

# Luxembourg Income Study Working Paper Series

Working Paper No. 1

An Introduction to LIS

Timothy Smeeding, Günther Schmaus and Serge Allegrezza

June 1985

*(scanned copy)*



# "AN INTRODUCTION TO LIS - THE LUXEMBOURG INCOME STUDY"

Timothy M. Smeeding\*  
Gunther Schmauss\*\*  
Serge Allegreza\*\*\*

June 1985

LIS-CEPS Working Paper #1

Walferdange, Luxembourg

We would like to thank the Ford Foundation, the Government of Luxembourg, the Special Collaborative Program 3 at the University of Frankfurt, and Computer Resources Incorporated (CRI) for their support in completing this project. We would like to thank Al Kahn and Sheila Kammerman for helpful comments on specific elements of the LIS dataset, and also Richard Hauser and Michael O'Higgins for clarification and comment.

\*Project Director, LIS; Professor of Economics, University of Utah; and Research Associate, Center d'Etudes de Populations, De Pauverté, et de Politiques Socio-Economiques (CEPS), Luxembourg.

\*\*Technical Programming Director, LIS; Senior Programming Analyst, Computer Resources International; Research Associate, CEPS, Luxembourg

\*\*\*Research Associate, CEPS, Luxembourg.

## I. Introduction

Over the past decade the use of household income survey data in policy analysis has increased dramatically. Today the capacity to describe the effects of existing policy and simulate the effects of changes in policy is well-established in most modern nations with elaborate welfare states. However, these analyses tend to be parochial except for the fact that the techniques are similar from country to country. The next step in improving policy analysis can come from moving to a cross-national focus using comparable income surveys in a number of countries. To this end, we have assembled a databank of income surveys that can be used by scholars and policy analysts to study the effects of different kinds of programs on poverty, income adequacy in retirement, and the distribution of economic well-being generally. This project is named the Luxembourg Income Study or LIS.

The purpose of this paper is to introduce the reader to LIS. The historical background and relevance of the project are covered in Section II of the paper. Section III and two technical appendices cover the basic technical issues in arriving at comparable datasets: definitional issues (dataset, year, population coverage); income sharing units; income components and their definitions; demographic variables and their definitions. The net product of this exercise is a matrix of data comparability: common income or demographic variable definitions by country dataset.

One important issue that can only be touched upon at this juncture is that of data quality. The information that we do have is contained in section IV. Our future plans to add more datasets, to update and expand

the LIS database, and to provide worldwide access to the LIS dataset resource are covered in the concluding section of the paper.

## II. Historical Background and Project Relevance

The conception of the LIS project came during the deliberation at the Clark/Luxembourg Conference on Poverty in Fall, 1982. Some participants in that conference were highly experienced in the microanalysis of income distribution data sets for their own countries. It became apparent during the conference deliberations that it would be possible to pool the knowledge and experience in these various countries to create internally and externally consistent data sets for comparative studies which are far superior to those currently in existence.

Under the sponsorship of the government of Luxembourg, LIS was begun in April 1982. The purpose of the project was to gather in one central location, the Center for Population, Poverty and Policy Studies (C.E.P.S.) at the Institut Pedagogique, in Walferdange, Luxembourg, sophisticated microdata sets which contain comprehensive measures of income and economic well-being for a set of modern industrialized welfare states. The eventual purpose of the project is to make the datasets available for public use, with core funding to maintain and to add to the datasets in terms of additional countries, later years data, broader measures of well-being (including nonmoney income and wealth, and social indicators.

Two-day meetings in Luxembourg during April 1983 and August 1983 were directed at shaping the project. As a result of these meetings under the direction of the project director, the technical staff, and the project originators,<sup>1</sup> and with the help of the country coordinators and, since

September 1984 with the timely support of the Ford Foundation, we have prepared a databank covering seven countries -- Canada, West Germany, Israel, Norway, Sweden, United Kingdom, and the United States. The procedure of preparing the datasets is described in the next section of the paper.

The data bases which emerge from this procedure consist of income microdata sets prepared to a common plan, based on common definitions of income sources and family and household characteristics. This resource should prove extremely useful in both basic and applied social and economic research concerned with issues such as:

1. The distribution of household income and the relative income positions of the old and the young; urban and rural residents, and other groups of policy interest, e.g. single parents.
2. The distribution of earnings for both men and women, and their change over the worker's lifecycle, including the transition to retirement.
3. The effect of transfers and taxes on the distribution of economic well-being, on productive efficiency, and on consumption and savings patterns.
4. Characteristics of low income populations and the effectiveness of various kinds of transfers in reducing the incidence of low income.
5. Comparative studies of the workings of the welfare state and its policies towards the disabled.

Recent cutbacks in social program spending combined with the concurrent recession, first in England, then in the U.S., and finally in Europe have increased joblessness, poverty and inequality in Europe and in the USA. The future of the modern welfare state is in doubt. Yet no one is able to adequately compare the structure of income transfer systems across countries in order to assess their relative effectiveness in dealing with these problems. LIS would allow researchers to make such

comparisons in a straightforward manner. When LIS moves beyond the current experimental stage, it will provide a databank which can be perpetually updated and expanded to include the most recent data available for any and all nations with high quality income microdata sets which choose to participate.

### III. Basic Technical Issues

In order to produce the final harmonized LIS dataset, several procedures were required. The purpose of this section of the paper is to review these procedures, indicating the strengths, weaknesses, and limitations of each dataset and procedure.

Definitional Issues. The LIS procedure began with each country dataset expert filling out a database questionnaire which indicated the size, comprehensivity of population and detail of income components for each dataset. Each respondent listed what he or she thought were the major strengths or weaknesses of these datasets for the type of comparable social policy analyses to which we hope to contribute. Our criteria for accepting a dataset were several and tradeoffs were often made in order to arrive at a representative sample of datasets and countries. Still, several datasets were not up to standard and were, for one reason or another, excluded. Most important for each acceptable dataset was substantial detail related to income by source, especially public transfer income. In order to provide a flexible tool for analysis of the workings of the welfare state, this was the most crucial item. Also of concern was the timeliness of the dataset (i.e. nearness to 1979 -- the modal year for other surveys<sup>2</sup>), its quality (as measured by response rates and other indicators of nonsampling error), its size, the

income accounting unit, and geographical location. On the other hand, trepidation concerning just who we (the project team) were and just what LIS was all about led some governments to forbid export of suitable datasets to the LIS project center. We hope to add these datasets to LIS in the near future once our project and its aims are established. A final and overriding concern was budget. By the time the Ford Foundation entered the project in September 1984, there simply were not enough funds to add to the datasets that we had already chosen (see Section V on other feasible datasets). There was only enough money to conduct a feasibility study based on the seven country datasets currently included.

