained. One final note of caution is that once formulaic results were obtained in the most recent comparison (OECD 1997, 1999), the robustness assessment technique was to compare their synthetic estimates to those obtained directly from LIS insofar as was possible![6]

**Conclusion**

The LIS offers a large set of choices and one set of guidelines for preparing internationally comparable income distribution estimates. Others offer different definitional guidelines, but produce very similar results (e.g., European Commission 2000). Because of experience with LIS, international bodies are now able to produce greater comparability using the formulaic approach, whereby a specific set of choices are requested and the data prepared according to plan by bodies from each nation. However, because all of the choices made by each nation cannot be simultaneously compared, this technique does not insure comparability. Only when nations make their microdata available to others in safe, secure ways can we be assured that the greatest degree of comparability can be obtained, subject to the limits imposed by the data themselves.

## V. Beyond DPI: What's Next for Comparable Household Income Microdata

If we want to move from the left-hand column of Table 1 to the right hand column one must confront a set of issues which are many, but not so many as to be unmanageable. As the Harrison (2000), Weinberg et al. (2000), and Church et al. (2000) papers have suggested, there

is a reasonable degree of consistency between macro and micro analysts and amongst nations based on current microdata survey practices and definitions. The "next steps" in improving comparability demand that we move beyond cash DPI and into less well-known waters. This segment of the paper discusses these waters with the greatest emphases on those issues which are liable to have the largest payoff, and suggesting the dilemmas and difficulties in producing cross-nationally comparable estimates.

**Overview**

Moving beyond cash DPI is a daunting task. The differences are both practical and theoretical. On the practical side, *none* of the 27 LIS nations' surveys contain all (or even most) of the "ideal" variables on the right-hand-side of Table 1. The reasons are straightforward. For instance consider indirect taxes. Estimates of indirect taxes paid are normally achieved by imputation and not from data collection. Surveys must measure both income and expenditure (or one of the two must be imputed) in order to allocate expenditure-based taxes across households. Moreover, these imputations depend on several "tax incidence" assumptions upon which may be little or no agreement among economists or policy analysts. For example, consider the incidence of indirect taxes on rental housing, taxes on employers, and taxes on corporations. These taxes may fall on suppliers (e.g., profits), on workers, or on demanders (consumers). Thus, stockholdings (those who earn profits) and expenditures on taxed items must also be captured by the survey in order to simulate the incidence of such taxes. Limited experimentation with these simulations (e.g., Rosenberg 1989; Rosenberg and Bell 1992; Harris 1999) indicates that these choices affect the tax burden in so far as one can agree in a pattern of incidence for such taxes and on which taxes to include. However, the EUROMOD project has made some progress on these issues (Hancock 1997). Additional anomalies also arise, e.g., eight LIS nations provide

data on "church taxes" or other (ultimately) voluntary payments made to churches or religious institutions.  Should these also be counted as taxes paid?

A second category, measurement and valuation of noncash social transfers in-kind, or STIK, provide a similar but more important set of alternatives, which are addressed below in some detail.  Estimates of noncash benefits depend on a list of benefits to include (see Table 1 for the LIS list, but other lists may vary), a measurement concept for the market value of benefits (e.g., an insurance valuation for medical services versus benefit receipt), and a measure of cash value that the household places on these benefits (market value, government cost, or recipients' value).  Each of these choices can have an impact on measured income inequality (see Smeeding et al., 1993).  Other forms of noncash income (such as imputed rent for owned houses) also need to be considered.  Here, our experience and that of others (e.g., Eurostat 1998, 2000a) is that each nation has its own definitions of this income component and even simple international rules (e.g., a low interest rate applied to housing net worth) may provide biased estimates of the value and distribution of imputed rent.  Thus, achieving these measures on an internationally comparable scale is a daunting task.

Experimental combined estimates of indirect taxes and noncash benefits have been performed for two nations, France and the United Kingdom, by Gardiner et al. (1995).  These estimates clearly indicate the importance and sensitivity to measurement technique for *each* of these components.  Others who have worked in this area report similar finds (Smeeding 1982; Harris 1999; ABS 1996).  A new project to build a European community-wide microsimulation model, EUROMOD, (see Sutherland 1997) may someday work out a set of mutually acceptable rules for estimating the incidence of indirect taxation and noncash benefits for a range of estimates.  However, this set of rules is not ready at present.

So, what do we do?  Rather than halt all comparisons, something which we know we cannot accomplish in practice, we should continue to press forward on two fronts: first, bringing all nations up to the current minimally acceptable level of comparison, i.e., cash disposable income for rich countries (and cash disposable income plus production for own consumption in poorer countries); and second, continuing to work *from the bottom up* to improve data quality, comparability, imputation, and microsimulation capabilities to deepen and broaden international comparisons.

At the same time, we can continue to improve the *quality* of the data that we use.  For instance, comparisons of micro aggregates with macro estimates of income from capital indicate that household surveys do a poor job in capturing capital (property) income flows.  Realized capital gains or losses are counted in only a few LIS surveys, and deferred/unrealized capital income is not at all captured.  Survey questions on wealth, debt and net worth, and data on flows from this wealth would allow better capture of full capital income measures (e.g., Juster and Smith 1997).  Other items, e.g., subtracting child support *paid (*as well as adding child support received) and subtracting interest *paid* (as well as recording interest received) might be a good recommendation for survey takers at this time.  Child support paid is easily collectible from both parties in the transaction (though these estimates may be at odds!) while "interest paid," if not deductible from income taxes, creates difficult survey measurement issues.

