ETLA ELINKEINOELÄMÄN TUTKIMUSLAITOS
THE RESEARCH INSTITUTE OF THE FINNISH ECONOMY
Lönnrotinkatu 4 B, 00120 Helsinki 12, Finland, tel. 601322

Keskusteluaiheita
Discussion papers

Fredrik Nygård*

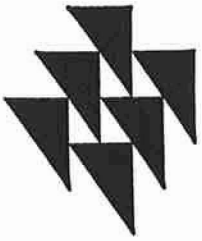
RELATIVE INCOME DIFFERENCES

IN FINLAND 1971-1981

No 185 18.11.1985

* Department of Statistics,
Åbo Akademi

ISSN 0781-6847



ABSTRACT

The paper deals with relative income differences among households in Finland 1971, 1976, and 1981. The analysis is based on micro data from the household budget surveys conducted by the Central Statistical Office.
After first graphical presentations of the income differences, the differences are related to other household characteristics. A pooled cross-section regression model is finally applied.

1. INTRODUCTION

The leading idea of this paper is to try getting 'behind' the observed Finnish income distribution and to adopt a principally dynamic approach to illustrate how a number of background variates have influenced the income position of a household.

Before spelling out the empirical framework in any greater detail, it should be noted that the discussion will rely heavily on the following concepts:

(A) A main reason for attaching significance to the income distribution can be traced back to its welfare implications (welfare depends on consumption possibilities and these in turn on income). When discussing income distribution matters, the distribution of available income, closely related to the notion of consumption possibilies, should consequently be placed in the forefront.

(B) From a welfare point of view the appropriate income recipient unit seems to be the family/household, as members of a family may share their consumption goods and pay for the purchases from their joint income. Hence, even if incomes are earned largely by individuals, the 'consumption possibility' approach focuses on the pooled incomes, which families/households control.

(C) It would be highly preferable if we could follow the development of annual incomes from a dynamic perspective,

instead of confining the discussion to a set of annual 'snapshots'. A genuine dynamic analysis of trends in the income distribution would require panel data, following the same families/household and their incomes over a number of years. However, panel data on incomes in Finland is currently not available.
On the other hand, the possibility of deriving average income profiles from cross-section income data covering several years (by disaggregating the income recipients according to age) should not be overlooked. A similar approach will be chosen below.

(D) When comparing the income positions between households we will mainly confine the discussion to their relative position in the income distribution.
Hence, if y_{it} denotes the income of the i^{th} household during year t and μ_t is the corresponding annual mean income in the whole household population, we will derive the distribution of incomes relative to the mean, y_{it}/μ_t, or log differences of household incomes from the mean, $\log(y_{it}/\mu_t)$, from the original y_{it}-distribution.

(E) The empirical analysis of income differences below will be based on survey sample data. As a consequence, the results must be interpreted as estimates involving sampling errors. However, the adopted data base lacks sufficient information for a rigorous treatment of the induced sampling errors and the analysis will mainly be carried out without any reference to the precision of the estimates.

2. THE DATA

An analysis of the personal income distribution in Finland during the last decade(s) may principally be based on two data sources:

(a) the taxation records compiled by the National Board of Taxation (Verohallitus),

(b) the household and income distribution surveys conducted by the Central Statistical Office (Tilastokeskus).

At first sight the taxation records, covering the time period from 1920 onwards (with only a few exeptions), may appear as an appropriate base for studying trends in the income distribution. However, the records depend heavily on current fiscal legislation (exemption limits and tax-exempt income components, joint/separate taxation of married couples, etc) and many shifts in reported incomes can be traced back to changes in the taxation rules. Moreover, as the taxation record uses the individual as the income recipient, records have to be matched if we wish to study the income distribution among families/households.

The income distribution surveys have been conducted annually since 1977, with forerunners in the household budget surveys 1966, 1971, and 1976, and are based on samples consisting of some 10000 households. Besides the influence of sampling errors, the comparability between these surveys are to some extent restricted by some changes in the adopted income concepts and sampling procedure. Yet, the income distribution surveys (at least the later editions) must be considered as

the most reliable Finnish data source to a study of income differences between households.

Table 2.1a summarizes the development of average available household income, transformed to real 1984 standards by the cost-of-living index, derived from the 'official' estimates published by the Central Statistical Office (CSO). The Table, based on reults from the household budget surveys 1966, 1971, 1976, and 1981, and the income distribution surveys 1977-80, adopts the 'traditional' CSO-classification of households into socio-economic groups.

Table 2.1a in here

Table 2.1b in here

As can be seen from the Table, the real household income has developed according to an increasing profile. This increasing tendency is further accentuated when examining the available household income per capita in Table 2.1b, to be explained by the simple fact that the households have on average grown smaller over the period 1966-1981. We return to this point in Section 5.

In Figure 2.1a and 2.1b the level of real available income within the socio-economic groups are represented graphically for the three years 1971, 1976, and 1976.

Figure 2.1a in here

Table 2.1a Finland 1966-1981: Real (1984 standards) available income per household; 1000 FIM.

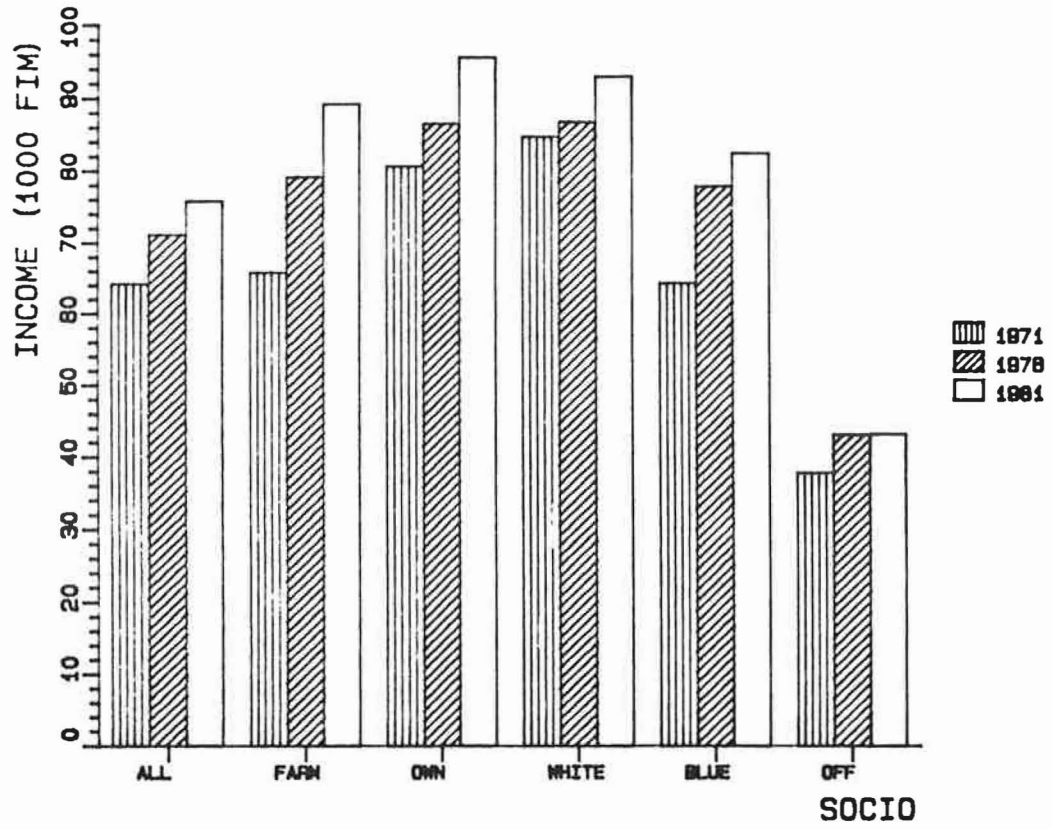
	1966	1971	1976	1977	1978	1979	1980	1981
All households	57.5	64.2	71.7	70.2	71.6	72.8	72.8	75.8
Farmers	50.6	65.6	79.1	82.3	84.9	85.4	94.6	89.2
Own-account	77.6	80.6	86.5	84.6	93.9	93.3	92.0	95.6
White collar	84.4	84.7	86.8	87.4	89.3	93.5	91.8	93.0
Blue collar	54.5	64.3	77.8	76.7	77.4	79.7	78.9	82.4
Inactive	30.5	37.8	43.1	38.3	40.2	37.4	38.2	43.2

Table 2.1b Finland 1966-1981: Real (1984 standards) available household income per capita; 1000 FIM.

	1966	1971	1976	1977	1978	1979	1980	1981
All households	17.2	21.6	25.8	26.4	27.3	28.3	28.8	29.2
Farmers	11.0	15.9	20.0	21.4	22.0	22.6	24.9	24.0
Own-account	20.5	21.9	24.4	25.1	28.3	28.6	27.9	29.1
White collar	25.7	28.9	30.6	31.7	32.0	33.5	33.7	34.6
Blue collar	16.2	20.0	25.0	25.3	26.3	27.2	27.8	29.0
Inactive	16.0	19.4	24.4	23.8	24.4	23.5	24.4	26.3

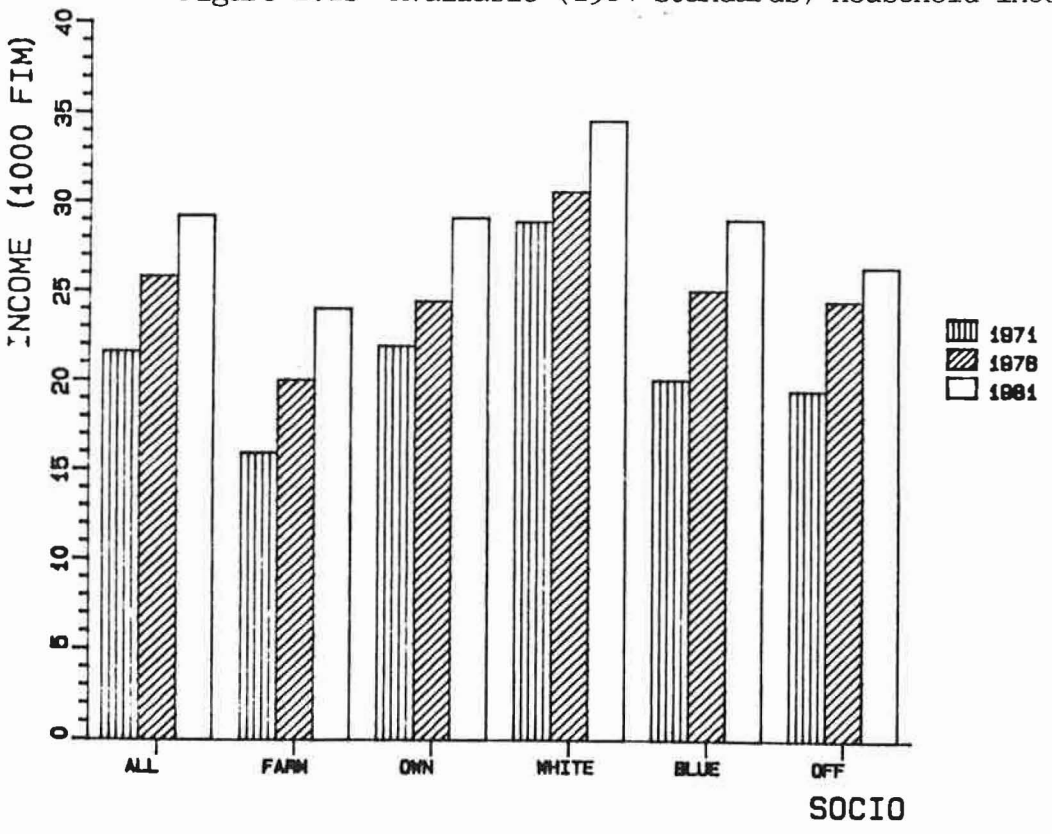
Sources: Household Budget Survey for 1966
Household Survey 1971, 1976, 1981
Income Distribution Statistics 1977-1980.