Table 1 contains an overview of the datasets which are now contained in the LIS database. Data on country, dataset name and size, income year, data sampling frame, and representativeness of the population is also included. The dataset coordinator(s) in each country are also listed. All datasets contain detailed information on income (by source), taxes, and household or family composition. All but the Canadian and German datasets are for 1979; these two are for 1981..<sup>3</sup> All but Germany, Norway, and the UK collect annual survey income data. The UK income data come from an expenditure or budget survey while in the case of Norway data comes from a sample of income tax files.<sup>4</sup> In the UK and in Germany, weekly and monthly data on earnings and usual income sources are collected along with some types of annual income (e.g. dividends, royalties, profits, etc.). The data are then adjusted to normal annual income terms based on length of time unemployed, retired, etc., and finally, this annual figure is divided by 12 (3) to calculate normal monthly (quarterly) income. Each country has considerable experience with this method (Ramprakash, 1975; Stephenson, 1980) and,

Table 1  
An Overview of LIS Datasets

Country	Dataset Name, Income Year (and Size) <sup>1</sup>	LIS Coordinators	Population Coverage <sup>3</sup>	Sampling Frame <sup>8</sup>
USA	<u>Current Population Survey, 1979</u> (69,000)	Tim Smeeding Lee Rainwater Martin Rein	97.5 <sup>4</sup>	Dicennial Census
Israel	<u>Income and Expenditure Survey,</u> 1979 (2,300)	Lea Achdut Yossi Tamir	89.0 <sup>5</sup>	Electoral Register
Norway	<u>Norwegian Tax Files,</u> 1979 (10,400)	Stein Ringen Leif Korbøl	98.5 <sup>4</sup>	Tax Records
Canada	<u>Survey of Consumer Finances,</u> 1981, (37,900)	Gail Oja Michael Love	97.5 <sup>4</sup>	Dicennial Census
U.K.	<u>Family Expenditure Survey,<sup>2</sup></u> 1979 (6,900)	Michael O'Higgins Geoffrey Stephenson	96.5 <sup>6</sup>	Electoral Register
Germany	<u>Transfer Survey, 1981<sup>2</sup></u> (2,800)	Richard Hauser Irena Stolz Gunther Schmaus	91.5 <sup>7</sup>	Electoral Register
Sweden	<u>Swedish Income Distribution - Living Survey, 1979</u> (9,600)	Peter Hedstrom Robert Erikson	98.0 <sup>4</sup>	Electoral Register

<sup>1</sup>Number of actual household units surveyed.

<sup>2</sup>The UK and German surveys collect subannual income data which is normalized to annual income levels. See text for explanation.

<sup>3</sup>As a percent of total national population.

<sup>4</sup>Excludes institutionalized and homeless populations.

<sup>5</sup>Excludes rural population (those living in places of 2000 or less), institutionalized, homeless, people in kibbutzum and guest workers.

<sup>6</sup>Excludes those not on the electoral register, the homeless, and the institutionalized.

<sup>7</sup>Excludes households with foreign-born heads, the institutionalized, and the homeless.

<sup>8</sup>Sampling frame indicates the universe from which the relevant household population sample was drawn.

subject to certain well-known limitations of virtually all survey income data sources (e.g. O'Higgins, 1980), UK and German normal monthly and quarterly income compares well with annual income estimates from other sources.

The databases vary considerably in sample size with the very large U.S. and Canadian data sets allowing for quite detailed income and demographic breakdowns, while the smaller German and Israeli datasets are problematic when examining certain specific groups, e.g. particular types of one parent families (see Hauser and Fischer, 1985). Finally, all of our datasets exclude the institutionalized and the homeless, thus reducing the sample size to about 97 or 98 percent of the national population. When measuring poverty or low income this may prove problematic (Atkinson, 1985). In addition the Israeli data excludes the rural population and those living in kibbutzum; while the German dataset excludes foreign heads of household. Again one should be aware of these differences in examining relative poverty and employment status across countries. However, comparisons of Israeli poverty and income inequality with urban only poverty and income inequality in two of our sample countries resulted in no difference in rank order between these countries. Documentation of uncounted illegal aliens or differential Census undercount of particular groups (e.g. blacks in the USA) is not available.

Income Sharing Units. One crucial element in comparing the composition of incomes within one country, much less across several countries is the issue of income accounting units (e.g. see Fiegehen and Lansley, 1975). Table 2 presents the situation with respect to the LIS countries. As in all cases, flexibility of data is a prime goal for

Table 2  
Income Sharing Units

<u>Country</u>	<u>Household</u> <sup>1</sup>	<u>Family</u> <sup>2</sup>	<u>Both</u>
Canada	X	X	X
Germany	X	x <sup>3,4</sup>	x <sup>4</sup>
Israel	X	x <sup>3</sup>	
Norway		x <sup>5</sup>	
Sweden		x <sup>5</sup>	
U.K.	X	X	X
U.S.A.	X	X	X

<sup>1</sup>"Household" = one or more persons who share common living quarters whether related or unrelated.

<sup>2</sup>"Family" = two or more persons living together (sharing common living quarters) who are related by blood, marriage, or adoption, or a single individual not living with relatives. One person families are often termed "unrelated individuals", but we call them (one person) families here for simplicity sake.

<sup>3</sup>In Israel and Germany, multiple family households can be identified but, at least in the case of Israel, we are unable to separate income sources or amounts among family members within a given household.

<sup>4</sup>We expect to be able to separately tabulate families, households, and therefore both types of units in Germany by fall 1985.

<sup>5</sup>The Norwegian and Swedish families differ slightly from the family definition given in 2 above. See text for further explanation.

LIS. We would ideally like to leave the choice of "family" (all persons living together who are related by blood, marriage, or adoption) vs. "household" (all persons, related or unrelated, who share the same living arrangements) to the researcher and to the research problem at hand. However this is not always the case. Because we do have family or household size for all datasets, the person can always be used as the unit of analysis.<sup>5</sup>

A casual glance at Table 2 indicates that, with minor exceptions, all datasets contain family data, while only five contain household data. Four datasets contain both.

The anomalies are as follows:

1. In Israel and Germany we can separate households with single families from those with multiple families, but we cannot separately identify the second family in multi-family households.
2. We can do no better than to quote Radner (1984) on the Norwegian "family": "All persons who lived in the same dwelling and had the same surname are grouped in the same family. However, one family never comprises more than one married couple. Spouses are grouped in the same family and dependents are grouped together with their supporters regardless of their surname."
3. Since 1972 (when Swedish family definition was last identical to the definition that we have above) the Swedes have defined their family as "either two adults who have lived, or normally should have lived, at least half the income year in the same dwelling irrespective of marital status and with or without children; or a single adult with or without children." (Adults are considered to be persons 18 years of age or more.)

The most significant differences are in Germany, Israel, and Sweden. The Norwegian family is very close to our family definition except for the grouping together of unmarried persons with those whom they support (or who they are supported by). Because the rationale of mutual income support is ultimately the objective of defining income sharing units in the first place, the Norwegian family definition which is based on mutual

support regardless of marital status may, in fact, be superior to the standard family definition given above.

In Germany and Israel, where we cannot individually identify families in multiple family households, the problem is not so severe as it may seem at first glance. First of all, only 2.4 percent of German households and 2.2 percent of Israeli households have multiple families. Thus less than 5 percent of families (about 4.8 percent in Germany and about 4.4 percent in Israel) live in multiple family households. Secondly, overall poverty rates and summary measures of inequality are virtually identical for single family households as compared to all (including multiple family) households in these families.<sup>6</sup>

In Sweden, the major problem is that of adult children (e.g. students) who are listed as separate families even if they are mainly dependent on their parents for income support and even if they live with them in the same household. This "adult unit" definition is based on Swedish law wherein age 18 signifies legal independence. In fact, defining poverty as equivalent disposable income less than half of median equivalent disposable income (see Smeeding, et al., 1985; Tables 4,5), of the 410,000 poor persons in Sweden, 62,000 or 15.1 percent of the poor are single persons age 18-24 who were in school at the time of the survey. The overall poverty rate of 5.0 percent in Sweden would fall to 4.2 percent were all such students nonpoor on the usual family income definition basis. However, because we are not sure how many of these persons actually live with or are otherwise supported by other persons, we cannot be sure that they are nonpoor.

Of course, given the microdata basis of LIS, the researcher is free to follow any one of several paths in dealing with these units' problems. For instance:

- a) only the three (five) countries with completely consistent family (household) definitions could be utilized. This would exclude Norway and Sweden at the least, were the household definition used, and Israel and Germany as well, were the very strict family definition used.
- b) Only consistently defined single family households could be utilized including all countries but Sweden and Norway, but excluding multifamily households in all remaining countries.
- c) Persons in families, or families per se (with the exceptions noted above), could be utilized, including all countries.

We do not think that the differences in family definition are severe enough to preclude c), but each LIS researcher can make his/her own decision of which countries to include or exclude, and on what income accounting unit basis.