**Frontier Issues in Improved Measures of Income Inequality**

In our judgement, four issues on the forefront of household income comparison debates are key to moving forward.  These are:

A.  Better estimates of income from capital, or "property income," especially realized capital gains,

B.  Imputed rent for owner occupied housing, an item which bridges (a) above and (c) below

C. Social transfers in kind (STIK) or noncash government benefits

D. The use and development of purchasing power parity (PPP) indices to make better real income comparisons across nations.

We discuss each in turn and note here that these issues are at the forefront of others' discussions of more complete and more comparable income distributions as well as ours (e.g., Walton 1999; Eurostat 1998, 2000a).

**A. Property Income and Capital Gains.** Household surveys are notoriously bad at measuring income from capital. Our SNA statistics tell us that capital income (over and above self employment income and all other forms of earnings, but gross of undistributed institutional income, e.g., pension fund accumulations) are close to 25 percent of GDP. Yet annual household income surveys report that income from capital is 5-8 percent of total household income. If we start with annually realized capital income, the gap is between the SNA and household surveys is much smaller because many of the differences (e.g., returns to the nonhousehold sector or to pension plans) are excluded. Most of the remaining omission is due to unrealized property income and from underreporting of cash property income payments on income surveys. The quality of cash property income (gross interest, rent, and dividends) reporting is poor due to inaccurate recall, infrequent receipt or failure to sample the rich.

Household income surveys are also lax in their treatment of interest paid. Because of recent changes in asset ownership and unsecured consumer debt in nations such as the United States, both positive and negative income flows from capital have increased in recent years. A netting out of interest paid on credit card debts in the United States and in other nations, would show a negative net interest flow for many low income households (e.g., see Lupton and Stafford 2000), because their current debt payments exceed their current capital income flows.

One specific source of property income, which is increasingly important, is *realized* capital gains. While capital gains are not counted as national income for purposes of the SNA, they are of growing importance in OECD countries. Selling off assets that have risen in value can sometimes enable a household to meet its everyday needs for food, clothing, shelter, and the like, particularly among the aged. The typical treatment of *unrealized* capital gains in measuring income is to ignore them. One could, in principle, impute an income stream for those assets that do not pay interest or dividends (we address this issue more specifically below for owner-occupied housing, the largest such household asset). Such a general approach may be considered the more theoretically correct as it measures unrealized but available command over resources. But if one were mainly interested in whether a household can meet its everyday needs (as with the annual income concept) the relevant approach is to count only realized capital gains and losses. While counting realized capital gains and losses may produce large changes in income that should be prorated over a longer period with appropriate price deflators, it would be useful to both improve reporting on income from capital and to include realized capital gains and losses in our income measures.[7] Among the nations currently involved in the Canberra Group, Sweden is one of the few which currently counts realized capital gains as part of its official income definition.

One final note of caution is that once capital gains and losses are included in the annual income measure, the pattern of change in income inequality may become very uneven and pro-cyclical. For instance, consider the top line in Figure 1 for Sweden where capital gains alone made a large and uneven difference in measured inequality compared to the trend based on income net of capital gains. The addition of this income item clearly results in more instability and cyclical sensitivity in the resulting estimates of cross-national income inequality than is found in the line below it (which excludes this income component). It is also important to note

that Sweden derives its estimates of capital gains directly from income tax registers. If we were to use surveys to ask basis price and selling price, the respondent burden, high income under-sampling, and refusal issues would most certainly be large.

**B. Net Imputed Rent for Owner-Occupied Dwellings.** The issue of imputed rent is one of great importance to income distribution studies. First, imputed rent is very important in many nations, including those with below group average incomes. For instance, in Spain, 86 percent of households are homeowners (Eurostat 2000), while in other richer northern European nations (e.g., Germany) the fraction of homeowners is much smaller, around 50 percent (Smeeding et al. 1993). Second, homeownership (and owner occupation) confers an annual flow of consumption services which may offset other costs and thus deserves to be included as annual income as we define it above. Third, rental housing is often subsidized as well. If renters pay below market rents, with market rents made up by governments, there is an implicit rental subsidy in non-owned units as well. All three forms of imputed rent may be important in nations such as those of Central and Eastern Europe where "public ownership" of housing is still widespread.

The main problem is the accurate measurement of imputed rent. In theory, imputed rent is the difference between the cost of renting ones living arrangements (in a competitive market) minus the cost actually incurred in owning the home (or renting it at a below market price). Thus one needs estimates of the gross rental value of the unit, minus owners costs such as taxes, depreciation (or repair and upkeep), interest charges, property taxes, utilities, and other shelter costs. Proper estimation of imputed rent therefore requires a great deal of additional information about the unit itself (quality, size, location, unit features such as bathrooms, space, etc., are all required) if we are to estimate market rent. Further, and the owner's actual costs (taxes, upkeep,

utility charges, etc.) must also be assessed since true imputed rent is the difference between these two items. (See also Eurostat 1998 and 2000a on their approach to this issue.)

Net imputed return on the equity in one's own home could also be estimated as the annual benefit from converting one's net home equity into an annuity. If included in income, one must be careful that it is measured in a way that leads to greater international standardization instead of nation-specific measures of its value. One suggestion is to use a low government interest rate times the net value of home equity (Smeeding, et. al, 1993). Yet one must still be wary of unreasonably high land values in certain large cities (e.g., Tokyo, Hong Kong, New York) that would distort the valuation of housing services for residents there. This method, while producing "comparable" estimates may yield unreasonably high estimates of imputed rent. For instance, low income elderly homeowners in the United States who own their homes outright (no mortgage) still spend 30 to 40 percent of their incomes on shelter costs due to property taxes, repairs, utilities, upkeep, etc. (Johnson and Smeeding 2000). Thus the more complicated method of estimating market rental value net of costs might be required if the easier but "comparable" method fails to provide accurate estimates (see Smeeding 1982).