Figure 2.1a Available (1984 standards) household income.



REAL (1984 STANDARDS) AVAILABLE INCOME PER HOUSEHOLD 1971-1981.

Figure 2.1b Available (1984 standards) household income per capita.



REAL (1984 STANDARDS) AVAILABLE HOUSEHOLD INCOME PER CAPITA 1971-1981.

Figure 2.1b in here

The data base to be adopted in this paper consists of micro-data from the samples used by the CSO in the household budget surveys (kotitaloustiedustelut) 1971, 1976, and 1981, with effective sample sizes of 8816, 7971, and 7368 households, respectively. The survey designs were based on traditional probability sampling, without any panel-type links between the three samples, so that the 1971, 1976, and 1981 samples all include different households (possibly with a few exceptions due to pure chance).
The survey households were defined in a multigenerational setting as consisting of all persons who live together and jointly spend their income. Hence, in addition to the nuclear family a household may include other persons (presumably grandparents).

In co-operation with the CSO, the original 1971 and 1976 sample data was reworked (unifying a number of classification rules and including/excluding some income components) in order to improve the internal comparability between 1971 and 1976 microdata. However, the 1981 sample is included without any corresponding modifications. As a consequence, the data from 1971 and 1976 are not strictly comparable to the 1981 records, the major difference being a slightly narrower income coverage in the 1981 data. Even if the transformation to relative incomes, frequently used in the sequel, may be expected to improve the comparability (cf Parkkinen [1985]

who applies a similar argument), the differences in coverage should be kept in mind when interpreting the results of the following sections.

Figure 2.2a gives a first outline the distribution of available household income in nominal amounts according to our data base. Note that the 1981 distribution, due to its slightly different income coverage, probably should be shifted somewhat upwards.

Since the changes in the shape of the distribution (growth of the average nominal income accompanied by increasing dispersion) may be linked to the inflationary development over the decade 1971-81, Figure 2.2b presents the three distributions with the horizontal axis transformed to real available household income. The transformation is based on the cost-of-living index, with 1984 as a reference standard, implying that 1971 incomes should be multipied with a factor 3.911, the 1976 incomes with a factor 2.065, and the 1981 incomes with 1.271. As can be seen, there are considerably less differences between the distributions according to this representation.

Figure 2.2a in here

Figure 2.2b in here

Figure 2.2a The distribution of nominal available household income in Finland 1971, 1976, and 1981.

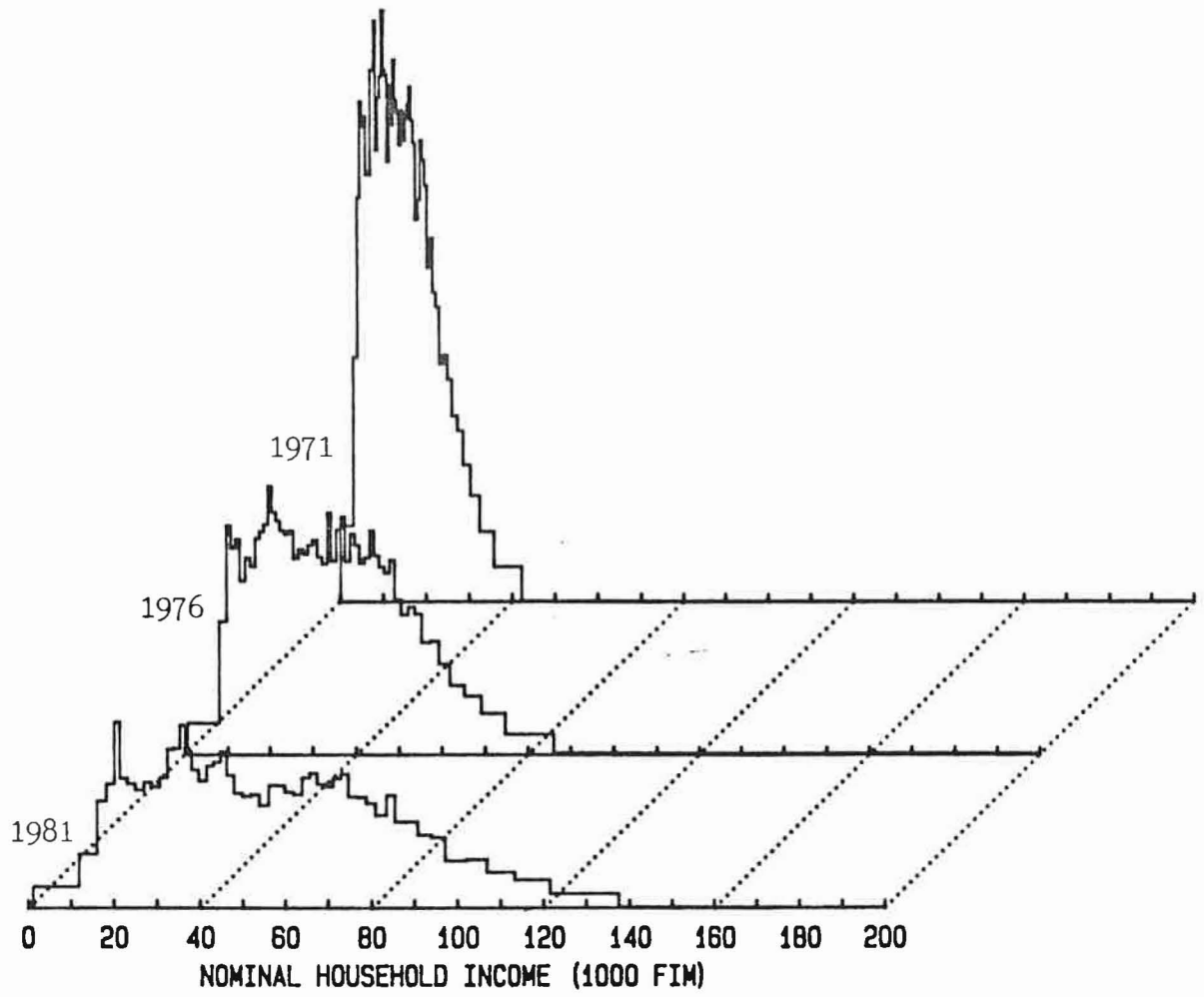
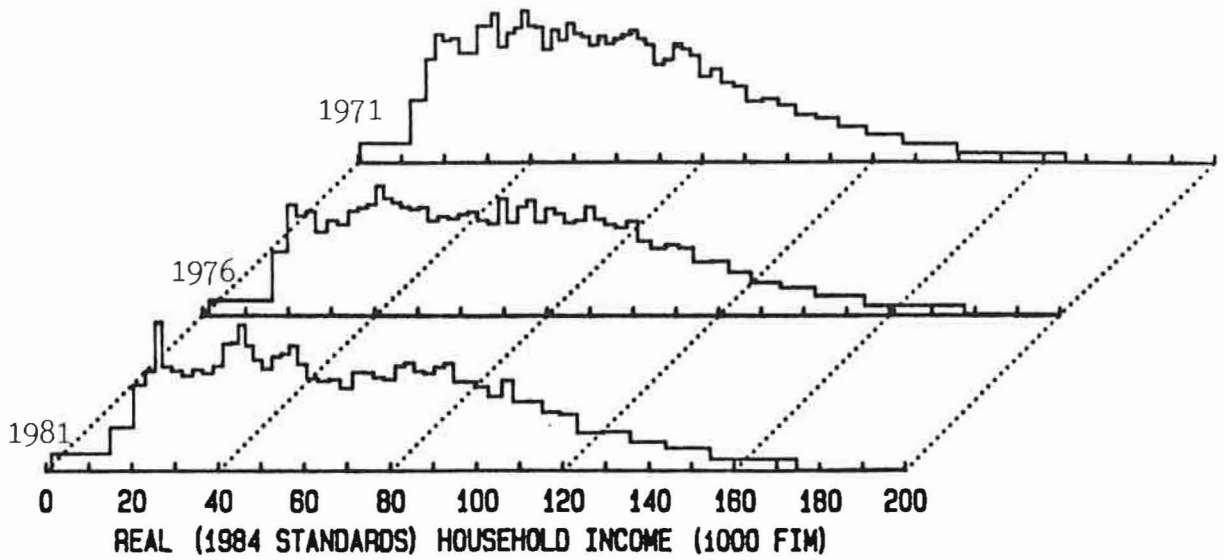


Figure 2.2b The distribution of real available household income (1984 standards) in Finland 1971, 1976, and 1981.



3. RELATIVE INCOME DIFFERENCES - GRAPHICAL DISPLAYS

As noted in the introduction, this paper will primarly deal with relative income differences. Now, if y_{it} denote the annual income of a household in year t, its relative income is given by $u_{it} = y_{it}/\mu_t$, where μ_t is the average income in the household population during the same year, and most measures of income inequality (e.g. the Gini coefficient) may be interpreted as a weighted sum of the u_{it}'s. However, any monotone transformation of the relative income u_{it} could as well serve as an indicator of the relative income position of the household. In this paper we will frequently use the transformation

$$z_{it} = 100\log(u_{it}) = 100\log(y_{it}/\mu_t) = 100(\log(y_{it})-\log(\mu_t))$$

for descriptive purposes. The z_{it}'s, principally the log differences between household incomes and the mean income, will be referred to as the relative incomes in log-scale, and their use is mainly motivated by the fact that a log-scale in many cases faciliates a graphical interpretation of income distribution data. In this context it should be noted that z_{it} is close to zero for households with an income close to the annual average, and that the magnitude of z_{it} roughly may be interpreted as the difference, expressed as a percentage, between the household income and the mean income.