Income Definitions. One major task faced by LIS was to aggregate (or to disaggregate) country-specific income elements into internationally consistent income categories. Detailed definitions of each type of income variable (V's) are given in Appendix 1. Table 3 contains basic income variable aggregations and definitions. Only general definitional guidelines are explained below; the interested reader is referred to Appendix 1 for greater detail. Before we describe these income aggregations, however, it should be stressed that international comparability and consistency rather than perfection is our goal. For instance, imputed rental value for owner-occupied homes (income variable,

V9.) is defined differently for each country with such information. Alternatively, home value (V10.) is consistently defined across five datasets. Thus the researcher can consistently define implicit rental value, for instance as some percent of home value, across the five datasets containing this information. This procedure would in effect ignore each country's own preferred implicit rent formula in favor of a definition that is consistent across all datasets.

At this time due to budgetary constraints, we have concentrated our attention on annual cash income components. Noncash income, wealth, consumption, savings, and other indicators of well-being (e.g. subjective feelings on well-offness, health indicators, neighborhood amenities, etc.) are not dealt with in any great detail. We hope to broaden our income definitions, at least to include basic components of nonmoney income: food, housing, health care and education, in the near future. But first the basic research value and feasibility of LIS must be established.

The Income Aggregations in Table 3 conform to basic OECD and/or United Nations Statistical Guidelines (1977) with little variation. Each categorization procedure involved several pieces of correspondence between the project director and one or more country coordinators, in addition to a one to three day face to face session in Luxembourg before each country's dataset was made to conform to the income variables and aggregations listed in Table 3 and Appendix 1. Based on these definitions and aggregations, several points are worth noting here:

Table 3  
Basic LIS Income Variable Definitions: Aggregation and  
Component Variables<sup>1</sup>

<u>Aggregation</u>	<u>Component Variables<sup>1</sup></u>
A1. Wages and Salaries	V1. (Wage and Salary Income)
A2. Self Employment Income (SEI)	V4. (Farm SEI) & V5. (Nonfarm SEI)
A3. Earned Income	A1. + A2.
A4. Cash Property Income	V8. (Interest, Rents, Dividends, Royalties, Annuities, etc.)
A5. Factor Income	A3. + A4.
A6. Payroll Taxes	V7. (Payroll Taxes on SEI) + V13 (Employee Payroll Taxes)
A7. Direct Taxes	V11. (Personal Income Tax) + A6.
A8. Social Insurance Transfers	V16. (Sick Pay) + V17. (Accident Pay) + V18. (Disability Pay) + V19. (Social Retirement) + V20. (Child Allowances) + V21. (Unemployment Pay) + V22. (Maternity Pay) + V23. (Military or War-Related Benefits) + V24. (Other Cash or Nearcash Social Insurance)
A9. Means Tested Transfers	V25. (Cash Payments) + V26. (Nearcash Payments)
A10. Public Cash Transfers	A8. + A9.
A11. Employment Related Pensions	V32. (Private Sector Employment Related Pensions) + V33. (Public Sector Employment Related Pensions)
A12. Private Transfers	V34. (Alimony and child Support) + V35. (Regular Private Transfers)
A13. Other Income	A11. + A12. + V36. (Other Cash Income)
A14. Gross Income	A5. + A10. + A13.
A15. Disposable (Net) Income	A14. - A7.
A16. Retirement Income	A11. + V19. (Social Retirement)
A17. Market Income	A5. + A11.

---

<sup>1</sup>Income variable definitions given in Appendix 1.

1. With respect to direct taxes, we do not count personal property or wealth taxes (V12.) as direct taxes, nor do we count the church tax (particularly popular in Germany, and reported as V14., other tax, in Appendix 1) as a direct tax. Because we are unable to separately estimate the proportion of property taxes on renters (due to the uncertain incidence of the property tax) property taxes on owner occupied homes (or other durables) are not subtracted. Church taxes are considered voluntary uses of income (user charges) by the LIS project team.
2. Employer payroll taxes are also available (V2.). Some researchers may want to define gross wages and salaries to include them because of the differential mix between employer and employee payroll taxes across countries, and because in the case of payroll taxes on self employment income (V7.), both the employee and employer portion are by definition deducted. Still, our definition of Disposable Net Income (A15. in Table 3) is net of all types of payroll tax. Also a true measure of gross employee compensation should include nonmandatory employer contributions (V3.), e.g. employer contributions to voluntary employee health or life insurance policies in the U.S. or Canada which are covered by employer and/or employee payroll taxes in other countries.
3. We define Retirement Income (A15. in Table 3) to include employment related pensions (V32. + V33.) and social retirement (V19.) because in European countries the latter are very large while the former are usually small or nonexistent. Because both types of pensions are funded by employment related contributions, either voluntary or involuntary (payroll taxes), on both employer and employee, it may be best to simply define the aggregate as retirement income. In Sweden, for instance, social retirement (Social Security) and occupational pensions for government workers cannot be separated. Since the public sector in Sweden is so large, this differentiation may prove difficult when comparing the role of occupational pension with the role of social retirement.
4. Nonrecurring lump sum payments such as realized capital gains, lottery winnings, inheritances, and/or insurance settlements are not counted as income but are recorded in the LIS dataset where available (see V37. in Appendix 1).
5. The basic differentials in transfer payments are between social insurance (or nonmeans tested) transfers and emergency (or means tested) benefits; and between cash and noncash transfers. Within this classification system, however, we have defined nearcash transfers (V24. if not means tested, or V26. if means tested). Nearcash benefits include all forms of transfers that are, in a strict sense, in-kind payments (i.e. they are tied to a specific requirement such as school attendance) but have a cash equivalent value equal to the market value. In effect these are disguised cash transfers. For instance, in the USA this includes food stamps and Low Income Energy Assistance (LIEA), which do not increase food or

energy consumption, only substituting food vouchers (or direct cash for LIEA) for cash payments that would have otherwise been made. In Germany this includes cash student allowances (which require school attendance and are means tested on parents', students', and spouses of students' incomes). In the UK, Sweden, and Germany it includes cash allowances which are supposed to help reduce rental housing costs but which do not increase housing consumption. In Germany if cash student allowances were not means tested, they would be recorded under scholarships in V24. Moreover, V24. may include other nonmeans tested near cash benefits such as training allowances. In contrast, in the U.S. and Germany, public housing benefits (i.e. for those living in a publicly owned housing unit at below market rent) are not included here. Instead, they are treated as noncash transfers (V27.) because they do not have a cash equivalent value equal to their market value. Of course, this is a fairly subjective area, and if there is any doubt that the cash equivalent value did not equal the market value, the transfer was counted as noncash transfer (V.27., V28., V29., or V31.) and not as nearcash transfer (V24. or V26.).

Additional details on income component definitions are given in Appendix 1. Several papers employ the concept of equivalent income whereby cash income is adjusted by the number of equivalent adults in a family. A full discussion of this adjustment is contained in Smeeding, et al. (1985). Finally, gross wage and salary income and hourly wage and salary income are separately reported for family or household heads or spouses (see V39., V40., V41., and V42. in Appendix 1).<sup>7</sup> For the most part these variables are obtained by dividing reported hours worked by reported wage and salary income, and so they are rough estimates of average hourly earnings.

As is always the case with income survey data, negative or zero incomes sometimes result. For instance, negative incomes may appear in cases where large business losses swamp modest (or immodest) positive income amounts. Overall zero incomes are usually a product of income nonresponse since virtually all households have some positive cash income. Each researcher is left to deal with zero or negative incomes as he or she sees fit. For six countries the percentage of families with

zero or negative incomes ranged from 0 percent (Israel) to 1.1 percent (USA), with Germany having 2.7 percent of all families with zero incomes.<sup>8</sup>

Demographic Variables. Appendix 2 contains a listing of the 30 socio-demographic variables (D's) common across LIS countries. Most of these variables are self-explanatory, containing information on the household or family (e.g. size, location, tenure, number of earners, number of children, etc.) or data on the household (family) head or spouse (e.g. age, sex, marital status, industry, period of employment, work status, etc.). Also we show variables on occupation, education and disability.<sup>9</sup> If a country dataset contained a country-specific definition of poverty, it was also added to the dataset. However, the poverty data are most assuredly not consistent across countries.