This treatment would not address the issue of equal treatment for other assets that yield unrealized capital income. Nor does it address the imputed value of other consumer durables, such as automobiles. Consistency suggests that imputed rent from these goods night also need to be counted along with that of owned homes.

Finally, the case of imputed rent, the service yielding asset can be bought and sold on real estate markets. Hence, the value to the consumer is close to the market value of the service flow, since the owner could presumably sell the housing unit and rent it back from the new owner were it more profitable to do so. While imputed rent can therefore be valued to consumers at its market price; this is not always the case with nonmarketed in-kind benefits such as social

transfers in-kind, including publicly-subsidized or publicly-owned housing which is rented at below-market value.

**C. Social Transfers In-Kind.** Canberra Group discussion in 1996 and 1998 (Smeeding and Weinberg 1998) suggested that household income should include some value for publicly provided transfers in kind, such as health care and elementary and secondary education, including early schooling (pre-school) when provided as a right of citizenship (e.g., école maternal in France) or when publicly subsidized. Most countries also give their citizens other types of in-kind social insurance benefits. The most popular are government social and health care services for the elderly, disabled and benefits for public education tied to previous government employment (e.g., educational support for veterans in the United States).

Most countries also provide in-kind assistance to their low-income populations. They provide some of these in near-cash form, such as food (food stamps in the United States) and cash housing allowances (the United Kingdom, Sweden), and one might easily count them as cash (see Section III and IV above), as we do in LIS DPI. Beyond these near-cash benefits, some other noncash social transfers in-kind are also aimed at the poor. Heating (cooling) subsidies and food subsidies are also found in some nations. These also include public housing units and related benefits in-kind, such as free health care for the poor. These benefits also differ from near-cash benefits in that they have a value to the recipient that is sometimes very hard to estimate.

A serious concern for cross-national comparisons is developing a consistent set of benefits to include, and then a consistent methodology to value these programs for recipients. All health care systems are not alike, nor are all education systems. Those who are sick should not be considered as getting more "income" from subsidized health programs than those who are not. Furthermore, the quality of programs, particularly education, is likely to vary within

countries.  Measuring the quality of universal noncash benefits and then valuing them in money terms is quite difficult on a cross-national basis (Smeeding et al. 1993).

One major concern in measuring the value of all in-kind benefits is that recipients—particularly low income recipients—may be willing to accept lesser amounts of cash income instead of noncash benefits.  In theory, one could convert these benefits to cash using their Hicksian cash equivalent value, not their market value or cost to governments (see Smeeding 1977).  However, correctly estimating this value is problematic because counterfactual behavior, i.e., unsubsidized expenditures on government-provided goods and services such as basic education or health care, are not observed.  In many circumstances, legislators have chosen to directly provide assistance for particular needs rather than providing cash that the recipient could spend how he or she wanted to.  Valuation issues arise and are magnified due to the lower cash incomes of recipients, underlining the fact that the recipient may be willing to trade the rights to his or her benefits for a lower amount of scarce cash income than the cost of those benefits.  On the other side, the accounting transparency and connectedness of SNA and income distribution statistics warrant valuing in-kind benefits at their market value (or cost to government).[8]

The authors note that already three of Canberra participant nations (United Kingdom, Denmark, and Australia) publish annual estimates of the effects of government benefits and taxes on household incomes, including health benefits, education, and housing benefits.  Indirect taxes as well as direct taxes are included to round out the picture (Harris 1999; Australian Bureau of Statistics (ABS) 1996; Jorgenson and Pedersen 2000).  In the United Kingdom these benefits amount to a full one-third of public spending (Sefton 1997), and to a roughly equal amount in Australia.  The effect of these benefits (as imputed and valued) on inequality is very large.  In Australia, the ratios of the income share of the top 20 percent to the bottom 20 percent falls from a range of 5.5 to 5.7 for disposable income to 3.0 to 3.5 for disposable income plus in-kind

benefits (Castles 2000).  In the United Kingdom the final figure is four to one (Harriss 1999). Hence, these benefits are likely to have important effects on income distribution measures, depending on how low they are valued.

In general Denmark, Australia, and the United Kingdom's studies value in-kind or social benefits at their cost to the government as did the earlier LIS study of six nations (Smeeding et al. 1993).  In general households with children (who have a large imputed education benefit) and retired households (who receive a high insurance value from health-care subsidies due to their lower average health status and greater needs for care) benefit at the expense of younger, single persons and childless couple units.  Smeeding et al. (1993) found this same result for a group of six countries including Australia and the United Kingdom.  Because social benefits such as health care and education tend to be of relatively equal cost to parties which receive them, and since their cost (or market value) is a higher fraction of income for low income households these benefits drastically reduce income inequality.  In particular single parents, larger low income families with children, the disabled, and the low income elderly benefit the most.

Unfortunately, imputing benefits at cost to low income households creates a situation where many units receive more in social transfer income in-kind than in cash income!  This creates a dilemma because most such households, if given an equal cash benefit, would spend it very differently.  The welfare implications of a $30,000 "total income" household with $15,000 worth of education benefits, $5000 worth of health benefits, and $10,000 in cash income for a single mother with three children (one of whom is disabled), compared to the same household with a completely flexible $30,000 cash disposable income, begs the question directly.

