

## **Sweden 1995: Survey Information\***

\*The information presented here is based on the Robustness Assessment Report on Income Distribution Data for Sweden by Kjell Jansson, Department for Population and Welfare, Program for Income and Wealth, Statistics Sweden, Örebro, 1998.

### **Table of contents:**

- A. General Characteristics
- B. Population, sample size and sampling methods
- C. Data collection and acquisition
- D. Definition of the survey units
- E. Availability of labor market information
- F. Quality of data

### **A. General characteristics**

#### **Name**

Inkomstfördelningsundersökningen 1995 (Swedish abbreviation HINK)  
(Income distribution Survey)

#### **Purpose**

The main purpose of the survey is to provide data for statistics on income distribution in Sweden. Seen the extensiveness of the information gathered by the survey, several other applications of the survey exist (micro simulation for income analysis, calculations with regard to the effects of changes in the tax-benefits system). The main users are the Swedish Government and Parliament and researchers at Statistics Sweden.

The survey was conducted for the first time in 1973 and it is conducted every year. Data are available 15-16 months after the end of the survey year.

Time series are available from 1975 onwards with a break in 1990/1991 due to a tax reform.

#### **Administrative Unit responsible**

Statistics Sweden  
Program for Income and Wealth  
S-70189 Örebro  
Sweden  
e-mail: [kjell.jansson@scb.se](mailto:kjell.jansson@scb.se)

Phone: +46 19 17 65 40  
Fax: +46 19 17 70 87

## **B. Population, sampling size and sampling methods**

Population size: December 31, 1996

8 844 499 individuals, 4 666 000 family units and 4 140 000 households

### Coverage of the population in the survey:

Excluded from the survey are:

- People living in institutions (hospital, care or nursing homes). If the interviewer will get in touch with relatives of a person who is living in an institution or the person him/herself, the person living in the institution will be excluded from the survey.
- People with no essential income variables in the total population register (they have most likely already left the country). Estimation of persons in this particular case: 30.000 persons
- Foreign armed forces and diplomats: only some of them are excluded from the survey and this is considered having minor effects on the distribution.

### Sample frame

The main sample frame is the total population register; this register covers all people who are registered in Sweden. It is performed by county tax authorities and used for administrative purposes. Statistics Sweden receives this register on a continuous basis and mainly uses it for population statistics. All persons having a personal identification number (personnummer, everyone who was born in Sweden and immigrants who will have or already have a permanent residence permit have a personal identification number) are included in the register. People moving out of the country to live abroad are no longer in the register.

### Sample size

There were 16260 family units included in the sample in 1996. The adults in these units were registered in the total population register both at the beginning and the end of the year in which the survey was done.

### Sample design

The sample was stratified in order to cover self-employed, farmers and the elderly. These categories were sampled with a higher fraction than others were.

### Weighting procedure

Weights are calculated in function of the sampling frame

### Changes to the survey over time

During the 80's the design was a two-year panel study. It was stratified nearly in the same way as the design in 1996. The design with two-year panels gave better estimates on income changes between two consecutive years.

There have also been some changes in design due to a co-ordination with the survey on Housing and Rents. One of the changes was the stratification of the survey into more than 25 strata in 1993 and 1995.

Also in recent years there has been an over-sampling of older people in order to estimate some costs for special services for older people. This does not affect the overall estimate of income dispersion for the whole population.

## **C. Data collection and acquisition**

Data are collected for individuals and added up to family units. Most of means-tested transfers are for the family unit but will be registered on one of the members.

No top or bottom coding has been done. Negative income originates from self-employment, this is allowed in the survey so it is possible that there is negative income after tax.

### Method of data collection

Most of the information with regard to the income variables is collected from various administrative records; a small amount of information comes from copies of tax return forms. Therefore the response rates are almost 100%.

An interview is undertaken to supplement the data from the register.

Data on household composition, employment, housing and some expenditure: housing costs, childcare, medical treatment, medicine, dental care are collected in a telephone interview. This interview is done in the period January-March following the survey year. Interviews by proxy are allowed. The response rate for these interviews has decreased during recent years.

If more than one source is used for the collection of data, both sources are linked through the personal identification number.

## **D. Definition of the survey units**

### Units of observation for income data

Income data are collected for individuals and afterwards added up to family unit level. Most of the registered means-tested transfers are paid to the family unit as a whole but they will be attributed to one of the family unit members in the dataset.

### Children

Persons between 0 and 17 years of age are defined as children. Income is assigned to each person so it is possible that income has been assigned to this category.