Figures 3.1a and 3.1b give a first illustration of the relative income differences in log-scale. In Figure 3.1a the distributions of pre-tax household income are outlined, and Figure 3.1b represents the corresponding distributions of

available household income.

Figure 3.1a in here

Figure 3.1b in here

The Figures show three points clearly: First, the main difference between the distribution of pre-tax income and that of available income is that the latter is more peaked. Second, the income distributions from 1971, 1976, and 1981 are on the whole remarkably similar. Third, from 1971 to 1976 the distribution seems to have concentrated to incomes slightly above the mean income at the expense of high incomes particulary, without any corresponding trend being evident between 1976 and 1981, and this applies to both the distribution of pre-tax and available income.

Figures 3.2a and 3.2b, in turn, illustrate the the income distributions in the traditional form of Lorenz diagrams.

Figure 3.2a in here

Figure 3.2b in here

As implied by the Lorenz curves, both the distribution of pre-tax income, Figure 3.2a, and the distribution of available income, Figure 3.2b, have from 1971 to 1976 developed towards

Figure 3.1a Pre-tax household income: relative income differences.

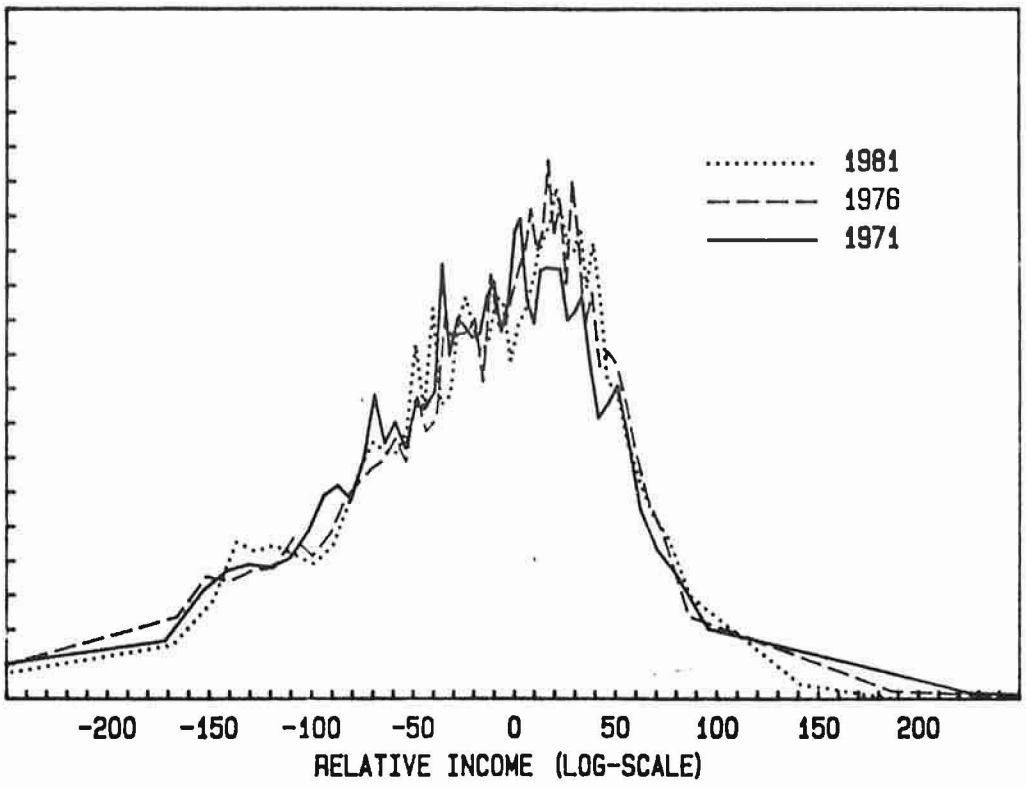


Figure 3.1b Available household income: relative income differences.

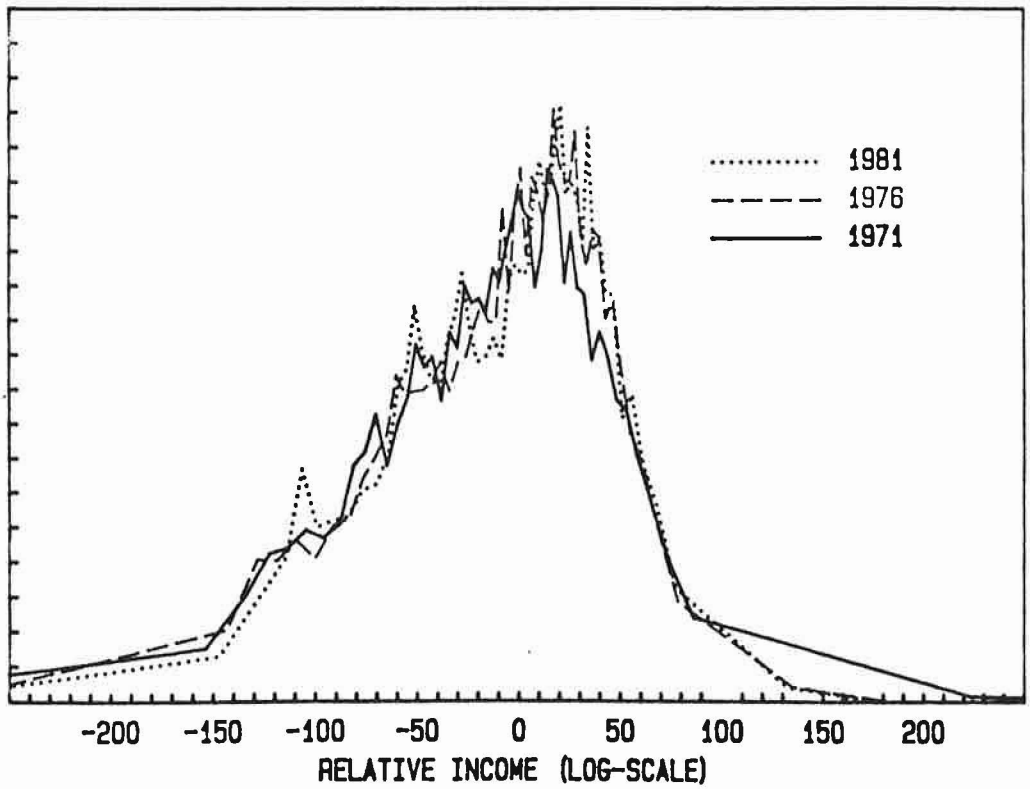


Figure 3.2a Pre-tax household income: Lorenz curves.

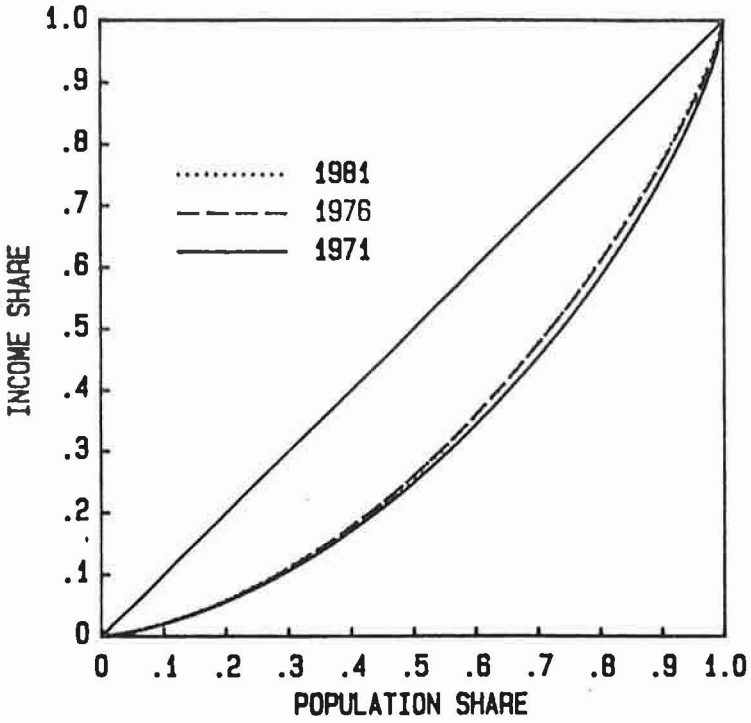
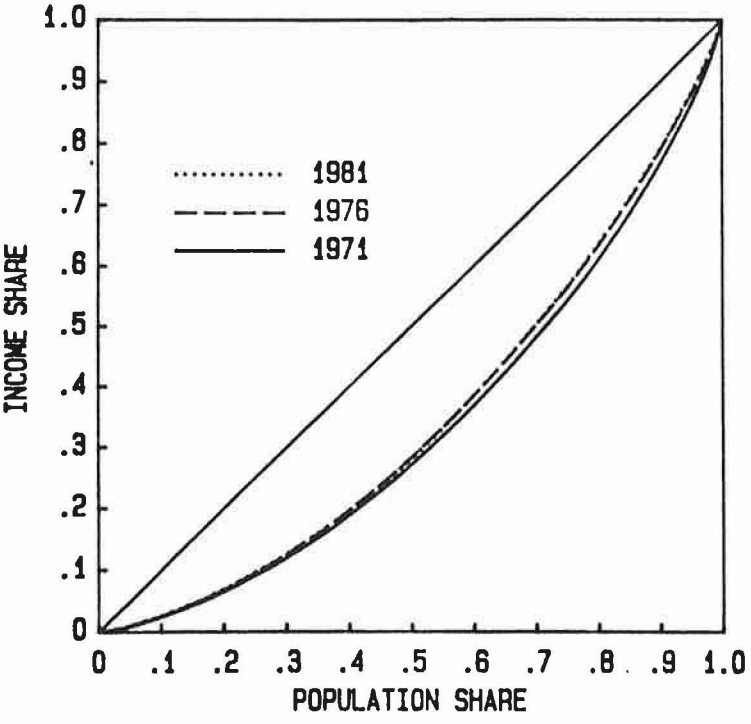


Figure 3.2b Available household income: Lorenz curve.



lesser inequality between household incomes. Again, no similar pattern can be found between 1976 and 1981. In fact, the curves for 1976 and 1981 practically coincide in the Lorenz diagrams.

Similar conclusions regarding the trend of income inequality in the household population can be drawn by comparing Gini coefficients: The Gini coefficient of pre-tax income drops from its 1971 value of .361 to .337 in 1976, a value still holding in 1981, whereas the coefficient of available income develops according to the pattern .326, .302 and .301 over the three years.

Figure 3.3 outlines the distribution of available income in a diagram similar to Pen's parade (with the relative income, originally suggested by Pen [1971], on the vertical axis substituted by its log-transformation).