Over and above the issues laid out above, a series of methodological issues need also be addressed is a comparable way.  If we are to add benefits of in-kind transfers to households, the

large majority of this work needs to be completed by imputation to household income micro-data. Because receipt of benefits by one household or another will change their ranking in the distribution, it is nor possible to rank by cash DPI (for instance) and then just add in some "averages" for household benefits while maintaining the same ranking. Each income addition or substitution requires another ranking. Smeeding (1977a) estimates that failure to re-rank reduced measured inequality by about 25 percent. That is, the 1972 share of the bottom quintile with all benefits counted in "total" income and without re-ranking, produced a share of 8.0 percent of total income. Re-ranking reduced the share of total income to 6.0 percent.

Individual (or household) assignment of benefits also requires several difficult choices regarding conceptual and practical issues. For instance health benefits are best imputed as insurance values, not in terms of benefits actually received (i.e., the sicker you are the more you get, as with actual healthcare consumption). Insurance values vary significantly by health status, age, gender, and other factors. Most analysts argue that these differences should be taken into account.[9]

In the area of housing benefits, the value need be determined using a method consistent with the imputed rental value methods described in (B) above. For instance, the governments cost of a rent subsidy can be estimated analogously to imputed rent, i.e., by the difference between the market value of a rental unit and the amount which the tenant actually pays for that unit (Smeeding 1982). Hence, market rental values need be assigned tenants to use in conjunction with their rent paid (often some fraction of income). In the area of elementary and secondary education, one must be careful to include the variance in expenditure per pupil across geographic areas while also including some estimate of the value of school buildings, computers, and other capital inputs.

Finally, we must realize that not only does one country have to reach agreement on such methodological and practical issues, *all* countries need to agree. And then all must follow commonly accepted guidelines and formulae to achieve these measures of "market value" or "government cost" of the in-kind benefits. And finally, a method of measuring recipient values need be determined!

In this section of the paper we have taken a first look at some of the frontier issues in noncash income distribution measurement in rich nations. Many other items, e.g. wages in-kind, free use of cars by executives, other employer benefits, etc., have omitted so that attention could be focused on the "hot button" issues discussed here. One more such issue remains.

## D.　　"Relative" to "Real:" The PPP Issue

Most comparisons of income distribution across countries are presented in relative terms. That is, high and low incomes are compared to median incomes *within* countries and relative positions within nations are then compared across countries. The distance between the income at the 10$^{th}$ and 90$^{th}$ percentile, or the decile ratio is constructed using relative income within one nation and then extended to the same relative income comparisons in other nations.

However, researchers and policymakers are also concerned with absolute or "real" levels of well-being and how poverty and inequality are shaped across nations in real terms. In other words, how well off are populations in one country measured in terms of their "real" standard of living compared to people in another country?[10] In order to make such comparisons, researchers need to transform relative incomes into real incomes.

In order to transform money incomes into real incomes one needs information on relative prices and quantities of the goods and services upon which these incomes are spent. These are what are called "Purchasing Power Parities" or PPP's. Castles (1997) argues that PPP's are more

relevant to the analysis of poverty and inequality than to macroeconomic aggregates, despite the fact that macroeconomists are the primary users of PPP's.

For some time, economists have made transformations of one currency to another using PPP's to generate "real GDP per capita," for instance (e.g., Brungger 1997; Summers and Heston 1991). PPP's are therefore microeconomic concepts used with aggregate data generated primarily from National Income Accounts (NIA) coupled with surveys of "average" market baskets of goods and services in dozens (OECD) or over 100 nations (Penn World Tables, World Bank). Given these aggregated income-SNA-based PPP data, we are interested in their use with household income microdata, for instance, net disposable (after tax and transfer) income per equivalent adult, derived from household income surveys. The problem addressed in this section of the paper is how to go from one (e.g., GDP per capita) to the other (e.g., equivalence adjusted household disposable income)?[11]

Two critical observations seem in order. First, PPP's share all the problems that price indices share (e.g., differences in consumer bundles across population groups, substitution biases) magnified across 25 to 130 nations. Thus, any set of PPP's face significant intertemporal and cross-national challenges.

Second, when PPP's are scaled up from consumer bundles to national aggregates like GDP, many items over and above household consumption are included. The grossing up bias is somewhat alleviated by production of an "Individual Consumption by Households" (ICH) sub-index which excludes capital goods and collective expenditures by governments. Happily this PPP index is not to different from the GDP-PPP index which is most often used (Brungger 1997; Bradbury and Jantti 1999; Rainwater and Smeeding 1999). However, even the ICH data includes a set of goods and services such as health care, education (including early childcare) and

31

housing, which are purchased differentially by governments (via taxes) and households (via disposable income) in different countries.

Our experience therefore suggests that household income micro-data provides a less complete measure of real income than do macroeconomic concepts and measures of "total income." Because household income micro comparisons are relative to the median, not the mean; because they often use a different equivalence scale than "per capita" income; and because noncash income in the form of social transfers-in-kind are often excluded from microdata, while they are counted as "government consumption" in the PPP's, which are used to adjust them, there is considerable uncertainty in how one accurately used PPP's with microdata.[12] In general, the process of moving from a PPP adjusted mean GDP per capita (or other aggregate income measure, e.g., the ICH per capita measure of total consumption) aggregate income concept, to median disposable household income per equivalent adult, the common microeconomic concept used in household distributional comparisons (e.g., ARS 1995) suggests several important adjustments to PPP's which can dramatically affect the results.

Whatever their technical difficulties, PPP's are much preferred to exchange rates for making cross-national comparisons. If the cost of a given basket of goods can be put in common currency, the conversion of one to another gives a real purchasing power measure of the local currency, a measure which can deviate quite a bit from the exchange rate since the latter is affected not only by the domestic cost of living but also by the relative demand for a country's products, capital market and currency trading, and international trade.