One key demographic variable (D5) controls the type of record being queried. It can take on three separate values:

D5 = 1 means that this is both a household and a family record

D5 = 2 means that this is a household but not a family record

D5 = 3 means that this is a family but not a household record

Thus, if families are desired only records with D5 = 1 or D5 = 3 are counted; if households, only records D5 = 1 or D5 = 2 are counted; and if only single family households are to be considered, records with D5 = 1 alone can be counted. Again the researcher has the flexibility to request the type of record and population analyzed (within the dataset limitations shown in Table 2).

Matrix of Comparability. The net results of this procedure are contained in Tables 4 and 5 which summarize the availability of separate

variables for each country dataset in matrix form. Separately coded variables are denoted by an S, while (income) variables that are combined with other variables in a way that does not permit direct separation by source are listed with a C.<sup>10</sup> The blank spaces indicate either that the country has no such income type or that the dataset which we employ does not contain these data (even though the country does have such an income type e.g. payroll taxes in Canada).

The first important thing to notice is that not all variables are available for all countries. For instance, if maternity/paternity allowances (V22) are important to the analysis, they can only be separately identified in the UK, Israel, and Sweden. In Norway, Germany, and Canada they cannot be efficiently separated from wages and salaries. The U.S. does not record such data separately from wages because paid maternity/paternity benefits are not prevalent.

Almost all countries contain a great deal of information on income, especially public transfer income. A dataset created by an economist in collaboration with several social policy analysts to be used primarily to study the workings of the welfare state might be expected to produce such data! All major forms of regular cash income are reported on all surveys. In general direct taxes are also well reported (except for mandatory employee contributions -- or payroll taxes -- in Canada). There is only sparse noncash benefit data at this time, but in many cases, this represents only our narrow concentration on cash income sources at this point. That is, several of our data sources contain data on receipt of public noncash transfer income, but they have not yet been made comparable mainly due to time and budget constraints. We hope to add these for some countries during 1986.

The demographic variables in Table 5 are fairly common across all datasets. In particular the age, family/household size, and other living arrangement data should make the LIS data highly popular among demographers as well as social policy analysis. However, even some of these variables require careful recoding of variables for consistency. For instance, marital status is well defined -- if the researcher wants only to identify husband-wife families as compared to others. However, if the researcher wants to find if single female heads are divorced, separated, never married, widowed, or something else (e.g. abandoned), all of these may not be available for all datasets (e.g. see Hauser and Fischer, 1985).

#### IV. Data Quality

One of the most crucial elements of comparison among countries' income data sets is relative data quality. Nonsampling errors in surveys are often many and complex. Overall survey response errors, income item nonresponse, and net income underreporting are all of concern. Here we concentrate only on the latter two problems.<sup>11</sup>

Comparative income studies are critically dependent on relative overall and item specific data quality. Unless the degree of response and net reporting errors are relatively the same, it is difficult to make accurate comparisons of income inequality or relative income positions of various groups. For instance, if a particular type of income is underreported more than another type, and if that income type is critically related to the overall degree of inequality (e.g. property income or self employment income) or to the relative income position of say the elderly vs. the nonelderly (e.g. property income or occupational

Table 4

LIS Income Variable Summary Matrix<sup>1</sup>

<u>Variable<sup>2</sup> Name &amp; Number</u>	<u>U.S.A.</u>	<u>U.K.</u>	<u>Norway</u>	<u>Canada</u>	<u>W. Germany</u>	<u>Israel</u>	<u>Sweden</u>
<b>A. <u>Earnings</u></b>							
V1. Wages & Salaries	S	S	S	S	S	S	S
V2. Mandatory Employer Contributions	S		S		S	S	
V3. Nonmandatory Employer Contributions	S	S					C
V4. Farm Self-Employed	S	S	S	S	S		S
V5. Non-Farm (Other) Self-Employed	S	S	S	S	S	S	S
V6. Earnings In-Kind		S	S				C
V7. Mandatory Contributions for Self-Employed	S	S	S		S	S	S
<b>B. <u>Property Income</u></b>							
V8. Cash Property Income	S	S	S	S	S	S	S
V9. Noncash Property Income	S	S	S			S	S
V10. Home Value	S	S	S		S	S	S
<b>C. <u>Taxes</u></b>							
V11. Income Tax	S	S	S	S	S	S	S
V12. Property/Wealth Tax	S	S	S		S	S	S

<sup>1</sup>Code: S = separately available; C = available but combined with other income sources in a way that does not permit direct separation.

<sup>2</sup>See Appendix 1 for variable definitions.

Variable <sup>2</sup> Name & Number	U.S.A.	U.K.	Norway	Canada	W. Germany	Israel	Sweden
V13. Mandatory Employee Contribution (Payroll Tax)	S	S	S		S	S	S
V14. Other Direct Taxes					S		S
V15. Indirect Taxes							
D. <u>Social Insurance</u>							
V16. Sick Pay		S		C			S
V17. Accident Pay	S	S	C	V	S		S
V18. Disability Pay	S	S	C	C	S	S	C
V19. Social Retirement	S	S	C	C	S	S	C
V20. Child or Family Allowances		S	S	S	S	S	S
V21. Unemployment Pay	S	S		S	S		C
V22. Maternity/Paternity Allowance	C	S	C	C	C	S	S
V23. Military, Veteran's War-related Benefits	S	S	C	C	S	S	
V24. Other Social Insurance		S	C	C	S	S	C
E. <u>Means-Tested or Emergency Benefits</u>							
V25. Cash Benefits	S	S		S	S	S	
V26. Near Cash Benefits	S	S	S	S	S		S
F. <u>Noncash Transfers</u>							
V27. Food Benefits	S						
V28. Housing Benefits	S					S	S
V29. Medical Benefits	S						

Variable <sup>2</sup> Name & Number	U.S.A.	U.K.	Norway	Canada	W. Germany	Israel	Sweden
V30. Heating Allowance							
V31. Education Benefits	S						
G. <u>Employer Pensions</u>							
V32. Private Employee Pension	S	S	S	C	S	S	C
V33. Public Sector Employer Pensions	S	S	C	C	S		C
H. <u>Private Transfers</u>							
V34. Alimony or Child Support	S	S	S	C	S		
V35. Other Regular Private Benefits	S	S	S	C	S	S	S
V36. Other Cash Income	S	S	S	C	S	S	
J. <u>Lump Sum Income</u>							
V37. Realized Lump Sum Income In Cash		S	S		S		
K. <u>Total or Net<sup>3</sup> Income Question</u>							
V38. Total Income Question		S			S	S	
L. <u>Other Income Categories</u>							
V39. Head's Wage-Salary	S	S	S	S	S	S	S
V40. Head's Wage Rate <sup>4</sup>	S	S		S	S	S	S
V41. Spouse's Wage-Salary	S	S	S	S	S	S	S
V42. Spouse's Wage Rate	S	S		S	S	S	S

<sup>3</sup>Refers to single separate question on total income.

<sup>4</sup>Wage rates are estimated roughly in most surveys, i.e. by dividing wages and salaries by hours worked.