Figure 3.3 in here

When the population 'marches by', starting from the household with the lowest relative income and ending with the highest, we observe that the 1971 parade falls short of both the 1976 and the 1981 parade up to a point when slightly more than 80 per cent of the population have marched by. After this point the 1971 parade exceeds the 1976/81 parades. Obviously, this is again an indication of the reduced inequality between 1971 and the two later years.

As have been seen above, the traditional graphical methods

Figure 3.3 Available household income: Pen's parades.

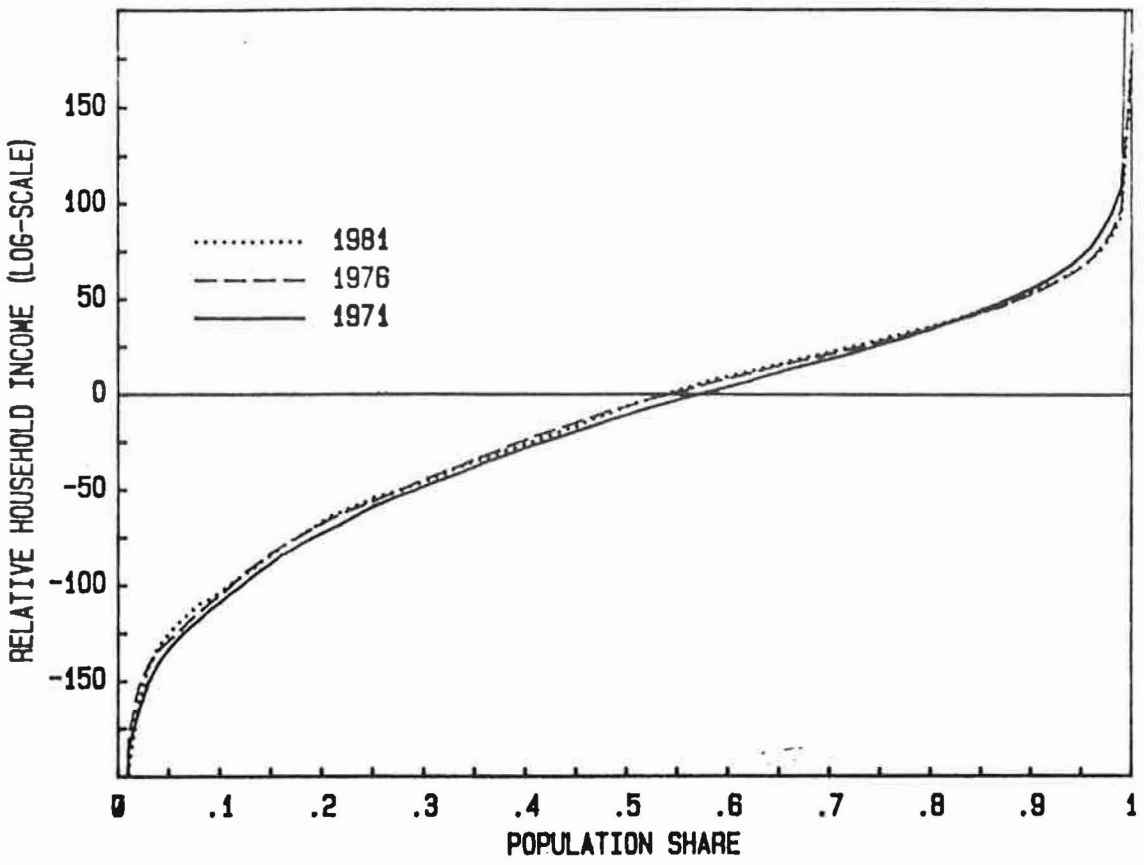
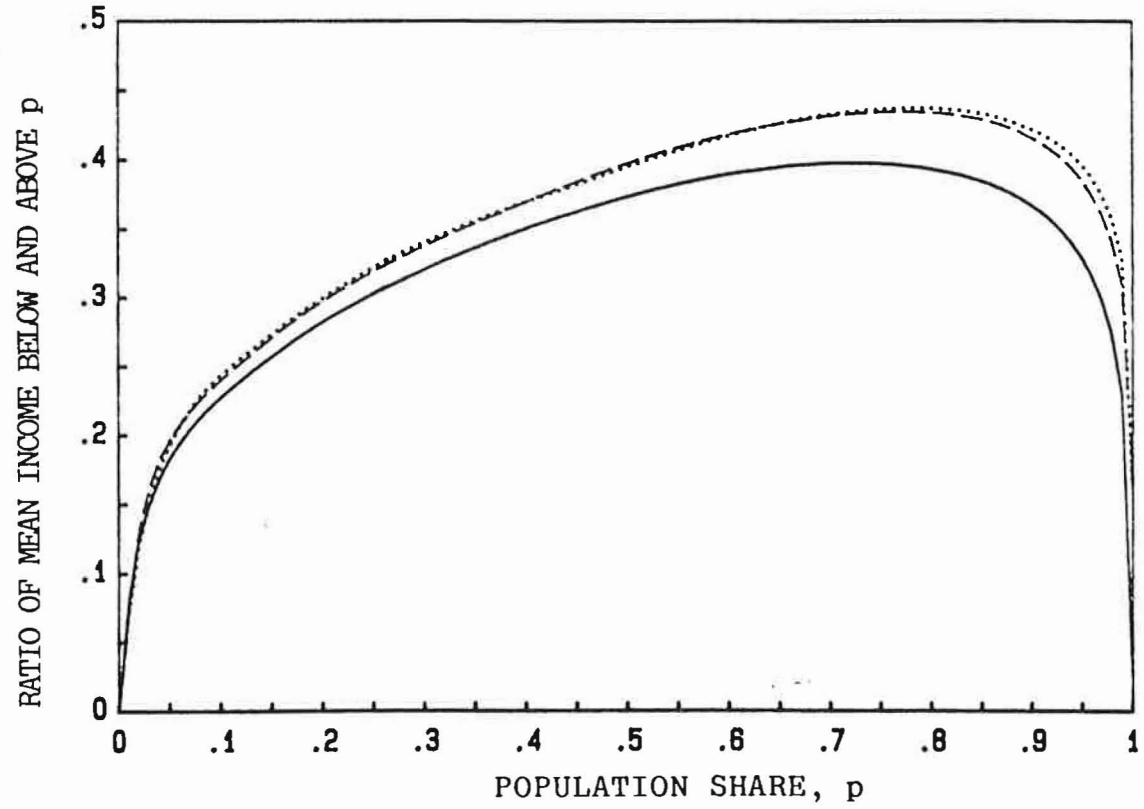


Figure 3.4 Available household income: Aaberge's inequality curves.



for displaying income distribution data do not succeed in separating between the 1976 and 1981 distributions: It is hard to get a summary interpretation of the differences between 1976 and 1981 from the frequency distributions of Figure 3.1b; the Lorenz diagram 3.2b is so 'smooth' that the difference between the 1976 and the 1981 curves disappears visually (this would still be true even if we made the diagram several times larger); and although there is a slight difference between the 1976 and 1981 parades in Figure 3.3, it may be hard to interpret.

To get a visual interpretation of the difference between the 1976 and the 1981 distributions, a graphical device suggested by Aaberge [1982] is presented in Figure 3.4.

Figure 3.4 in here

Similar to the Lorenz diagram and Pen's parade, the horizontal axis denotes population shares, p, ranked according to available income. The vertical axis, in turn, represents the ratio between the mean incomes among households falling to the left and to the right, respectively, of a specific p-value. As can be seen from the Figure, the 1976 and 1981 'inequality' curves are situated above the 1971 curve, again implying a reduction in income inequality. Moreover, the 1981 curve is slightly above the 1976 curve for population shares exeeding .80. Hence, the mean income among e.g. the 80 per cent poorest households as compared to the mean income among the 20 per cent richest households has increased from 1976 to 1981 and in this sense inequality has slightly been reduced.

4. THE EFFECT OF TAXATION

As was seen in above, the distribution of household income tended towards higher equality between 1971 and 1976: Measured by the Gini coefficient the inequality of available income was reduced by 7.4 per cent (from .326 to .302), and the inequality of pre-tax income by 6.6 per cent (from .361 to .337). Since the available income of a household is obtained from its pre-tax income by subtracting taxes (and other paid transfers), the drop in the inequality of available income depends on changes in the distribution of pre-tax income and on shifts in effective taxation schemes. Now, the slightly smaller drop in pre-tax income inequality, as compared to available income inequality, suggests that the effective taxation was more equality promoting in 1976 than in 1971.

This tentative conclusion is supported by a decomposition of the Gini coefficient of the available income according to pre-tax income and paid transfers. The decomposition result is presented in Table 4.1 (decomposition rules are discussed in Nygard and Sandstrom [1981], and Lerman and Yitzhaki [1985]).

Table 4.1 in here

As can be seen from the Table, the relative inequality reducing effect of paid transfers has increased from 37.8 per cent in 1971 to 45.9 per cent in 1976, followed by a

Table 4.1 Decomposition of the Gini coefficient of available income (relative contributions within brackets).

	Gini coefficient	Contribution from pre-tax income	paid transfers
1971	.326 (100.0)	.449 (137.8)	-.123 (-37.8)
1976	.302 (100.0)	.440 (145.9)	-.138 (-45.9)
1981	.301 (100.0)	.431 (143.4)	-.130 (-43.4)

Figure 4.1 Tax rates at different levels of real pre-tax income (1984 standards).