Researchers have made intensive use of the statistical series on real Gross Domestic Product developed by Summers and Heston (1991) in their Penn World Tables (PWT mark 5.6). These tables give us the real GDP from 1950 to 1992 of most of the countries in the world. For example, the tables provide a measure of the real GDP per capita of each of the LIS countries as

a percentage of the real GDP of the United States in the same year.  For comparisons of the distribution of economic well-being since 1992, researchers now make use of OECD purchasing power parities because the OECD PPP's are coming to be more widely used, and continue where the PWT leaves off (past 1992).  The later 1990's OECD PPP's are based on the Summers and Heston methodology but produce somewhat different parities.   For LIS countries, the discrepancies are not large except in the case of Australia and Sweden where the OECD parities suggest buying power about 10 percent less than those of the Penn World Tables.

These experiences suggest that one must be cautious of which set of PPP's one uses and their vintage. The Summers and Heston series, PWT Mark 5.6 differ from those produced by the OECD (e.g., Brungger 1997) and those of the World Bank (e.g., Ward 1997).  The differences are due to the number of countries included (PWT vs. OECD) and/or the concept of aggregate income employed (e.g., GDP for PWT, OECD; GNP for the World Bank).  And the more recent OECD PPP's often differ from the older ones. thus, one article published in *Science* (Smeeding and Torrey 1988) based on 1980 OECD PPP's produces very different results than if one then repeats the exercise 5 to 10 years later using 1985 or 1990 OECD PPP's (e.g., Blackburn 1994). Much of this difference is no doubt due to improvements in the PPP's themselves.

Ideally one would like a series of PPP's for only the countries and years of interest to the researcher, but this is not always possible.  For some purposes the OECD series is therefore preferred to those of the World Bank because the OECD base their PPP's on OECD nations and Central and Eastern European nations only.  For a wider range of countries one should also consider the World Bank series.  One additional note concerns absolute price differences across years.  Implicit in the PPP's are a set of inflation/deflation indices that adjust prices for years between market basket measurements.  These are based, we believe, on changes in the base country's PPP's, e.g., the United States CPI.  But we feel they could also be based on changes in

national prices as well. Such a problem is particularly at issue when measuring "real" income in a nation which is experiencing hyperinflation, e.g., Russia in the 1990's, or when market baskets change greatly over time, such as in time series analyses of real income in a cross-national perspective (Atkinson et al. 2000). One must therefore be careful to select PPP's which are of the latest vintage, or of the vintage closest to the years for which the household income microdata will be compared.

Finally, our research suggests that results of using PPP's with microdata may be quite sensitive to differences in the amount of unmeasured income in surveys—the larger the unmeasured income in a country relative to that in the base country, the lower the well-being of that country will appear to be. If there were no unmeasured income in the surveys of any set of countries, and the size and distribution of in-kind income was the same, a set of PPP adjustments would all have about the same unbiased effect on a nation. Thus, it is these differences with which we must primarily concern ourselves. Smeeding and Rainwater (1999) find that, among the LIS countries whose surveys are suspected of having differentially larger amounts of unreported income—Belgium, France, Italy, The Netherlands, and Spain—ratios of LIS market income to GDP differ by up to 18 percent from the United States (as a test country) ratio of LIS market income to GDP. The average difference for most other countries is less than 5 percent. Thus, one must be careful of the measurement error in both micro and macrodata when making these comparisons (see also ARS 1995, chapter 2; Smeeding and Weinberg 1997).

In principle, PPP's should be able to overcome many of these biases in time. Clearly the OECD has carefully read the Castles Report (1997) and is making progress at improving the comparability of their PPP data for use with household income microdata. However, when using PPP's to change microdata-based disposable incomes for "real" income distribution or "real" poverty comparisons, one must tread very carefully indeed.

# VI.    Summary and Conclusion

Statisticians who produce national estimates of household income inequality begin from a different perspective than that of a national accountant.  They are charged with collecting individual household survey data on economic well-being and then using that data to produce estimates of economic inequality within their nation.  In so doing, they face a series of choices, a series of definitional issues, and a series of adjustments to data for sampling and non-sampling error.  A wide range of national priorities (and national practitioners) emerge from this process. Needless-to-say, because not everyone makes the same choices, or has the same quality data, cross-national differences will occur.

We must realize that it is not possible to produce error-free cross-nationally comparable data.  The best we can hope for is to reduce the range of difference and to improve the ratio of "signal" (true differences) to "noise" (survey practice differences) in making cross-national comparison of inequality.

The Luxembourg Income Study has made important strides in this regard.  More recently the InterAmerican Development Bank's (IDB) Latin American statisticians (Szekely and Hilgert 1999), and World Bank statisticians (e.g., LSMS surveys) are also making some headway.  The widespread availability of actual household income survey micro-data is important to further improving these estimates and their comparability, including RAR reports for all nations (Harris 2000) as well as the raw data *per se*.

We currently have a consensus base measure for international income comparisons from a LIS-based study (ARS 1995) further used in recent OECD work (1999).  This consensus measure is the basis upon which statistical offices are trying to move forward in various arenas of more complete income measurement.  This chapter has attempted to bring the reader up to the

current state of debate from the micro-data producer perspective, both in terms of comparing DPI and in terms of the frontier issues that we have decided are most central to moving forward in expanding this definition. We have not included all items, or even a list of them, which could be discussed here instead concentrating on the issues most on the forefront of measurement from a micro-data perspective. We hope that this has been a useful contribution.

