VATT-KESKUSTELUALOITTEITA

VATT-DISCUSSION PAPERS

126

NONPARAMETRIC ESTIMATES OF AGE PROFILES OF HOUSEHOLD INCOME, CONSUMPTION, DIRECT TAXES AND PUBLIC TRANSFERS

Eugen Koev

Valtion taloudellinen tutkimuskeskus

Government Institute for Economic Research

Helsinki 1996

ISBN 951-561-177-6

ISSN 0788-5016

Valtion taloudellinen tutkimuskeskus

Government Institute for Economic Research

Hämeentie 3, 00530 Helsinki, Finland

J-Paino Ky

Helsinki, July 1996

KOEV, EUGEN: NONPARAMETRIC ESTIMATES OF AGE PROFILES OF HOUSEHOLD INCOME, CONSUMPTION, DIRECT TAXES AND PUBLIC TRANSFERS. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, 1996. (C, ISSN, 0788-5016, No 126). ISBN 951-561-177-6.

ABSTRACT: The report studies the distribution of income, expenditure, direct taxes and public transfers with respect to the age of household's head. The data consist of five household budget surveys from the period 1971-1990 collected by Statistics Finland. The age- and cohort profiles of the variables are estimated using a nonparametric technique, relying purely on information from the data. The relevant decision making and consumption unit in the report is chosen to be the household, so no attempt is made to associate the different components of income, consumption and transfers with particular members of the household. The examinations are done both by household and by consumption unit. The number of consumption units in the household is determined as recommended by OECD.

The age and cohort profiles are an interesting way of summarising data information and they can be also used as a first step in certain types of studies concerning welfare. The estimated age profiles of consumption are concave, as expected. Similar is the shape of age profiles of direct taxes and consumption. Disposable income is more evenly distributed than factor income, so direct taxes and income transfers equalise income among households of different age. Pensioner families are the most important recipients of income transfers, while households with children benefit most from public services. The results point out also to the fact, that the monetary value of benefits received by young families from society have in late eighties exceeded the amount of direct taxes paid by those families.

KEY WORDS: age profile; consumption and income distribution; public transfers; nonparametric estimation.

KOEV, EUGEN: NONPARAMETRIC ESTIMATES OF AGE PROFILES OF HOUSEHOLD INCOME, CONSUMPTION, DIRECT TAXES AND PUBLIC TRANSFERS. Helsinki, VATT, Valtion taloudellinen tutkimuskeskus, 1996. (C, ISSN, 0788-5016, No 126). ISBN 951-561-177-6.

TIIVISTELMÄ: Raportissa tutkitaan sekä tulojen, kulutuksen ja välittömien verojen että tulonsiirtojen ja yhteiskunnallisten palvelujen jakaumaa kotitalouden päämiehen iän mukaan. Aineistona on viisi Tilastokeskuksen kotitaloustiedustelua ajanjaksolta 1971-1990. Tarkasteltavien muuttujen ikä- ja kohortiprofiilit estimoidaan ei-parametrisesti, nojautuen yksinomaan aineiston sisältämään informaatioon. Raportissa päätöksenteko- ja havaintoyksiköksi on valittu kotitalous, joten tulo-, kulutus- ja tulonsiirtoeriä ei ole pyritty kohdistamaan erikseen kotitalouden jäsenille. Tarkastelut suoritetaan sekä kotitalous- että kulutusyksikkökohtaisesti. Kotitalouksien kulutusyksiköt määritellään OECD:n suosituksen mukaisesti.

Ikä- ja kohorttiprofiilien estimointi on kiinnostava tapa tiivistää aineiston sisältämää informaatiota. Lisäksi estimoinnit ovat luonnollinen alkuaskel hyvinvointia käsitteleviin jatkotutkimuksiin. Estimoidut tulojen ikäprofiilit ovat ennakko-odotusten mukaisesti konkaaveja. Samanmuotoiset ovat myös välittömien verojen ja kulutuksen profiilit. Käytettävissä olevat tulot ovat jakautuneet tasaisemmin kuin tuotannontekijätulot, joten välittömät verot ja tulonsiirrot tasoittavat tulojen jakaumaa eri-ikäisten kotitalouksien kesken. Eläkeläisperheet ovat suurimmat julkisten tulonsiirtojen saajat, kun taas lapsiperhet hyötyvät eniten yhteiskunnallisista palveluista. Tuloksista käy ilmi, että 1980-luvun loppupuolella nuorten perheiden yhteiskunnalta saamien etuuksien rahallinen arvo on ylittänyt näiden perheiden maksamat välittömät verot.

ASIASANAT: ikäprofiili; tulojen ja kulutuksen jakautuminen; tulonsiirrot; ei-parametrinen estimointi.