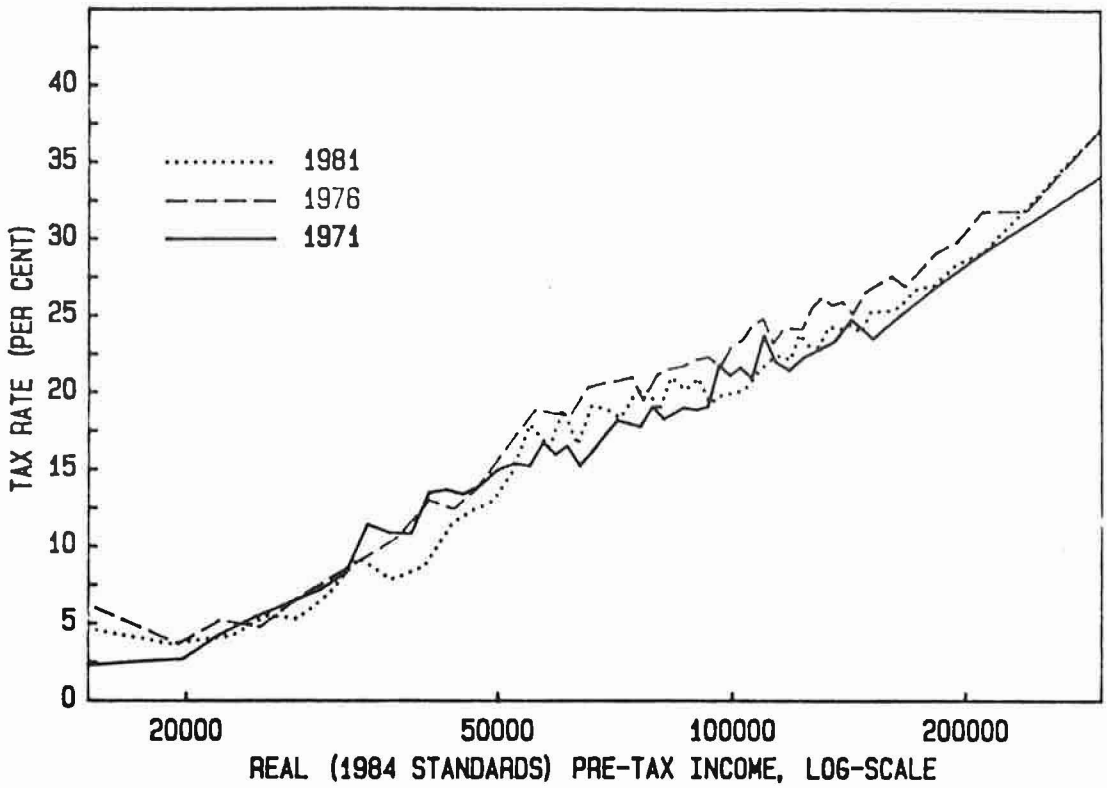
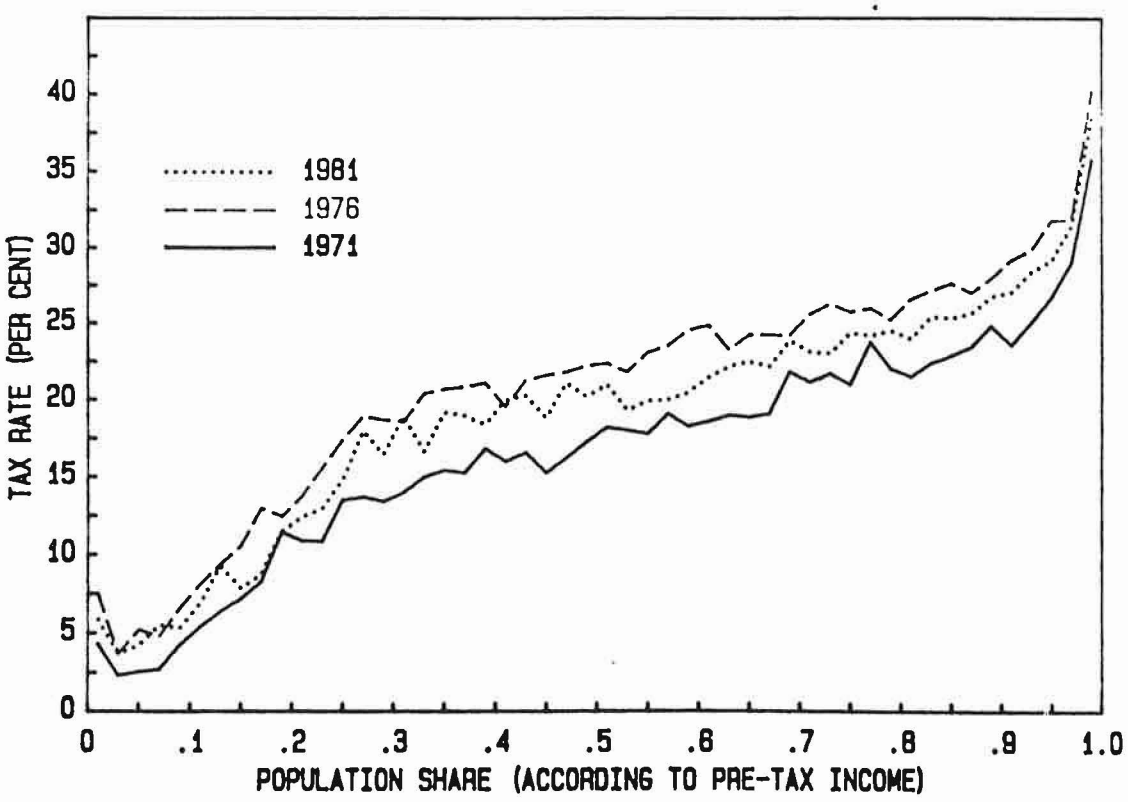


Figure 4.2 Tax rates at different population quantiles.



slight drop to 43.4 per cent in 1981.
Figure 4.1 outlines the actual tax rates 1971, 1976, and 1981 as a function of real (1984 standards) pre-tax income.

Figure 4.1 in here

Even if the tax rate profiles appear somewhat 'erratic', partly to be explained by the influence of sampling errors and partly by the fact that the pre-tax incomes on the horizontal axis refer to different households, the tax progressivity has clearly increased from 1971 to 1976: For real pre-tax incomes above 50000 FIM the 1976 tax rate exceeds the 1971 by some 3 per cent on the average. On the other hand, from 1976 to 1981 the tax rates show a decreasing tendency, the 1981 situation being rather close to the 1971. Figure 4.2 gives another visual interpretation of the tax rate changes from 1971 to 1981. In this Figure the horizontal axis represents population shares, the households being ranked according to pre-tax income.

Figure 4.2 in here

According to this representation, the difference between the 1971 and the 1976 situation becomes extremely clear: Due to the increasing real pre-tax income a growing part of the household population is subject to relatively high tax rates, the average tax rate having increased by some 5 per cent

between 1971 and 1976. From 1976 to 1981 taxation has again been alleviated, the 1981 tax profile lying roughly half-way between the 1971 and 1976 profiles.

5. POOLING THE DATA

5.1 CHANGES IN THE HOUSEHOLD POPULATION 1971-81

The data base includes, in addition to income amounts, information about some 'background' characteristics of each household. On one hand, there are variables referring to the household as a whole: Number of persons belonging to the household, number of children, and number of earners working at least half-time. On the other hand, there is a set of variables characterizing the head of the household: the age, socio-economic group, sex, and education.
Having this information for the years 1971, 1976, and 1981, we may construct household 'cohorts' by a disaggregation of the population according to age of the household head and pooling the data.

Table 5.1 presents a first result of the disaggregation. In the Table the average number of persons, half-time earners, and children are given for households belonging to different age groups. Relying on the 'representativity' of the CSO-samples (this is indeed a heavy assumption, the average sample size within a age group being some 600-700 households for each year) we may derive household 'cohort' profiles over the years.

Table 5.1 in here

To illustrate, the households belonging to the age group 20-24

Table 5.1 Some household characteristics 1971, 1976 and 1981 according to age (head of the household).

Age group	Average number of (A) Persons			(B) Half-time earners			(C) Children		
	1971	1976	1981	1971	1976	1981	1971	1976	1981
20-24	2.28	2.08	1.75	1.39	1.20	.96	.52	.40	.22
25-29	3.02	2.72	2.44	1.52	1.45	1.46	1.10	.83	.63
30-34	3.67	3.24	3.03	1.60	1.55	1.53	1.68	1.32	1.18
35-39	4.00	3.69	3.38	1.63	1.64	1.64	1.99	1.68	1.46
40-44	4.04	3.77	3.40	1.80	1.71	1.76	1.83	1.51	1.24
45-49	3.68	3.53	3.10	1.86	1.82	1.77	1.27	1.00	.78
50-54	3.07	2.98	2.75	1.70	1.60	1.66	.77	.56	.39
55-59	2.69	2.54	2.31	1.50	1.38	1.32	.49	.31	.20
60-64	2.21	2.09	1.95	1.01	.88	.83	.19	.14	.07
65-69	1.89	1.72	1.66	.46	.26	.22	.08	.04	.03
70-74	1.79	1.57	1.57	.31	.15	.08	.11	.04	.02
75-79	1.73	1.44	1.52	.30	.09	.06	.05	.01	.02
20-79	3.00	2.80	2.58	1.38	1.29	1.27	.96	.77	.64

in 1971, will in 1976 be represented by the age group 25-29, and in 1981 by the group 30-34, and looking at the number of half-time earners we for this 'cohort' we find a profile of 1.39, 1.45, and 1.53 half-time earners on the average. In Tables 5.2-5.4 corresponding information about socio-economic groups, sex, and education is given.

Table 5.2 in here

Table 5.3 in here

Table 5.4 in here

Some trends are readily observable from the Tables:

(a) The household size, Table 3.1, has grown smaller during the decade 1971-81, largely due to a smaller number of children.

(b) Even if earnings activity, Table 3.1, on the average seems to have declined, from 1.38 to 1.27 half-time earners per household, the reverse is true for the younger 'cohorts'.

(c) Regarding the distribution over socio-economic groups, Table 3.2, the heavy reduction in the share of farmer households, and the corresponding increase in the white collar share, should be noted. Another significant feature is the growing part of inactive households among the youngest (below 34 years) and oldest (above 65 years) age groups, also reflected in the number of half-time earners within these

Table 5.2 The households 1971, 1976 and 1981 according to age and socio-economic group (head of the household).

Percentage

Age group	(A) Farmers			(B) Own-account			(C) White collar		
	1971	1976	1981	1971	1976	1981	1971	1976	1981
20-24	2.0	1.4	1.4	.8	.6	1.1	37.9	34.2	24.7
25-29	4.7	2.7	3.7	2.4	2.9	1.1	39.7	43.2	41.9
30-34	8.7	5.2	5.1	4.8	4.2	4.6	38.6	46.0	46.4
35-39	10.9	7.4	6.5	8.5	4.3	4.6	32.7	44.8	44.2
40-44	16.8	9.5	8.3	9.2	5.2	7.1	25.1	35.8	42.7
45-49	19.2	13.2	10.4	6.2	6.1	6.2	23.7	30.3	38.2
50-54	19.1	13.8	12.6	6.0	6.8	5.6	21.2	25.2	34.5
55-59	20.0	16.0	12.1	5.9	5.3	4.2	18.2	20.6	23.5
60-64	15.8	12.2	8.9	4.0	3.1	5.0	7.7	13.2	14.5
65-69	8.8	5.4	5.0	1.7	1.6	1.2	1.1	1.6	2.2
70-74	6.1	3.8	2.0	1.5	.8	1.0	.4	1.4	.3
75-79	3.0	2.0	.9	.0	.0	.4	.9	.0	.0
20-79	12.3	8.2	6.7	4.8	3.9	3.9	23.0	28.5	30.9

Percentage

Age group	(D) Blue collar			(E) Inactive		
	1971	1976	1981	1971	1976	1981
20-24	54.2	52.5	44.2	5.1	11.2	28.6
25-29	50.1	47.7	46.8	3.1	3.5	6.5
30-34	46.1	42.8	39.5	1.8	1.8	4.4
35-39	46.3	41.9	42.5	1.6	1.6	2.3
40-44	44.3	44.1	38.2	4.6	5.3	3.8
45-49	43.1	43.7	39.3	7.8	6.8	5.8
50-54	39.7	40.2	35.2	13.9	14.1	12.0
55-59	33.6	33.1	34.7	22.3	24.9	25.4
60-64	18.6	18.9	15.6	53.8	52.6	55.8
65-69	2.8	3.0	.6	85.5	88.4	91.0
70-74	1.5	.0	.0	90.5	94.0	96.7
75-79	2.2	.0	.0	93.9	98.0	98.7
20-79	35.7	34.4	32.4	24.2	25.0	25.9

Table 5.3 The households 1971, 1976 and 1981 according to age and sex (head of the household).