# Endnotes

1. See Smeeding (1997) in Canberra (1997); Smeeding and Weinberg (1998) in Canberra (1998); Rainwater and Smeeding (1999) and Ward (1999) in Canberra (1999).

2. Note that while expenditure surveys offer the possibility of measuring expenditure, income *and* change in net worth, their respondent burden is very high. In these surveys, income measurement is often a secondary priority to expenditure weights for the CPI (not even for measuring true consumption). Thus, one must be wary of the quality of such income estimates. Yet great progress has been made in this area based on imputation and formal modeling. See Sutherland (1997) for a discussion of EUROMOD which accomplishes many of these more difficult taxes and transfer items.

3. Thus, LIS follows the practice of making national LIS datafiles available directly to the donating "nation" itself.

4. See, for instance, the LIS summary inequality measures at www.lis.ceps.lu.

5. See also Szekely and Hilgert (1999) for a set of choices imposed on Latin American datasets and the progress they have made in this regard.

6. The reader may ask: Why didn't they just use LIS? The answer is that the OECD (1999) study sought a wider range of countries, including those barred from LIS participation by national law, rules, or custom (e.g., New Zealand, Japan); those not choosing to join LIS for financial or other reasons (e.g., Turkey, some Central European nations); and a wider range of years, i.e., a longer series for time trend analyses.

7. The Eurostat (2000) working group voted positively on this issue. Other issues of periodic capital "income" such as inheritance and life insurance benefits received might also be usefully addressed in this framework.

8. Careful researchers note that in some cases market costs might exceed government costs for such goods as health insurance due to reduction of sales and marketing costs. Hence, these two concepts may also differ substantially.

9. See Smeeding et al. (1993) for a discussion of these differences. Also it may be noted that accounting for needs in estimating benefits suggests the equivalence scale used in these measures should also be sensitive to these needs.

10. Of course, the $90^{th}$, $10^{th}$, and $50^{th}$ deciles *within* a country do not change if all percentiles are changed by an equal amount once they are transformed into another currency. Having transformed all currencies into a single "absolute" currency, however, one is in a position to compare $10^{th}$, $90^{th}$, and $50^{th}$ percentiles directly across nations.

11. Readers interested in the full discussion of these issues should see Rainwater and Smeeding (1999), Castles (1997), and Bradbury and Jantti (1999, Appendix).

12. One solution suggested by Castles (1996) is that we might construct PPP's based on baskets of goods that exclude health care and education, having subtracted those expenses from incomes. The alternative is to include these incomes-in-kind (solving the measurement issues discussed in section V.C. above). Unfortunately, we do not have the data to make either of these adjustments at this time.

# References

Atkinson, A.B., A. Brandolini, and T.M. Smeeding. 2000. "Producing Time Series Data for Income Distribution: Sources, Method and Techniques," paper presented to the 26th General Conference, IARIW. Cracow, Poland, August 29.

Atkinson, A.B., L. Rainwater, and T.M. Smeeding. 1995. *Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study (LIS)*. Paris: OECD.

Australian Bureau of Statistics. 1996. *The Effects of Government Benefits and Taxes on Household Income 1993-94 Household Expenditure Survey*, Catalogue #6537.0. Canberra, Australia: Australian Bureau of Statistics.

Blackburn, M.L. 1994. "International Comparisons of Poverty and Extreme Income Poverty," *American Economic Review*, 84(2): 371-374.

Bradbury B. and M. Jäntti. 1999. "Child Poverty Across Industrialized Nations," Innocenti Occasional Paper, Social Policy Series No. 71, UNICEF-ICDC, Florence, September, Also, LIS Working Paper No. 205, September.

Brungger, H. 1997. "The Use of Purchasing Power Parities (PPP's) in International Comparisons." In *Papers and Final Report of the Expert Group on Household Income Statistics*. Canberra, Australia: Australian Bureau of Statistics, February, pp. 285-302.

Buhmann, B., L. Rainwater, G. Schmaus, and T.M. Smeeding. 1988. "Equivalence Scales, Well-Being, Inequality, and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study (LIS) Database," *Review of Income and Wealth* (June): 115-142.

Burkhauser, R.V., J. Merz, and T.M. Smeeding. 1996. "Relative Inequality and Poverty in Germany and the United States Using Alternative Equivalency Scales," *Review of Income and Wealth,* 42(4) (December): 381-400.

Canberra International Expert Group on Household Income Statistics (Canberra Group). 1999. *Papers and Final Report of the First Meeting on Household Income Statistics, Ontario, Canada, 7-9 June 1999.* Ottawa, Ontario: Statistics Canada. November.

Canberra International Expert Group on Household Income Statistics (Canberra Group). 1998. *Papers and Final Report of the First Meeting on Household Income Statistics, Voorburg, The Netherlands, 9-11 March 1998.* Voorburg and Heerlen: Statistics Netherlands, May 1998.

Canberra International Expert Group on Household Income Statistics (Canberra Group). 1997. *Papers and Final Report of the First Meeting on Household Income Statistics, Canberra, Australia, 2-4 December 1996.* Belconnen, ACT: Australian Bureau of Statistics, February.

Castles, Ian. 1997. "Comments on Papers by Heinrick Brungger," paper presented to the Canberra Group on Household Income Statistics. Canberra, Australia, 2-4 December.

Castles, Ian. 1996. "Review of the OECD-Eurostat PPP Program," STD/PPP(97)5, Economic Studies Branch, OECD, Paris, France.

Church, J., P Everaers, M. McDonald, L. Pietsch, and D. Weinberg. 2000. "Overview of the proposed Standards for Income Distribution Statistics," paper presented to the IARIW 26th General Conference. Cracow, Poland, August 29.