CONTENTS

Page

1 INTRODUCTION

This paper studies the age and cohort profiles of household income, consumption, direct taxes and receipts of public transfers with respect to (w. r. t.) the age of the head of household[1]. The data used are extensive household surveys collected by Statistics Finland in 1971, 1976, 1981, 1985 and 1990. The age profiles are estimated for each cross-section. A cross-sectional age profile provides information on how the average level of the respective variable differs between households of different age at a particular point in time. The data used makes it possible to examine changes in the age profiles over a period of 19 years. A cohort is formed by the households of same age in a particular year. Although there is no information about the same households in different years, it is natural to think, that eg. by comparing the survey data on the households aged 20 in 1971 to the households aged 25 in 1976, one gets an idea of the change undergone by the "typical" 20 years old household in 1971 during the period of five years.

Previously cross-sectional age profiles of household disposable income and consumption have been estimated from the same data by Sullström and Riihelä (1993). In this paper we consider also the age profiles of factor income and disposable income plus (public) transfers in kind. Thus one can examine the effects of taxation and public transfers on the distribution of income among households of different age. The age profiles of direct taxes and receipts of public services are separately estimated here.

The derived age and cohort profiles represent an interesting summary device which is used in this paper for descriptive purposes. However, its results can be utilised in more sophisticated statistical analysis to isolate the average effect of age on the variables of interest. They provide the appropriate information if for example welfare inequality among households is to be decomposed into inequality between households of different age and inequality within households of same age. In addition, the estimated profiles offer convenient means of assessing the life-cycle effects and the interactions between consumption and the different sources of income and

---

[1] For convenience the expression "age profile of variable X " is used synonymously with "distribution of the variable X w.r.t. age of household's head". The expression "household age" is used to denote the age of household's head.

can be possibly used as a first step in building semiparametric models where the impact of age on consumption is not restricted by a priori assumptions concerning its shape.

The approach used in the present paper and in Sullström et. al. (1993) can be seen as augmenting and providing a rather different point of view on life-cycle profiles of public transfers than a recent study by Parkkinen et. al. (1996). In the latter receipts of public tranfers and consumption associated with them always accrues to a particular individual in the household. Here the relevant concept is taken to be the household. Therefore in this paper eg. child benefits, which are received by households with children, are considered not as a source of financing children's consumption specifically, but as a component of household's income, which is used to finance household's consumption in general. If one can addopt the view, that the household is a proper decicion making unit allocating goods and services "fairly" among its members the present approach has obvious merits. On the other hand, it has the disadvantage of leaving out those members of society, who live in institutions, eg. a substantial part of the old persons.
In particular the present study does not account for their share of public welfare services.

The paper proceeds as follows. In chapter one the estimation method, nonparametric regression, is discussed. The benefit of using a nonparametric technique is that no restrictive priori assumptions are made for the shape of the age profiles of the investigated variables. This guarantees that the estimation results provide an adequate summary of the information contained in the data. Chapter two is a short description of the data and a summary of the estimation results, which are presented in the Appendix.

2 KERNEL REGRESSION[1]

Kernel regression estimation (regression smoothing) is a method for nonparametric regression fitting. As nonparametric procedures in general, it is most easily applicable in cases where the attention is restricted to the relationship between only two variables, although it can be extended to cover multivariate cases as well. In the exposition we shall consider only the problem

[1] The results presented here are based on Härdle (1990) and Härdle (1992).

of regressing the dependent variable, Y, on a single explanatory variable, X. The regression curve at point x is defined as:

(1) $$g(x) = E(Y|X = x).$$

In classical parametric regression estimation a functional form of g(x) is specified up to a vector of parameters, in other words it is assumed known on the basis of some theory or "experience" and the estimation procedures aim at obtaining estimates of the unknown parameters. The best known example is provided by the standard case in which the regression curve is assumed to be a straight line:

(2) $$g(x) = E(Y|X = x) = a + bx.$$

Here the unknown parameters to be estimated are *a* and *b*. The assumption that the shape of the regression curve is known is too restrictive, especially when little prior knowledge is available. In such cases it seems sensible to let the data determine the shape of the regression curve, at least for the beginning. That is precisely what nonparametric regression procedures do, since they rely only on sample information. Once the shape of the regression curve has been revealed by the nonparametric fit, a parametric specification may be considered and the nonparametric regression may be used to test the performance of different parametrisations. The nonparametric estimate of the regression curve may as well be the final one.

In the following it is assumed that $\{(X_i, Y_i)\}^n_{i=1}$ are independent identically distributed (i.i.d.) variables, whose probability distribution is continuous. The starting point of the kernel regression estimation is rewriting (1) in the form:

(3) $$g(x) = \int \frac{yf(x,y)}{f(x)} dy = \frac{\int yf(x,y)dy}{\int f(x,y)dy},$$

where $f(x,y)$ denotes the density function of the joint distribution of X and Y and $f(x)$ denotes the marginal density function of X. (thus (3) is the definition of the conditional expectation of