Table 5

LIS Demographic Summary Matrix<sup>1</sup>

<u>Variable<sup>2</sup>Name &amp; Number</u>	<u>U.S.A.</u>	<u>U.K.</u>	<u>Norway</u>	<u>Canada</u>	<u>W. Germany</u>	<u>Israel</u>	<u>Sweden</u>
D1. Age Head	S	S	S	S	S	S	S
D2. Age Spouse	S	S	S	S	S	S	S
D3. Sex Head	S	S	S	S	S	S	S
D4. Number Persons	S	S	S	S	S	S	S
D5. Relatedness Recode	S	S	S	S	S	S	S
D6. Number Earners	S	S	S	S	S	S	S
D7. Farm-Nonfarm	S	S	S	S	S		S
D8. Ethnicity Head	S			S	S	S	
D9. Race Head	S					S	
D10. Educational Head	S	S		S	S	S	S
D11. Education Spouse	S	S		S	S	S	S
D12. Occupation Train- ing Head					S	S	
D13. Occupation Train- ing Spouse					S	S	
D14. Occupation Head	S	S	S	S	S	S	S
D15. Occupation Spouse	S	S	S	S	S	S	S
D16. Industry Head	S	S	S		S	S	S
D17. Industry Spouse	S	S	S		S	S	S

<sup>1</sup>Code: S = separately available; C = available but combined with other income sources in a way that does not permit direct separation.

<sup>2</sup>See Appendix 2 for variable definitions.

<u>Variable<sup>2</sup> Name &amp; Number</u>	<u>U.S.A.</u>	<u>U.K.</u>	<u>Norway</u>	<u>Canada</u>	<u>W. Germany</u>	<u>Israel</u>	<u>Sweden</u>
D18. Type Worker: Head	S	S		S	S	S	S
D19. Type Worker: Spouse	S	S		S	S	S	S
D20. Location (Urban- Rural)	S	S	S	S	S	S	S
D21. Marital Status Head	S	S	S	S	S	S	S
D22. Tenure: Owned- Rented Housing	S	S	S	S	S	S	S
D23. Head: Full Time- Full Year	S	S		S	S	S	S
D24. Spouse: Full Time- Full Year	S	S		S	S	S	S
D25. Disability: Head	S	S		S	S		
D26. Disability: Spouse	S	S		S	S		
D27. Number Children Under 18	S	S	S	S	S	S	S
D28. Age Youngest Child	S	S		S	S	S	
D29. Poverty Status	S					S	
D30. Poverty Income Cutoff	S	S		S		S	

pensions) cross national comparisons may be quite misleading. For this reason, it is important to discuss the issue of relative income data quality.

Table 6 represents the five conceptual levels of income reporting to which we will refer, and the level at which each LIS country dataset lies. Income concepts in the lower numbered rows are presumably more complete than are those in the high numbered rows. Moving up the rows, bottom to top, we begin with the amount of income actually reported by the population, excluding entire noninterviews but leaving partial or "item" nonresponse intact (row 6). Currently only the German data are in this state.<sup>12</sup> The next step is edited income (row 5) whereby all item nonresponse is corrected for (i.e. there are no more income nonresponses). These adjustments may take many forms: "hot-deck" imputation (U.S.), "cold-deck" imputation (Germany), or limited comparisons to administrative records (Sweden).<sup>13</sup> The U.S., Canadian, Sweden, U.K. and Israel data are in this condition. Row 2 contains the administrative amount, with the in between category in row 3 being the amount of income recorded by tax-based surveys. While not all survey income sources are taxable in all countries, we implicitly assume that taxes are more reliably reported than are survey incomes as is the usual case (Radner, 1983). In the second row, incomes are grossed up to the total amount recorded by some administrative intermediary, usually and preferably national income accounts or administrative records of government agencies. Due to tax cheating, tax based records in row 3 are somewhere in between edited survey and administrative amounts and are presumably less accurate than the latter. The Norwegian data are at this

(row 3) level. However, because of high tax rates and a very wide tax base, tax noncompliance and the black economy may be large in Norway. If this is the case, the Norwegian income data quality may be closer to that of the survey data edited amounts (row 4) than to the quality of administrative amounts (row 2). The differences between the top row, true income, and the administrative amounts are usually due to those amounts of income which are not recorded by the national accounts at all, i.e. the underground or "black" economy. Because no data set records true income, the critical income level is the second highest one, the administrative level.

Three additional comments are in order. First, in some countries, national income accounts may not be all that much better than survey data. That is, the quality of the administrative data may be in some question to begin with. For instance, wage and salary income in the Canadian Survey of Consumer Finances is 101.6 percent of the comparable National Accounts estimate (see Table 7). Secondly, before survey data (or tax data) can be compared to administrative data, the latter must be adjusted to produce estimates of identical income concepts and populations. That is, administrative data are often gross of income amounts not accruing to households and/or lump sum income amounts which are not relevant when estimating income (e.g. see Appendix 1, and income variable V37). Also, income received by those households not in the survey population, e.g. nonresidents, the deceased, and the institutionalized, must be adjusted for. These adjustments are crucial. For instance, Atkinson and Micklewright (1982) show that the 1977 comparable English "Blue Book" (National Account) estimate of

Table 6

Differential Income Data Quality:  
A Conceptual Breakdown

<u>Row</u>	<u>Income Concept</u>	<u>Differences</u>	<u>( LIS Country Dataset )</u>
1.	True Income		
		Black Economy	
2.	Administrative Record Income		
		Tax Cheating	
3.	Tax Reported Income	-----	(Norway)
		Reporting Error	
5.	Edited Survey Income	-----	(USA, UK, Sweden Canada, Israel)
		Item Nonresponse	
6.	Reported Survey Income	-----	(Germany)

occupational pensions is only 3334 million or 54.9 percent of the total administrative amount (6070 million) once these adjustments are made. Finally, rows 4, 5, and 6 are based on microdata while row 2 amounts are based on aggregate or macrodata. Thus, if one finds for instance, that in the US total wages and salaries are 97.4 percent of the adjusted administrative amount, this does not mean that all individuals have reported 97.4 percent of their true wages and salaries. Thus it must be stressed that overall readings of data quality (e.g. those in Table 7) do not provide all of the necessary ingredients for adjusting microdata for reporting errors (see Radner, 1983). In particular they do not allow the researcher to differentiate between nonreporting and underreporting (or overreporting) of individual income amounts.

Unfortunately because of the time consuming and painstaking nature of these comparisons (e.g. see Smeeding, 1982), only three LIS countries have been able to compare survey data with adjusted administrative record data at this time. These comparisons, for Canada, the UK, and the USA are shown in Table 7. In all three countries item nonresponse on survey income has been adjusted for, and reported income amounts have been weighted up to national population estimates.<sup>14</sup> The Canadian and US surveys are based on annual data, while the UK data are based on monthly data that has been annualized by a sophisticated normalization procedure (see Rampraksh, 1975).

Overall income estimates are about 90 percent of national income totals in all three surveys. However, specific item estimates often differ by some nontrivial degree. In the UK earnings are reported only at 92.8 percent rate, despite the fact that pay receipts are consulted in

about 75 percent of all cases (Kelmsley, et al., 1980). Most probably these differences reflect either nonreported secondary income (which may also be part of unreported tax income) or biases in the UK methodology for adjusting monthly earnings and/or unemployment to annual totals. The Canadian amounts are actually larger than the adjusted national account totals indicating either poor national account totals or survey sample weighting problems. Self employment income reporting does also differ substantially across surveys, though as Atkinson and Mickelwright (1982) suggest, it is often hard to estimate just what is meant by "self-employment income" either in surveys or in the national accounts. However, because self employment is less than 7.5 percent of total income in all three countries, this problem is not so great as it might be.<sup>15</sup>

Property income reporting plagues virtually all types of income surveys and the three reported here are no different in this respect. Because of its highly skewed pro rich distribution by income and by age, this differential reporting problem is to be carefully noted. For instance, adjustments for nonreporting of all types of income among the elderly in the 1973 Current Population Survey based on a record for record match with several sources of administrative data indicate that the overall incomes of the elderly would increase by 37 percent were it accurately reported as compared to about 9 percent for the population as a whole (Radner, 1983). Most of this differential was due to property income nonreporting among the high income elderly.