### Household/ family

Family unit: Cohabiting couples and singles with or without dependent children 0-17.

A family unit consists of people who permanently live together, family members living away from the family but who will be part of the unit in the near future (for example: children 15-17 attending educational institutions) are considered to be a member of the family unit. Children 18+ living in the same households as one/both of the parents are classified as single independent family units.

Students who have moved out of their parent's home to attend university or college are considered to be a separate family unit. This results in overestimated/ high (-er) poverty rates for this group.

People working away from home during the week or for short periods are considered to be a part of the family unit.

### Head of household/family

Head of the family is the person who has the highest earnings including pensions. If a person is self-employed he/she will be the head, irrespective of the income level.

The survey has so far collected data for family units, not households.

### Spouses

Relationships between persons within the household are not clearly identified, there is only information to distinguish between married, cohabiting, single and children. Neither are there any specific questions with regard to the sharing of budgets or income.

### Students

Students will be connected to the household in which they live. All income and transfers and student loans are collected. If a student lives away from the parents as a separate

household and he/she will not return, he or she will be classified as a separate household. If the student receives some transfers from the parents this will not be caught in the survey.

#### Implications

The restriction of family unit to cohabiting/married and singles with or without children 0-17 causes an overestimation of the number of economic units (in this case households) by some 12 %. Since the number of individuals is correctly estimated this will have for consequence that the average income of households will be underestimated and the income dispersion overestimated. It is estimated that around 50 % of family units in the first decile of equalized disposable income are children between 18-29 who are considered as independent units.

### **E. Availability of labor market information**

The picture of employment patterns in the income micro-dataset is consistent with the information from the Swedish Labor Force Survey or other data sources.

#### Self-employed

Income from self-employment is collected from registers and those data originates from business accounts. This information is not always very reliable since the time period in which the data are collected is mostly the same as a calendar year while some business accounts may have the bookkeeping period July-June. It is assumed that the income from self-employment is underestimated, self-employment income is also reported gross of taxes.

#### Working time

For people in the non-response, some rough estimates for working time have been made. Those who have earnings over a certain amount will be classified as full time earners.

### **F. Quality of data**

The National Accounts data and the Total Population register data compared.

Income Component	Grossed estimate from micro-data as a % of National Accounts	Comments	Implications for income distribution estimates
Wages and salaries	98%	NA uses the same data + an estimated addition of 2%	No opinion, the most will be in the black economy

Self-employment income	40%	There's a common opinion that there's a large part of black economy which is measured in NA but not in the survey	In average, people with income from self-employment will be in the lower part of the distribution, income about 75% of blue - collar worker. Dispersion will be overestimated
Occupational pensions	110%	Some minor differences in classification between occupational pensions and other pensions are the main causes of the difference NA/micro data	None
Investment income	(a) 96% (b) 145%	(a) just looking at rents, interests and dividends (b) including capital gain, this is included in micro data but not in NA	An expected increase of investment income abroad will not be counted, if not registered. Minor increase on dispersion would probably be if all investment income was measured
Transfers	(a) 105% (b) 101%	(a) private pension schemes and student loans are included in micro data but not in NA (b) If excluded private pension and student loans	
Other income		Inter household transfers and inheritance are not monitored in neither	Probably the dispersion would be lower if included

		NA or micro data	
Direct taxes	100%	No differences at all	

Source: Income Distribution Data for Sweden, Robustness Assessment report, Kjell Jansson, Statistics Sweden, December 1998.

#### Accuracy of income data

Since all income data is collected from administrative registers we could say that none of the income information was collected by proxy.

There was no imputation of income data.

#### Non-response

There are some non-responses biases on the estimate of family types. 23 % did not answer the telephone interview and for this non-response the survey adopts the classification in the registers. This will lead to an overestimate of single family units and single parents. No re-weighting is applied so this non-response will probably lead to an overestimate of the number of household as well as income dispersion.

All income variables are collected in the same way regardless if people answered the telephone interview or not. All units are in the survey. However, for some groups there might be some income which is not registered.

- Young people who have moved from their parents' home may continue to receive some transfers from their parents, however these transfers are not found in the administrative registers.
- Self employed people; the problem here is underreporting of income and possible underreporting to the tax authorities. There is no separate information available on the tax evasion but there's a large difference between operating surplus in NA and the equivalent amount in micro data.
- Immigrants, it is known that immigrants more often borrow money from relatives than from the bank. This indicates that there might be transfers between relatives and that some of the loans may be changed to gifts.

Income variables reflect the income that is known by the authorities. Income from black economy is not included or imputed.