Age group	Percentage with (A) Male head of household			(B) Female head of household		
	1971	1976	1981	1971	1976	1981
20-24	70.7	59.2	54.4	29.3	40.8	45.6
25-29	82.9	79.2	76.5	17.1	20.8	23.5
30-34	84.7	81.7	76.3	15.3	18.3	23.7
35-39	85.6	84.0	78.8	14.4	16.0	21.2
40-44	82.0	77.8	77.3	18.0	22.2	22.7
45-49	81.5	77.9	72.2	18.5	22.1	27.8
50-54	69.4	68.3	68.2	30.6	31.7	31.8
55-59	68.3	64.8	64.2	31.7	35.2	35.8
60-64	65.5	57.2	57.7	34.5	42.8	42.3
65-69	61.4	53.1	51.9	38.6	46.9	48.1
70-74	54.7	45.3	54.3	45.3	54.7	45.7
75-79	46.5	35.5	44.4	53.5	64.5	55.6
20-79	73.9	69.4	67.9	26.1	30.6	32.1

Table 5.4 The households 1971, 1976 and 1981 according to age and education (head of the household).

Age group	Percentage with (A) Primary education			(B) Secondary education			(C) Higher education		
	1971	1976	1981	1971	1976	1981	1971	1976	1981
20-24	48.2	37.3	15.9	46.7	56.9	79.8	5.1	5.8	4.3
25-29	48.3	43.1	18.6	39.3	42.3	63.9	12.5	14.6	17.4
30-34	58.3	45.1	28.9	29.8	39.0	55.3	11.8	15.9	15.8
35-39	64.2	52.2	36.8	25.7	33.5	47.9	10.0	14.3	15.3
40-44	74.2	63.8	46.3	17.1	24.8	37.9	8.7	11.4	15.8
45-49	73.0	72.0	60.3	19.5	20.2	29.9	7.5	7.9	9.8
50-54	78.5	79.3	64.7	15.6	15.8	25.7	5.9	4.8	9.5
55-59	80.4	77.2	74.8	14.9	16.2	19.1	4.7	6.6	6.4
60-64	81.7	82.6	76.0	11.7	13.3	18.3	6.5	4.1	5.6
65-69	82.6	84.4	78.9	11.6	11.2	16.8	5.8	4.3	4.4
70-74	85.1	82.4	78.3	8.7	11.5	17.2	6.2	6.2	4.4
75-79	82.9	89.2	79.4	9.6	7.9	14.1	7.5	2.9	6.5
20-79	70.1	64.8	50.3	22.0	26.0	38.8	7.9	9.3	10.8

groups, Table 3.1.

(d) The households headed by females, Table 3.3, has increased.

(e) The population shares with secondary or higher education, Table 3.4, has grown rapidly.

Bearing these significant structural changes of the household characteristics in mind, the stability of the relative income differences (cf. Figures 3.1a and 3.1b) turns out as quite startling.

5.2 HOUSEHOLD INCOME PROFILES

Table 5.5 presents the pre-tax income, the available income, and the Gini coefficient of available income 1971, 1976, and 1981 according to age group.

Table 5.5 in here

The incomes in the Table are given in nominal amounts, to be multiplied with the factors given in Section 2 when transformed to real (1984 standards) income amounts. The resulting real available household income profiles from the three cross-sections are outlined in Figure 5.1.

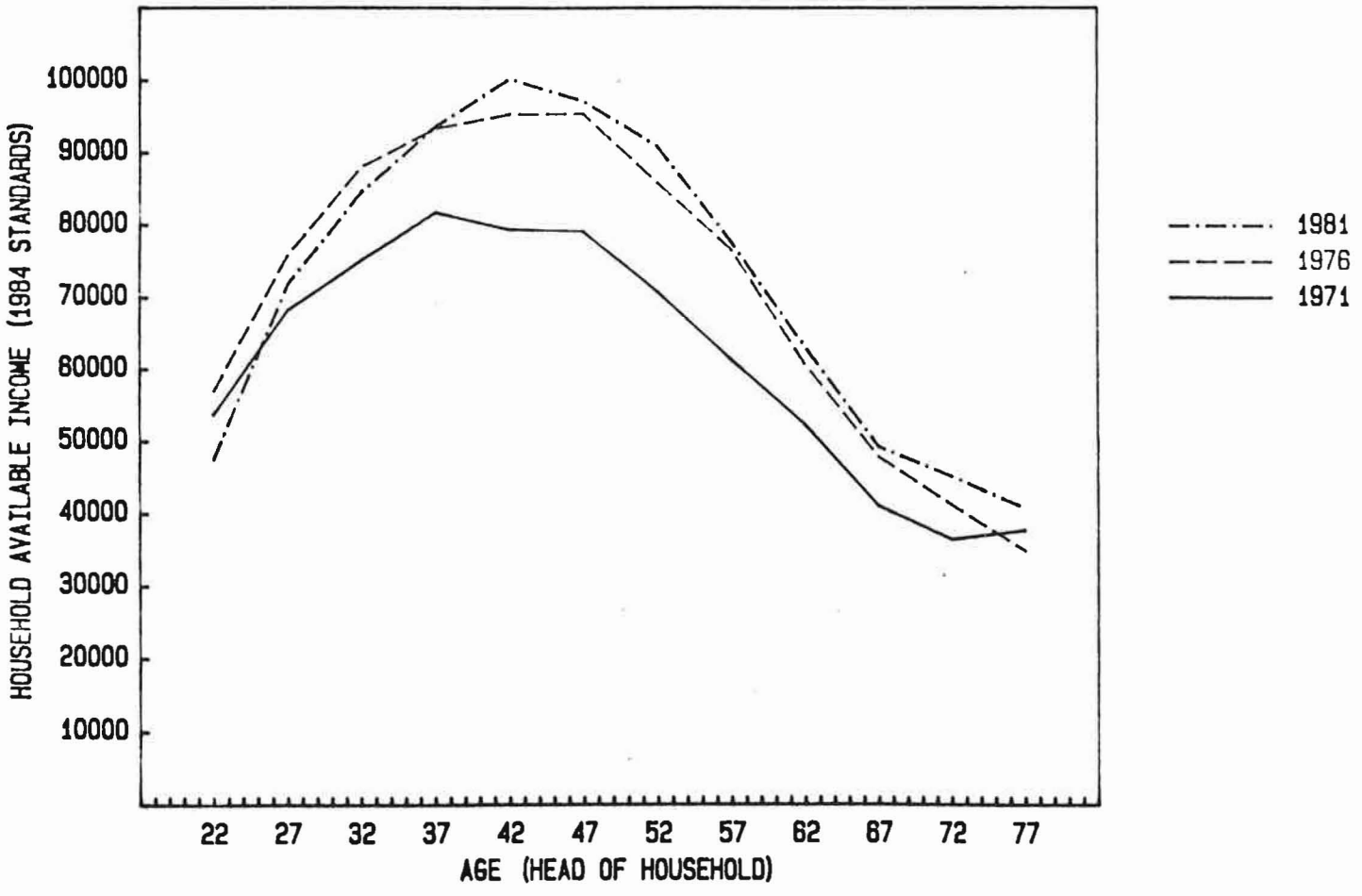
Figure 5.1 in here

The cross-section profiles are quite similar in shape, the major difference being a upward drift when passing from 1971

Table 5.5 Pre-tax household income, available household income, and the Gini coefficient 1971, 1976, and 1981 according to age (head of the household).

Age group	(A) Average pre-tax income (1000 FIM)			(B) Average available income (1000 FIM)			(C) Gini coefficient, available income		
	1971	1976	1981	1971	1976	1981	1971	1976	1981
20-24	17.1	35.5	46.4	13.7	27.6	37.5	.277	.265	.296
25-29	22.7	49.4	72.8	17.5	36.8	56.6	.248	.219	.225
30-34	25.2	59.1	87.4	19.2	42.6	66.6	.244	.215	.211
35-39	27.4	62.1	98.5	20.9	45.2	73.7	.273	.206	.210
40-44	26.4	62.8	106.1	20.3	46.2	78.9	.256	.228	.222
45-49	26.1	62.7	103.5	20.2	46.2	76.5	.287	.252	.255
50-54	23.3	55.8	96.1	18.1	41.6	71.6	.321	.284	.277
55-59	19.5	50.7	80.7	15.6	37.1	61.2	.325	.317	.304
60-64	16.2	37.7	63.9	13.3	29.3	49.6	.370	.333	.323
65-69	12.0	28.2	45.0	10.5	23.7	41.1	.354	.329	.289
70-74	10.3	23.7	41.1	9.3	19.9	35.4	.348	.313	.284
75-79	11.1	19.2	36.5	9.6	16.9	32.0	.376	.292	.291
20-79	21.0	49.4	79.0	16.6	36.9	60.5	.323	.298	.296

Figure 5.1 Cross-section real available household income profiles.



over 1976 to 1981. The similarity in shape is, once again, accentuated by transforming the income amounts to relative incomes in log-scale as in Figure 5.2 below.

Figure 5.2 in here

According to this representation the profiles for the years after 1971 start out from a lower relative income in the age group 20-24, this being compensated by a shift of the peak towards higher ages.

Figure 5.3 outlines the average income profiles for eleven household 'cohorts' over the three years, with the profiles for the youngest 'cohorts' leftmost in the diagram.

Figure 5.3 in here

The four youngest 'cohorts', where the head is born between 1932 and 1951 (the age groups 20-24, 25-29, 30-34, and 35-39 in 1971), have all increasing available income profiles, whereas the fifth 'cohort' (the age group 40-44 in 1971) is the youngest showing a reduction in available household income. The drop in available income between 1976 and 1981 for this 'cohort' may be contrasted to the information in Table 5.1, revealing that the reduced income is associated with a lower earning activity: In the 'cohort' the average number of half-time earners falls from 1.82 in 1976 to 1.66 in 1981.

Figure 5.2 Cross-section relative income differences according to age group.