While occupational pension income is consistently reported across all three surveys, transfer income also differs substantially. Again, this time because of the pro-poor nature of transfer income, one must

Table 7

Quality of Income Data for Three Countries:  
Ratio of Survey Estimates to Adjusted  
Administrative Data Estimates

Income Item	Country and Year		
	Canada (1981) <sup>1</sup>	United Kingdom (1977) <sup>2</sup>	United States (1979) <sup>3</sup>
Wages and Salaries	101.6%	92.8%	97.4%
Self Employment Income	78.2	75.7	84.2
Property Income	60.5	55.3	45.1
Occupational Pension Income	85.4	83.9	81.5
Government Transfers	77.5	96.2	82.8
All Income <sup>4</sup>	92.4	89.8	89.0

Notes and Sources:

<sup>1</sup>Canada survey data from Survey of Consumer Finances for 1981; comparisons from unpublished tabulations based on family income data provided by Gail Oja, former director of Income Statistics, Statistics Canada.

<sup>2</sup>UK survey data from Family Expenditure Survey for 1977; comparisons as reported by Atkinson and Micklewright (1982) using in part methodology developed by Ramprakash (1975).

<sup>3</sup>U.S. survey data from the Current Population Survey for 1979; comparisons as reported in U. S. Bureau of the Census (1981, Table A-2).

<sup>4</sup>Based on sum of items presented above only. Some income amounts, e.g. alimony and child support or private transfers, have no administrative data to which the survey data can be compared.

carefully note these differences. While the UK monthly data do a much better job of overall transfer income reporting than do the other annual income-based surveys, there is some evidence of misreporting of transfer income by transfer type (Atkinson and Mickelwright, 1982). Also while the U.S. overall transfer income reporting rate is 82.8 percent, means tested benefits are only about 75 percent reported (Smeeding, 1982, Appendix F). In Canada, social assistance, provincial income supplements, and provincial tax credits which are also largely means-tested are only about 50 percent reported.

These differences must be carefully noted when comparing relative incomes across countries. For instance because of relatively better reporting of property income while at the same time having relatively worse reporting of transfer income, overall measures of income inequality in Canada may be more unequal than those in other countries due to relative income data quality alone. As the LIS project continues, we hope to be able to add all countries to Table 7, but at this time budget constraints prohibit a complete seven country comparison of this type.

#### V. Concluding Comments

The purpose of this paper was to introduce the reader to technical issues related to the construction of the LIS database. Additional information concerning how to use LIS can be obtained from the "LIS Users' Guide" (Rainwater and Smeeding, 1985). Those who apply to use the LIS database for a specific purpose by following the rules shown in the LIS Users' Guide can obtain additional technical details from the the LIS technical team (authors of this paper). Each user of the file will be expected to pay their own way in terms of data and computer time

necessary to prepare the specified tabulations. There are no set charges. Rather charges for data use are to be determined based on projected data use and ability to pay. These funds will be used as a self-renewing resource to replenish funds for file maintenance and continuation. All publications, reports, or papers resulting from LIS will be entered into the LIS Working Papers Series, of which this paper is the initial entry. In this way other interested researchers can see how prior analysts have chosen to utilize LIS for specific types of analyses.

While LIS is at this time only an experimental project, we hope to continue and to expand the project. Enthusiastic responses from various countries, major international research centers, and others have led us to consider expansion in the following ways:

1. Document the data file. Should we obtain funding to continue the project, our highest priority is to completely document the data file so that a complete definition of each income item in each country is automatically printed out each time a given variable and country dataset is used. Also additional information on data quality for current LIS country datasets (and expected new datasets) will be obtained.
2. Add other countries' data. Right now there are least nine additional countries who have expressed interest in joining our project and database on the same cooperative basis as the seven current LIS countries. Budgetary constraints and logistical problems have so far prohibited their addition. Given continued support, we expect to add most of the following countries over

the next year: Netherlands, Ireland, Denmark, Australia, Switzerland, Finland and possibly also Japan, France, Italy, and an additional German dataset.<sup>16</sup>

3. Expand the list of income and demographic variables. Many of the current and expected LIS datasets contain additional detailed information concerning such topics as receipt of noncash income, ages of all children, unemployment status of parents, and possibly additional demographic data on household composition. In addition, some demographic variables will require recoding for consistency, e.g. educational and occupational status. We hope to obtain separate outside research funds to pursue these objectives over the coming years.
4. Prepare a LIS macrodata file. While we realize that we cannot make the LIS microdata directly available to users by providing microdata tapes, we plan to seek funding to prepare a machine readable statistical matrix of cross tabulations that can be used as a research resource for those desiring a compact set of floppy disk data which can be used with microcomputers. This project will not begin, however, until we have had considerable experience with the larger microdataset.
5. Update (and downdate) the entire dataset. After one or two years experience with the current dataset, and assuming its continued usefulness, we hope to obtain a consortium of funds to continue the project by updating the file to include both later years cross-sectional datasets from a large list of countries, including those such as Greece and Spain which are considering

income surveys based on the LIS model, and new European Panel Study databases currently underway in Luxembourg, Holland, Belgium, France, Sweden, Germany and the USA. This latter set of data will of course add the rich potential for intertemporal panel analyses of income and socioeconomic change across countries and over time. In addition, if enough researchers are interested, we may also go back to downdate LIS by adding data from 1960 or 1970 for each country to allow for cohort analyses.

The LIS project objective from the outset was to provide a flexible research tool for the comparative analyses of incomes, sociodemographics and social policy that could be used by researchers at low cost around the world. On this basis, we encourage your comments, suggestions and use of this resource.

## Appendix 1

### LIS Income Components Definition List

The pages that follow contain the income definition component list (by number, name, and comments/description). The reader is advised that as LIS progresses a more complete and refined dataset will emerge, including additional information on the value and sources of noncash income. In addition the attached information along with greater detail on each country's definitional nuances for each variable will be transcribed to machine readable format. Future LIS users will, in other words, have all of the detailed country and variable specific information that is available. For now, short of our handwritten coding sheets, the attached list must serve this purpose.

LIS Income Components: Basic List, Description, and Aggregations

<u>Variable Number</u>	<u>Variable Name</u>	<u>Comments/Description</u>
A. <u>Earnings</u>		
	V1. Gross Cash Wages and Salaries	This includes all forms of cash wage and salary income, including employer bonuses, 13th month bonus, etc., gross of employee social insurance taxes, but net of employer social insurance taxes
	V2. Mandatory Employer Contributions for Wage and Salary Workers	These include employer contributions for all types of social insurance: social security, health, and unemployment insurance, contributions
	V3. Nonmandatory Employer Contributions for Wage and Salary Workers	These include private pension, health, life insurance, or other contributions voluntarily agreed upon by employers and workers
	V4. Gross Cash <u>Farm</u> Self- Employment Income	Farm self-employment income gross of social insurance contributions
	V5. Gross Cash <u>Nonfarm</u> Self-Employment Income	Nonfarm self employment income (sometimes called entrepreneurial income) gross of social insurance contributions
	V6. In Kind Earnings	This category is intended to measure home production or in-kind income as a <u>substitute</u> for cash wages. We count only food commodities, home grown food, board, or housing received as pay. Employer luncheon vouchers, education vouchers, medical benefits, etc. are to be counted in V3 as voluntary supplements to cash wages.