Epland, J. and K. Jansson. 2000. "Data Presentation Issues," mimeo, Canberra Group Meeting, LIS, Luxembourg, May.

European Commission. 2000. "The Social Situation in the European Union, 2000." Brussels: Directorate General for Employment and Social Affairs, Unit E.1, April.

Eurostat. 2000. "Seminar on Income Methodology for Statistics on Households: Draft Long Conclusions." Working Group: Statistics on Income, Social Exclusion and Poverty," DOC.E2/SEP/10/2000. Luxembourg, Eurostat.

Eurostat. 2000a. "Summary of Criticisms and Suggestions Concerning Methods of Calculating Imputed Rents," DOC.PAN144/2000. Luxembourg, Eurostat, February.

Eurostat. 1998. "Procedure for Calculating Imputed Rents." Working Group: Statistics on Income, Social Exclusion and Poverty," DOC.PAN103/99. Luxembourg, Eurostat, December.

Franz, A., D. Ramprakash, and J.W.S. Walton. 1998. "Statistics on the Distribution of Income, Consumption and Accumulation of Households (DICAH)." Report to EUROSTAT. Vienna, London, and Luxembourg, August, mimeo.

Gardiner, W., J. Hills, J. Falkingham, H. Sutherland, and V. Lechene. 1995. "The Effects of Differences in Housing and Health Care Systems on International Comparisons of Income Distribution," Welfare State Programme Discussion Paper No. 110. London, UK: STICERD, London School of Economics.

Gottschalk, P. and T.M. Smeeding. 2000. "Empirical Evidence on Income Inequality in Industrialized Countries." In A. B. Atkinson and F. Bourgignon (eds.), *Handbook of Income Distribution*. New York: Elsevier-North Holland Publishers, in press.

Gottschalk, P. and T.M. Smeeding. 1997. "Cross-National Comparisons of Earnings and Income Inequality," *Journal of Economic Literature*, XXXV (June): 633-687.

Hancock, R. 1997. "Computing Strategy for a European Tax-Benefit Model," Microsimulation Unit Discussion Paper MU9704, United Kingdom: Department of Applied Economics, University of Cambridge.

Harris, G. 2000. "Quality Assessment Reports (RAR) for Micro and Macrodata," mimeo, Canberra Group Meeting, LIS, Luxembourg, May.

Harrison, A. 2000. "Reconciliation of Micro-Macro Concepts and Terminology," paper presented to the IARIW 26th General Conference. Cracow, Poland, August 30.

Harriss, T. 1999. "The Effects of Taxes and Benefits on Household Income," *Economic Trends*, 545 (April): 27-63.

Jenkins, S. 1991. "The Measurement of Income Inequality." In L. Osberg (ed.), *Economic Inequality and Poverty*. Armonk, NY: M.E. Sharpe, Inc., pp. 3-38.

Johnson, D. and T.M. Smeeding. 2000. "Who are the Poor Elderly? An Examination Using Alternative Poverty Measures," mimeo, Center for Policy Research, The Maxwell School. Syracuse, NY: Syracuse University, January.

Jorgensen, S. and T.B. Pedersen. 2000. "Danish Income Distribution and Net Distribution to Public Finances," mimeo, Denmark Ministry of Finance, Copenhagen.

Juster, T. and J. Smith. 1997. "Improving the Quality of Economic Data: Lessons from the HRS and AHEAD," *Journal of the American Statistical Association*, 92(444): 1268-1278.

Organization for Economic Co-operation and Development (OECD). 1999. "Trends in Income Distribution and Poverty in the OECD Area," DEELSA/ELSA/WPI(99) 15, Social Policy Studies Division. Paris: OECD, September.

Organization for Economic Co-operation and Development (OECD). 1997. "Terms of Reference for a Project Analyzing the Distribution of Household Incomes, Second and Revised Version," mimeo, Social Policy Division, OECD, December 1999.

Rainwater, L. and T.M. Smeeding. 1999. "From 'Relative' to 'Real' Income: Purchase Power Parities and Household Income Microdata, Problems and Prospects." In M. Sheridan, L. Rouillard and A. Choquette (eds.), *Third Meeting on Household Income Statistics; Papers & Final Report*. Ottawa, Canada: Statistics Canada, (June): 139-163.

Rosenberg, C.B. 1989. "The Redistributive Effects of the Tax System in West Germany, Sweden, and the United States: A Comparative Tax Incidence Study," LIS Working Paper No. 36, Luxembourg Income Study, Center for Policy Research, The Maxwell School. Syracuse, NY: Syracuse University, September.

Rosenberg, C. and C. Bell. 1992. "Comparing Consistency with Simplicity When Estimating the Incidence: Alternative Assumptions and Findings for Three Countries," LIS Working Paper No. 102, Luxembourg Income Study, Center for Policy Research, The Maxwell School. Syracuse, NY: Syracuse University, September.

Sefton, T. 1997. *The Changing Distribution of the Social Wage*. London, UK: STICERD, London School of Economics.

Sheridan, M. 2000. "Units of Classification and Analysis," paper presented to the Canberra Group Meeting. Luxembourg, May.

Short, K., M. Shea, D. Johnson, and T. Garner. 1998. "Poverty Measurement Research Using the Consumer Expenditure Survey and the Survey of Income and Program Participation," paper presented at the American Economic Association meetings. Chicago, IL, January.

Smeeding, T.M. 2000. "Changing Income Inequality in OECD Countries: Updated Results from the Luxembourg Income Study (LIS)." In R. Hauser and I. Becker (eds.), *The Changing Distribution of Income*. Berlin, Germany: Springer-Verlag, in press.