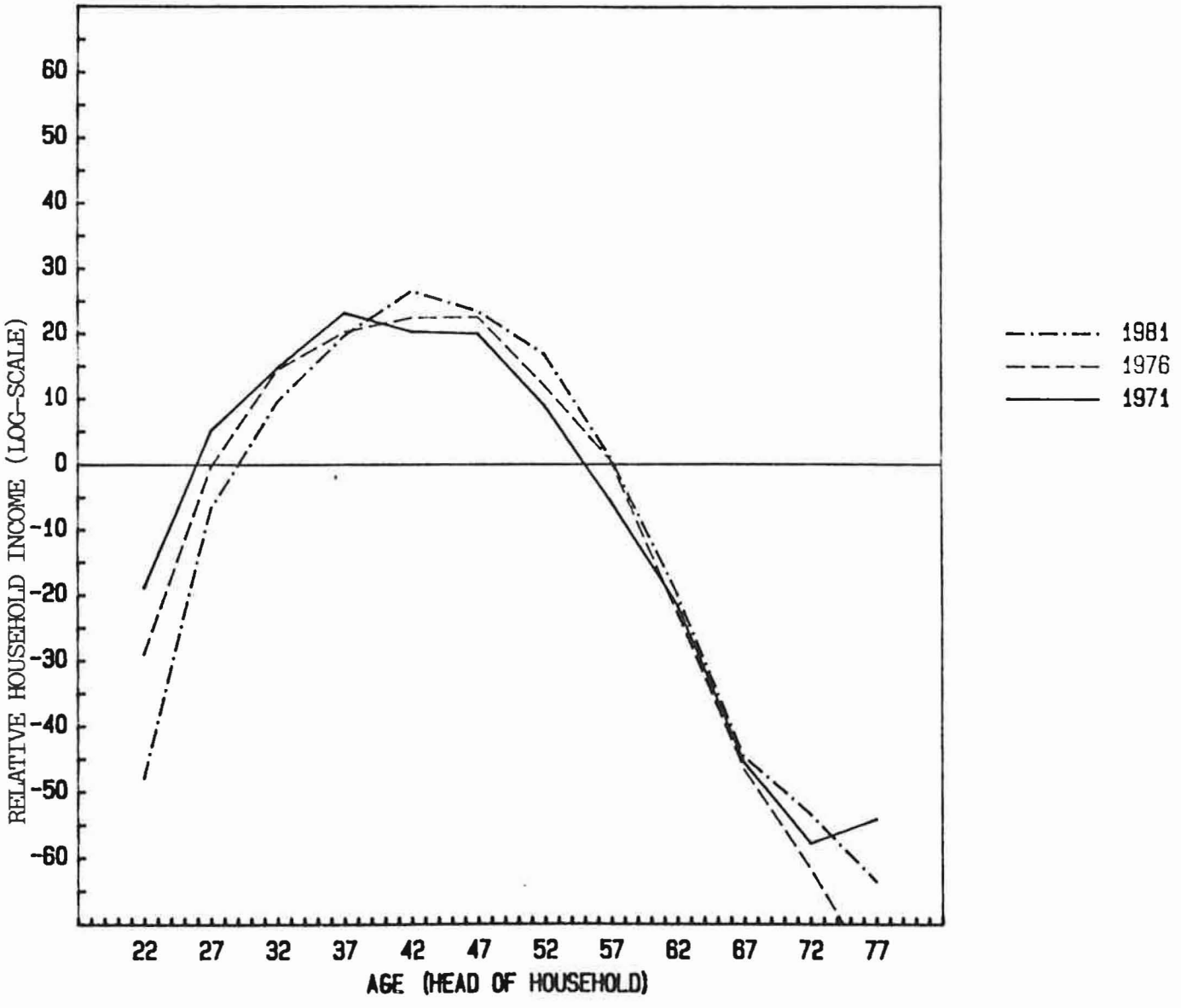


Figure 5.3 Available household income (1984 standards): 'cohort' profiles.

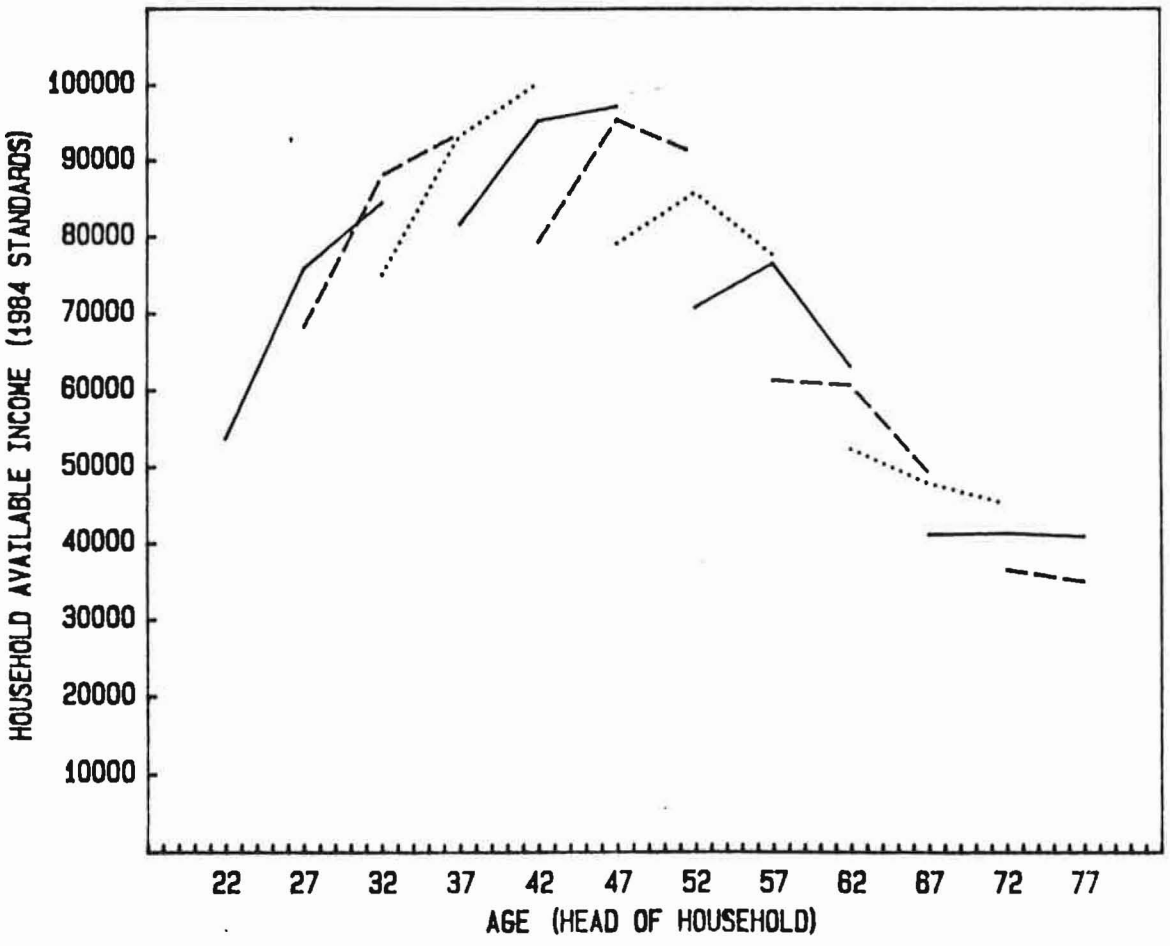


Figure 5.4, finally, gives the eleven 'cohort' profiles transformed to relative incomes in log-scale. The profiles fall remarkably close to one another, with relative incomes appearing to be reasonable approximated by a quadratic function of age.

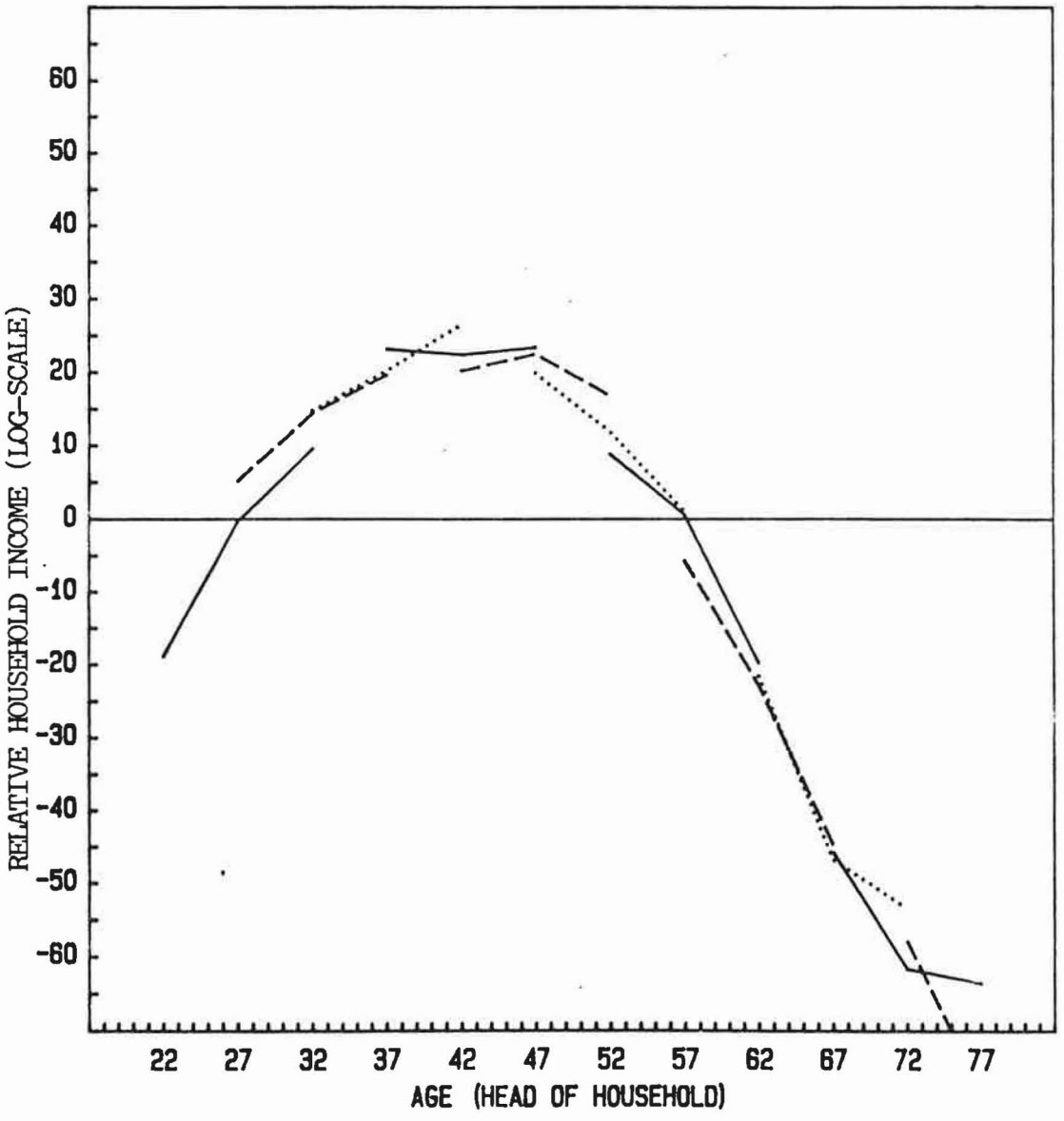
Figure 5.4 in here

5.3 A POOLED CROSS-SECTION REGRESSION

The discussion above indicates that there is a relationship between the 'age' of the household and its income. Similarly, the household income may be related to other 'background' characteristics of the household. Obviously, a descriptive study of these relationships could be based on calculations of average incomes conditional on the household characteristics to obtain a set of crosstabulations. However, since a detailed analysis requires a large number of crosstabulations the results will tend to be 'messy' and awkward to interpret. As a consequence, we will drop detailed crosstabulations in favour for an analysis of regession-type.