<u>Variable Number</u>	<u>Variable Name</u>	<u>Comments/Description</u>
A. <u>Earnings</u> (Cont.)	V7. Mandatory Contributions for Self Employed	All forms of social insurance contributions by the self employed: social security, medical insurance, unemployment, etc.
B. <u>Property Income</u>	V8. Cash property income	Includes cash interest, rent, dividends, annuities, royalties, etc., but excludes capital gains, lottery winnings, inheritances, insurance settlements, and all other forms of lump sum payments (see V37).
	V9. Noncash property income	Includes imputed rental income from owner occupied homes, autos, or other property.
	V10. Market Value of Own Home	Current estimated value of owned home.
C. <u>Taxes</u>	V11. Income Tax	Personal income tax liabilities
	V12. Property and Wealth Taxes	Annual property tax and/or wealth tax liabilities, not including death or inheritance taxes.
	V13. Mandatory Employee Contributions for Social Insurance (Payroll Taxes)	Mandatory employee contributions for all forms of social insurance: social security, health insurance, unemployment insurance, etc. Note that self-employment contributions are recorded in V7 above.
	V14. Other Direct Taxes	Any other type of <u>direct</u> consumption tax, church tax, etc. <sup>A1</sup>

C. Taxes (Cont.)

## V15. Indirect Taxes

Include an explanation of how these sales taxes, general in-direct consumption taxes, vat, or whatever, are calculated.

D. Government Cash Transfers: Social Insurance  
(Not Means Tested)

## V16. Cash Sickness Insurance Benefits (Sick Pay)

Include only if it can be separated from other earnings, otherwise counted in V1 as wages or salaries.<sup>A2</sup>

## V17. Cash Accident or Injury Payments (Accident Pay)

Include only short-term government stipends for injured workers

## V18. Cash Disability Benefits (Disability Pay)

Cash benefits for partial or total permanent disability (i.e. long term illness) or permanent injury.

## V19. Cash Old Age or Survivor's (Widow or Widower) Benefits (Social Retirement)

Only include here the cash Social Security benefit.

## V20. Cash Payments for Child Allowances or Family Allowances

This may include refundable tax credits as long as they are not means tested.

## V21. Cash Payments for Unemployment Insurance (Unemployment Pay)

Excludes means tested unemployment benefits, including only those that are not means tested.

## V22. Cash Payments for Maternity or Paternity Allowances

Excludes means tested or non-mandatory employer provided benefits.

<u>Variable Number</u>	<u>Variable Name</u>	<u>Comments/Description</u>
<u>D. Government Cash Transfers: Social Insurance</u> <u>(Not Means Tested) (cont.)</u>		
V23.	Military, Veterans or War Related Benefits	Cash veteran's or military benefits for old age, military disability, war separations, etc. Included also are cash benefits provided to dependents of the military, as long as they are not means tested.
V24.	Other cash or Near Cash Payments Not Listed Above (Other Social Insurance)	For instance, these would include education, training, or re-training allowances (and dependents allowance for these types of persons) as long as they are not means tested; also scholarships. <sup>A3</sup>
<u>E. Government Cash Transfers:</u> <u>Emergency (Means Tested) Benefits</u>		
V25.	Cash Emergency or Means Tested Payments (Cash Benefits)	Includes all mandatory cash transfers <u>not</u> tied to some form of in-kind benefit (e.g. not tied to food or education), including emergency assistance and benefits for long term unemployed ( <u>if</u> means tested).
V26.	"Near Cash" Means Tested Benefits	Includes all forms of transfers that are, in a strict sense, in-kind payments (i.e. they are tied to a specific requirement such as school attendance) but have a cash equivalent value equal or nearly equal to the market value. <sup>A3</sup>

<u>Variable Number</u>	<u>Variable Name</u>	<u>Comments/Description</u>
<u>F. Government Noncash Transfers</u> (Means or Nonmeans Tested)		
V27.	Food Benefits	All items are recorded at their market value or government cost. Tax subsidies and/or national price subsidies (e.g. Norwegian food price ceilings) are <u>not</u> included.
V28.	Housing Benefits	
V29.	Medical Care Benefits	
V30.	Heating or Energy Allowances	
V31.	Education Benefits <sup>A4</sup>	
<u>G. Employer Pensions</u>		
V32.	Private Pensions	These are employer payments for retirement that may (or may not) supplement Social Security. They include self employment pension plans, if they are designed to supplement Social Security.
V33.	Public Sector Pension Plans	These include pensions for public employees or self employed <u>if</u> they are designed to stand alone, i.e. not to be supplemented by some form of Social Security benefit for the aged or survivors. <sup>A5</sup>
<u>H. Private Sector Cash Transfers</u>		
V34.	Cash Alimony or Child Support	These are counted separately even if government mandated. They are otherwise not government handled (or actually government paid). <sup>A6</sup>
V35.	Regular Cash Private Interhousehold Transfers	Regular, continuing transfers from one household to another. These do <u>not</u> include one time cash gifts.

## Appendix 2

### LIS Demographic Variable Definition List

The list below defines our demographic variables. As LIS progresses each country and variable specific definition will be transcribed to machine readable form.

<u>Demographic Item</u>	<u>Variable Name</u>	<u>Description/Comments</u>
D1	Age Head	
D2	Age Spouse	
D3	Sex Head	
D4	Number of Persons in Household	
D5	Relatedness of Household members	This data item has been recoded from original data in cases where both household and family data are available. D5=1 means both a household and a family record; D5=2 means household but not family record; D5=3 means family but not household record.
D6	Number of Earners	
D7	Farm or Nonfarm Household	
D8	Ethnicity of Head	
D9	Race of Head	
D10	Years Education, Head	{ If not coded in years, recoded as minimum years needed to attain given level of education.
D11	Years Education, Spouse	
D12	Type of Occupational Training, Head	
D13	Type of Occupational Training, Spouse	
D14	Occupation of Head	{ Using ISCO codes where possible.
D15	Occupation of Spouse	
D16	Industry of Head	{ Using SIC codes where possible.
D17	Industry of Spouse	

<u>Demographic Item</u>	<u>Variable Name</u>	<u>Description/Comments</u>
D18	Class or Type of Worker, Head	Type of worker means government, blue collar, farmer, entrepreneur, etc.
D19	Class or Type of Worker, Spouse	
D20	Household Location	Urban (Metropolitan) or Rural (Nonmetropolitan)
D21	Marital Status of Head	
D22	Tenure (Owned or Rented Living Quarters)	
D23	Head Full Year, Full Time Worker?	{ These are yes-no questions.
D24	Spouse Full Year, Full Time Worker?	
D25	Disability Status, Head	{ These may either be coded yes=disabled or no=not; or degree of disability 0 percent=not; 100 percent=totally disabled.
D26	Disability Status, Spouse	
D27	Number of Children Under 18	Coded 0 if infant less than 1.
D28	Age of Youngest Child	
D29	Poverty Status of Household	{ As determined by Country-specific poverty status
D30	Poverty Income Level	
D31	Household Population Weight	{ These may all be the same weight or they may be separate, depending on the Country.
D32	Family Population Weight	
D33	Head Population Weight	
D34	Spouse Population Weight	

## Appendix End Notes

A1 Direct Consumption taxes are paid not at time of sale, but on an annual (or subannual) basis like the income tax (but excluding savings). These direct taxes are reported in V14. Indirect consumption taxes, e.g. the sales tax or value added tax, are included in V15 where such estimates are available.

A2 In Germany, for instance, sick pay appears in earnings and cannot be separated. In the USA it is not mandatory, employers have plans at their own discretion, and so even though it can be separated, it is not.

A3 For example, in the USA this includes food stamps and Low Income Energy Assistance (LIEA), which do not increase food (energy) consumption, only substituting vouchers (or direct cash for LIEA) for cash payments that would have otherwise been made. In Germany this includes cash student allowances (which require school attendance and are means tested on parents', student's, and spouses of student's incomes) and housing allowances (which do not increase housing consumption but act as partial means tested rent rebates). In Germany cash student allowances were not means tested, they would be recorded under Scholarships in V24 above. In effect V24 may include other non-means tested near cash benefits such as training allowances. In the USA, public housing benefits are not included here. Instead, they are treated as non-cash transfers (in V27 below) because they do not have a cash equivalent value equal to their market value. In Britain and Sweden they are included because they are mainly rent rebates which are capped at a certain level, but which do not increase housing consumption directly. See text for further discussion.

A4 Includes vouchers for school bills (or tuition) only. Excludes merit scholarships and student living allowances. These latter two are reported as near cash transfers, either means tested or non-means tested, in V24 or V26 above.