Smeeding, T.M. 1997. "Cross-National Comparisons of Income Distribution: The Needs of Microdata Users as Seen from the Perspective of the Luxembourg Income Study." In *Papers and Final Report of the Expert Group on Household Income Statistics*. Canberra, Australia: Australian Bureau of Statistics, February, pp.259-284.

Smeeding, T.M. 1982. *Alternative Methods for Valuing Selected In-Kind Transfers and Measuring Their Impact on Poverty*. U.S. Bureau of Census Technical Report #50. Washington, DC: Government Printing Office, April.

Smeeding, T.M. 1977. "The Anti-Poverty Effectiveness of In-Kind Transfers," *Journal of Human Resources,* XII (Summer): 360-378.

Smeeding, T.M. 1977a. "The Economic Well-Being of Low Income Households: Implications for Income Inequality and Poverty." In M. Moon and E. Smolensky (eds.), *Improving Economic Measures of Well-Being*. New York: Academic Press, pp. 155-183.

Smeeding, T.M., P. Saunders, J. Coder, S. Jenkins, J. Fontall, A. Hagenaars, R. Hauser, and M. Wolfson.. 1993. "Poverty, Inequality and Family Living Standards Impacts Across Seven Nations: The Effect of Noncash Subsidies for Health, Education and Housing," *Review of Income and Wealth*, 39(3) (September): 229-256.

Smeeding, T.M. and Barbara B. Torrey. 1988. "Poor Children in Rich Countries," *Science*, 242 (November): 873-877.

Smeeding, T.M. and D.H. Weinberg. 1998. "Toward a Uniform Household Income Definition." Paper prepared for the Canberra Group on Household Income Measurement meeting in the Hague, The Netherlands, March; revised October 1998.

Summers, R. and A. Heston. 1991. "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1989," *Quarterly Journal of Economics*, 105(2): 327-68.

Sutherland, H. (ed.). 1997. "The EUROMOD Preparatory Study: A Summary Report; Microsimulation Unit Discussion Paper MU9705, United Kingdom: Department of Applied Economics, University of Cambridge.

Szekely, M. and M. Hilgert. 1999. "What's Behind the Inequality We Measure: An Investigation Using Latin American Data for the 1990s," mimeo. Washington, DC: IDB, December 3.

United Nations. 1977. *Provisions Guidelines for the Distribution of Consumption, Income, and Household Wealth*. Geneva: United Nations.

U.S. Bureau of the Census. 1998. "Money Income in the United States: 1997 (with Separate Data on Valuation of Noncash Benefits)," *Current Population Reports,* Series P-60, No. 200. Washington, DC: Government Printing Office.

Van der Laan, P. 2000. "The International Expert Group in Household Income Statistics ('Canberra Group'), 1996-2000 prepared for the 26th General Conference of the IARIW. Cracow, Poland, August.

Walton, J. 1999. "Social Transfers In-Kind," paper presented to the Eurostat Seminar on Income Methodology and Statistics on Households, 13-14 December, 1999, Sessions DOC.1.

Ward, M. 1999. "Update on World Bank Measures in Income Distribution." In M. Sheridan, I. Rouillard, and A. Choquette (eds.), *Third Meeting on Household Income Statistics: Papers and Final Report*. Ottawa, Canada: Statistics Canada, June, pp. 267-285.

Weinberg, D., P. Everaers, and P. Sainz. 2000. "Current Status of Countries," paper presented the Canberra Group Meeting, Luxembourg, May.

# Table 1. Summary Income Variables: Beyond DPI*

| **Existing LIS Disposable Income** | **More Ideal Income Measures: Component Choices for Next Steps** |
|---|---|
| **A. Total Earnings**<br>Gross Wages and Salaries<br>Farm Self-Employment Income<br>Non-Farm Self-Employment Income | **A′. Total Compensation**<br>+ Mandatory Employer Contributions<br>+ Non-Mandatory Employer Contributions<br>+ Earnings In-Kind |
| **B. Factor and Other Cash Income**<br>Cash Property Income (Interest Received, Rent, Dividends)<br>Private Pensions<br>Public Sector Pensions | **B′. Total Factor and Other Cash Income**<br>+ Noncash Property Income (Imputed Rent)<br>+ Realized Capital Gains/Losses<br>- Interest Paid |
| **C. Social Insurance Transfers**<br>Sick Pay<br>Disability Pay<br>Social Retirement Benefits<br>Child or Family Allowances<br>Unemployment Compensation<br>Maternity Pay<br>Military/Vet/War Benefits<br>Other Social Insurance | **C′ or D′. +Social Transfers In-Kind**<br>+Food Benefits<br>+Housing Benefits<br>+Medical Benefits<br>+Heating Benefits<br>+Education Benefits<br>+Child Care Benefits |
| **D. Income or Means-tested Transfers**<br>Means-tested Cash Benefits<br>Near-cash Benefits | |
| **E. Net Private Transfers**<br>Alimony or Child Support Received<br>Regular Private Income Transfers | - Alimony/Child Support paid |
| **F. Other Cash Income** | |
| **[G. Total Gross Cash and Nearcash Income or (A., B., C., D., E., F.)] minus:** | |
| **H. Direct Taxes**<br>Mandatory Contributions for Self-Employed<br>Mandatory Employee Contribution<br>Income Tax | + Other Direct Taxes<br>+ Other Indirect Taxes<br>+Property/Wealth Taxes |
| **I. = LIS Cash and Noncash Disposable Income (DPI)** | **I′. Full Household Disposable Income** |

*Preliminary

**Figure 1.**
**Trends in Income Inequality: Gini Coefficients in Sweden**