To spell out the adopted regression model in somewhat greater detail, we start with the case of data from one year only. The dependent variable in the analysis will be the relative incomes (in log-scale), and to this we relate the set of 'background' characteristics as independent variables. The model will be applied to economically active households only,

Figure 5.4 Relative income differences (log-scale): 'cohort' profiles.



as the socio-economic group 'inactive' may be anticipiated to form a case of its own. From the data base we form five independent variables on ratio scale:

* Number of half-time earners as a proxy for earnings activity,
* Number of children,
* Number of non-earning adults, defined as the household size minus the number of half-time earners and children,
* Age, and age squared, as a proxy for seniority.

Moreover we introduce three sets of dummy variables to take account of

* Sex (two dummies),
* Socio-economic group (four dummies), and
* Education (three dummies).

In this way we obtain fourteen independent variables, and fifteen after introducing an intercept. However, a direct inclusion of the dummies into the regression will bring about singularity. Instead of using the mainstream method to ensure non-singularity, i.e. by forcing one regression coefficient within each dummy set to zero (being the same thing as excluding one dummy from each set), we for ease of interpretation rely on the method suggested by Klevmarken [1972] and restrict the problem by requiring the sum of regression coefficient within each dummy set to equal zero. For instance, we will include both a male and a female dummy under the restriction that the corresponding regression coefficients add to zero.

Even if there are some indications of interactions between

the independent variables, no interaction terms will be included in this analysis.

The effect of pooling the data over the three years could principally be handled by defining a fourth set of dummy variables referring to the years. But, again for reasons of ease of interpretation, we prefer to introduce the effect of the three years in a way similar to the treatment of interventions in time series analysis. To be more specific, we define three 'time' variates, z_1, z_2, and z_3, by

$z_1 = 0$ if year < 1971,
$\quad\quad 1$ if year ≥ 1971,
$z_2 = 0$ if year < 1976,
$\quad\quad 1$ if year ≥ 1976,
$z_3 = 0$ if year < 1981,
$\quad\quad 1$ if year ≥ 1981,

and replace each of fifteen independent variables in the one-year model above with three variables, obtained by multiplying the 'original' independent variable with z_1, z_2, and z_3, respectively. In this way we get a final regression model with 45 independent variables, in which each 'original' variable is associated with three regression coefficients. The first of these coefficients may be interpreted as a 'base' coeffient referring to year 1971, the second represents an additional component which added to the 1971 base gives the 1976 regression coefficient, and the third a second additional component which added to the 1971 and 1976 terms gives the 1981 regression coefficient. Or to put it in another way, the second of the three coefficients represents the change from

1971 to 1976, and the third the change from 1976 to 1981.

Table 5.6 gives the result of the 'descriptive' regression analysis, being based on a total of 18932 households, with the relative pre-tax income (log-scale) as the dependent variable and traditional OLS-estimation of parameters (using IMSL library subroutines).

Table 5.6 in here

Bearing the rather crude quality of the independent variables in mind, the model succeeds suprisingly well, as measured by R^2, in explaining relative income differences. Regarding the interpretation of the results, it should be noted that the regression coefficients may roughly be tought of as the percentage contribution of each variable to the relative household income. A closer inspection of the results give rise to the following tentative conclusions:

Half-time earners The influence of earnings activity on relative household income seems to have increased over time, the regression coefficient for the number of half-time earners growing from 36.07 in 1971 to 41.14 in 1981.

Inactive adults No significant trend over time.

Children The number of children seems slightly to have increased its effect on relative household income.

Seniority The changes in the regression coefficients for age and age squared are all significant, implying a shift of the relative income peak towards higher ages. The estimated

Table 5.6 Coefficient estimates of the pooled cross-section regression (t-values within brackets).

N = 18932 households

R^2 = .539

		1971 Base	1976 First addition	1981 Second addition
Intercept		-178.42 (- 35.95)	5.17 (.70)	-46.59 (- 5.20)
Household characteristics				
Number of half-time earners		36.07 (54.74)	2.42 (2.41)	2.65 (2.40)
Number of other adults		17.54 (25.10)	.92 (.91)	- .51 (- .48)
Number of children		3.10 (8.23)	1.13 (1.87)	1.45 (2.05)
Head of household characteristics				
Seniority:	Age	4.51 (17.80)	- .78 (-1.93)	1.72 (3.74)
	Age^2	- .051 (-16.88)	.011 (2.28)	- .016 (-2.95)
Sex :	Male	14.75 (22.78)	-3.02 (-3.30)	-1.66 (-1.78)
	Female	-14.75 (-22.78)	3.02 (3.30)	1.66 (1.78)
Education:	Primary	-24.71 (-30.50)	8.21 (6.94)	2.00 (1.60)
	Secondary	-10.59 (-12.97)	4.04 (3.50)	2.84 (2.53)
	Higher	35.50 (30.34)	-12.25 (-7.45)	-4.84 (-3.01)
Socio-economic group :	Farmer	-37.40 (-34.91)	3.16 (1.85)	7.76 (3.97)
	Own-account	5.28 (3.80)	1.97 (.91)	-3.28 (-1.38)
	White collar	26.37 (27.56)	-4.91 (-3.46)	-3.83 (-2.55)
	Blue collar	5.76 (7.37)	- .21 (- .18)	- .65 (- .50)

income peak in 1971 corresponds to an age of 44.2 years, in 1976 to 46.6 years, and in 1981 to 48.7 years (cf. Figure 5.4).

Socio-economic group In the base situation 1971 the farmer households have the lowest relative income, and the white collar household the highest. Towards 1981 the farmers improve their relative income position, at the 'expense' of white collar households.

Sex The relative income differences between households headed by females and males have significantly decreased over time.

Education Both the primary and the secondary education group have improved their relative income position at the 'expense' of higher education.

REFERENCES:

Aaberge, R. [1982]. Om malning av ulikskap. Report 82/9, Central Bureau of Statistics of Norway, Oslo.

Klevmarken, A. [1972]. Statistical Methods for the Analysis of Earnings Data. Almqvist & Wiksell, Uppsala, 1982.

Lerman, R.I. and Yitzhaki, S. [1985]. Income Inequality Effects by Income Source: A New Approach and Applications to the United States. Rev. Econ. Statist., 67, 151-156.

Nygard, F. and Sandstrom, A. [1981]. Measuring Income Inequality. Almqvist & Wiksell International, Stockholm, 1981.

Parkkinen, P. [1985]. Suomalaisten kotitalouksien tuloerot ja toimeentuloerot. Työväen taloudellinen tutkimuslaitos, Katsaus, 13, 1, 25-36.

Pen, J. [1971]. Income Distribution. Facts, Theories, Policies. Praeger Publ., New York, 1971

ELINKEINOELÄMÄN TUTKIMUSLAITOS (ETLA)
The Research Institute of the Finnish Economy
Lönnrotinkatu 4 B, SF-00120 HELSINKI 12 Puh./Tel. (90) 601 322

KESKUSTELUAIHEITA - DISCUSSION PAPERS ISSN 0781-6847

No 173 EERO PYLKKÄNEN - PENTTI VARTIA, Some Comments on Fine-Tuning Macro-Model Forecasts. 29.05.1985. 22 p.

No 174 CHRISTIAN EDGREN, Automatic versus Discretionary Tax-Policy Effects Revisited. 05.06.1985. 44 p.

No 175 TAUNO KALLINEN, Työn tuottavuuden taso ja ansiotaso Suomen, Ruotsin, Norjan ja Tanskan tehdasteollisuudessa vuosina 1973 ja 1982. 11.06.1985. 44 s.

No 176 J. LESKELÄ - J. MOILANEN, A Comparison of Recent Medium-Term Forecasts. 17.06.1985. 34 p.

No 177 PAAVO SUNI, Efektiivisen nimellisen valuuttakurssin mittaaminen ja mittarien tulkinta. 20.06.1985. 42 s.

No 178 MARKKU LAMMI, Alkoholi- ja polttoaineverojen kertymästä. 27.09.1985. 49 s.

No 179 YRJÖ O. VARTIA, Tunnetuissa pisteissä epäsäännöllisen funktion numeerinen integrointi. 21.10.1985. 8 s.

No 180 ERKKI KOSKELA, On Monopolistic Tax Evasion under Direct and Indirect Taxation. 23.10.1985. 17 p.

No 181 PEKKA ILMAKUNNAS, Stochastic Constraints on Cost Function Parameters: Mixed and Hierarchical Approaches. 31.10.1985. 26 p.

No 182 HIROKI TSURUMI - HAJIME WAGO - PEKKA ILMAKUNNAS, Gradual Switching Multivariate Regression Models with Stochastic Cross-Equational Constraints and an Application to the Klem Translog Production Model. 01.11.1985. 28 p.

No 183 HANNU TÖRMÄ, Panossubstituution rakennnemuutos Suomen teollisuudessa. 06.11.1985. 23 s.

No 184 FREDRIK NYGÅRD, Lifetime Incomes in Finland - Desk Calculations Based on Civil Servant Salaries 1985. 18.11.1985. 28 p.

No 185 FREDRIK NYGÅRD, Relative Income Differences in Finland 1971-1981. 18.11.1985. 22 p.

Elinkeinoelämän Tutkimuslaitoksen julkaisemat "Keskusteluaiheet" ovat raportteja alustavista tutkimustuloksista ja väliraportteja tekeillä olevista tutkimuksista. Tässä sarjassa julkaistuja monisteita on rajoitetusti saatavissa ETLAn kirjastosta tai ao. tutkijalta.

Papers in this series are reports on preliminary research results and on studies in progress; they can be obtained, on request, by the author's permission.

0033A/18.11.1985