A5 Some employer pensions have been implicitly combined with Social Security. That is, recipients of one are not expected to receive the other. For instance, in the USA, private employee pensions (which are designed to supplement Social Security) are reported in V32. But federal government employee pensions and railroad retirement benefits are included in V33 because they were designed to exclude Social Security benefits (even though some ex-public employees and ex-railroad workers are "double-dippers" who actually get both types of benefits). In Germany and in Sweden, public employee pensions are included in V33. Government intermediated pensions for self employed persons in Germany are also in V33. But private employer pensions and life insurance pensions for retirees and survivors are recorded in V32 for Germany, because these workers also receive German Social Security.

A6 Germany "Alimony or child support advances," which are made in a few cases to those awaiting alimony or child support payments from spouses (and are later collected from those spouses), are counted here even though they are actually government payments in a strict sense.

## End Notes

<sup>1</sup>Timothy M. Smeeding of the University of Utah directed the project from its inception at the invitation of the initiators: Gaston Schaber, Director of C.E.P.S., project host in Luxembourg and Professor at the University of Liege, Belgium; Professor Dr. Richard Hauser, University of Frankfurt; Professor Lee Rainwater of Harvard University; and Professor Martin Rein of M.I.T.; and in conjunction with the several country coordinators (see Table 1). Gunther Schmaus of Germany and Serge Allegreza of Luxembourg form the technical project staff in Luxembourg.

<sup>2</sup>1979 was chosen because only 1979 data was available for Norway, while the 1979 USA Current Population Survey (CPS) dataset offered a considerable breadth of tax and noncash income sources not found in other CPS.

<sup>3</sup>Because the 1981-83 recession was only beginning in 1981, exogenous economic conditions differences in data years should have little influence on the outcome of the study. While Germany's unemployment rate for 1981, 5.3 percent, was somewhat above that in 1979, it is widely thought that the group most subject to this phenomenon, the gastarbiten (guest-workers) are largely excluded from the German dataset.

<sup>4</sup>In Norway because of the comprehensivity of the income tax base, 99.2 percent of personal money income is subject to tax. In some countries, where income transfers are largely not taxed, the income tax base is at best a partial measure of economic status. But in Norway, only local government welfare and unemployment benefits are excluded from the income tax base.

<sup>5</sup>This requires the assumption of equal intrahousehold or intrafamily income sharing (but see Lazear and Michael, 1984). Because none of the datasets contain individual income receipts for all individuals within the family, "true" individual incomes are not generally known. The one exception is earnings and other sociodemographic data for the household head and/or spouse.

<sup>6</sup>When dealing with the family unit of analysis, household based income transfers (e.g. nearcash housing allowances in several countries or Food Stamps in the USA) are prorated according to relative family size within the household. Thus in a household with one three person family and one single person family, the former would receive 75 percent of the benefit and the latter 25 percent.

<sup>7</sup>In all husband wife families the male spouse is designated the head, not because of paternalistic sexist tendencies on the part of the LIS research team, but because at least two datasets were so coded and could not be adjusted otherwise. We apologize to female family heads who live with their spouses.

<sup>8</sup>In the German case, the large majority of these records are for elderly persons who report little or no cash incomes. Each author is dealing with this problem based on available demographic (age, family

size, etc.) data concerning these records. The German datafile will be edited for total income nonresponse later this summer before the data are available for public use. See footnote 12.

<sup>9</sup>These variables have not all been consistently coded and may require some additional work by interested researchers. However, given the separate codebooks for each country dataset, this work should not be terribly time consuming.

<sup>10</sup>The reader should note that not all C items are lumped in one single category. For instance in Sweden V3 and V6 are lumped together, as are V18 and V19 but the four are not all lumped into one combined category.

<sup>11</sup>Overall survey response rates are in the 70-95 percent range except for the Israel survey which has only a 50 percent response rate.

<sup>12</sup>Of the total 2975 German family income records 54 have severe income nonresponse problems which require adjustments. These will be done later this summer. Individual paper authors have been apprised of the difficulties and individual adjustments (e.g. not counting these records and reweighting, or imputing income amounts) have been made.

<sup>13</sup>In Sweden only some nonresponse items, e.g. transfer income, are directly compared to administrative registers. In other cases, nonresponse adjustments take other forms. "Hot-decking" imputation in the U.S. is done by finding the nearest record to the nonresponse based on several criteria (age, sex, family size, etc.) and assigning the amount reported in that record to the nonresponse. Cold-decking is accomplished by calculating average amounts for those reporting an income type by age, sex, and other income levels, and assigning the average amount to the nonrespondent.

<sup>14</sup>In the UK there are no "official" population weights. The sample is randomly selected from election registers. However, there is differential overall survey nonresponse by age and region. The income totals here have been weighted to reflect differential survey response rates by age. See Atkinson and Mickelwright (1982).

<sup>15</sup>In other countries with high self-employment income, from either farm or nonfarm sources as a proportion of national income, e.g. France or Italy, this problem is much more serious.

<sup>16</sup>The German national EVS dataset for 1978 is an excellent source of income and tax data. We are negotiating with the German Bundesamt (National Central Statistical Office) to obtain a top-coded version of this dataset.

## References

- Atkinson, A. 1985. "On the Measurement of Poverty." London School of Economics, mimeo. February.
- Atkinson, A. and J. Mickelwright. 1982. "On the Reliability of Income Data in the Family Expenditure Survey 1970-1977." Working Paper #40, SSRC Programme in Taxation, Incentives, and Distribution. London School of Economics. July.
- Fiegehen, G. and P. Lansley. 1975. "Household Size and Income Unit in the Measurement of Poverty" presented to the 14th General Conference of the IARIW. Aulanko, Finland. August.
- Hauser, R. and I. Fischer. 1985. "Relative Economic Status of One Person Families with Children," LIS-CEPS Working Paper #6. July
- Kelmsley, W. et al. 1980. Family Expenditure Survey Handbook, HMSO. (London: England).
- Lazear, E. and R. Michael. 1984. "Estimating the Personal Distribution of Income with Adjustment for Within-Family Variation." University of Chicago, Mimeo. October.
- O'Higgins, M. 1980. "The Distributive Effects of Public Expenditure and Taxation: An Agnostic View of the CSO Analyses," in C.T. Stanford, et al., eds. Taxation and Social Policy (Heinemann Publishers: London):29-46.
- Radner, D. 1983. "Adjusted Estimates of the Size Distribution of Family Money Income," Journal of Business and Economic Statistics. Vol. 1, #2, April:135-146.
- \_\_\_\_\_. 1984. Family Income, Age, and Size of Unit: Selected International Comparisons. Social Security Administration, Division of Economic Research, Paper #32. Washington, D.C.: U.S. Government Printing Office. February.
- Rainwater, L. and T. Smeeding, 1985. "LIS Users Guide." LIS-CEPS Working Paper #7. July.
- Ramprakash, D. 1975. "Distribution of Income Statistics for the UK 1972-73: Sources and Methods." Economic Trends, August.
- Smeeding, T. 1982. Alternative Methods for Valuing Selected In-Kind Transfer and Measuring Their Impact on Poverty. U.S. Bureau of Census. Technical Report #50. April.
- \_\_\_\_\_. et al. 1985. "Poverty in Major Industrialized Countries." LIS-CEPS Working Paper #2. July.

Stephenson, G. 1980. "Taxes, Benefits, and the Distribution of Income," in C.T. Stanford, et al., eds. Taxation and Social Policy (Heinemann Publishers: London):15-27.

United Nations. 1977. Provisional Guidelines on Statistics on the Distribution of Income. Series M, No. 61 (New York: New York).

U.S. Bureau of the Census. 1981. "Money Income of Households and Persons in the U.S.: 1980." Current Population Reports, Series P-60, No. 132. (Washington, U.S. Government Printing Office). November.