Y at X=x). Applying the theory of kernel density estimation, which is not discussed here, separately to the numerator and denumerator of (3) one obtains the so called *Nadaraya-Watson* estimator of (3) at point x (Nadaraya (1964), Watson (1964)):

$$\text{(4)} \qquad \hat{g}(x) = \frac{(nh)^{-1}\sum_{i=1}^{n} K(\frac{x-X_i}{h})Y_i}{(nh)^{-1}\sum_{i=1}^{n} K(\frac{x-X_i}{h})} = (nh)^{-1}\sum_{i=1}^{n} W_{hi}(x)Y_i \ .$$

The last equality is used to emphasise the fact that the kernel estimate of *g(x)* is essentially a weighted average of the response variable, the weights being determined by

$$\text{(5)} \qquad W_{hi}(x) = \frac{K(\frac{x-X_i}{h})}{(nh)^{-1}\sum_{i=1}^{n} K(\frac{x-X_i}{h})} = \frac{K(\frac{x-X_i}{h})}{\hat{f}(x)} \ .$$

In (4) and (5) K(.) is a *kernel*, *h* is the *bandwidth* and $\hat{f}(x)$ is the kernel estimate of the marginal density function of X at x. Kernels are weight functions with the following properties: they are continuous, bounded, symmetric and non-negative and integrate to one. From practical point of view the choice of kernel makes little difference, at least when it is made among several well-known kernels. What is important in kernel estimation is the choice of *h*. The bandwidth determines the "degree of smoothing" of the estimator. It can be easily shown that

$$\text{(6)} \qquad \begin{aligned} &\hat{g}(X_i) \to Y_i, \text{ if } h \to 0, \\ &\hat{g}(X_i) \to n^{-1}\sum_{i=1}^{n} Y_i = \overline{Y}, \text{ if } h \to \infty . \end{aligned}$$

Thus a very narrow bandwidth tends to reproduce the data and a very broad bandwidth tends to fit the average of the response variable. Intuitively it is clear, that the bandwidth should be chosen in such a way, that the systematic features (the shape) of the regression of Y on X are correctly revealed, but there is no unnecessary noise due to randomness in the fit. It should be pointed out, that the results presented in (6) are based on the assumption that there are no ties in the explanatory variable (that is $X_i \neq X_j \ \forall i \neq j$). If the explanatory variable is discrete or if the data is obtained from an experiment, where the values of X have been fixed at several pre-determined levels a lot of ties arise naturally. In this case the first statement in (6) will no longer be true. With bandwidth approaching 0 one will obtain the average value of the responses Y_j at the particular level of the explanatory variable.

It can be shown that the Nadaraya-Watson estimator is a consistent estimator of $g(x)$ if $h \to 0$ and $nh \to \infty$. In finite samples it is biased. The approximate mean squared error at $\hat{g}(x)$, assuming stochastic regressor, is given by (see Härdle, 1992, p 135, table 3.6.2):

$$\text{MSE}(\hat{g}(x)) \approx \frac{(nh)^{-1}\sigma^2(x)}{f(x)} \int K(u)\,du + \tag{7}$$
$$+ \frac{h^4}{4}\left(\int u^2 K(u)\,du\right)^2 \left[g''(x) + 2\frac{g'(x)f'(x)}{f(x)} \right]^2,$$

where $K(.)$ is the kernel and $\sigma^2(x) = Var(Y|X{=}x)$. The first term in (7) is the variance of the estimator. The second one is the squared bias. MSE for the non-stochastic case is slightly different, but the general implications remain the same (compare Härdle, 1990, pp 76-77, tables 3.6.1 and 3.6.2). The variance is a decreasing function of the bandwidth, while the bias is an increasing one. Thus an optimal bandwidth must balance bias against variance. The bias depends on the second derivative of $g(x)$ and is large around narrow "peaks" and "valleys" appearing in the true regression curve. From (7) it can be shown that the optimal bandwidth should decrease at the rate of $n^{-1/5}$. However, (7) obviously cannot be used for choosing the bandwidth, since it depends on several unknown parameters. Moreover, it measures the uncertainty of estimation only at a particular point and does not provide a guidance for the overall performance of the estimator. For this purpose some global measure for the discrepancy between the true and the estimated regression curves should be used. Typically squared distance measures, such as the Averaged Squared Error (ASE), the Integrated Squared Error (ISE), the Conditioned Squared Error (MASE) or the Mean Integrated Squared Error are used. It turns out, that choosing h so as to minimise any of these will asymptotically leads to the same level of smoothing. Thus it suffices to examine the choice of h by minimising say:

(8) $$ASE(h) = n^{-1}\sum_{i=1}^{n}(\hat{g}_h(X_i) - g(X_i))^2 w(X_i).$$

Since (9) depends on the unknown true regression curve, *ASE* has to be estimated. A "naive" estimator is

(9) $$\begin{aligned}&\hat{p}(h) = n^{-1}\sum_{i=1}^{n}[Y_i - \hat{g}(X_i)]^2 w(X_i) = \\ &= ASE(h) + n^{-1}\sum_{i=1}^{n}\varepsilon_i^2 w(X_i) - 2n^{-1}\sum_{i=1}^{n}\varepsilon_i(\hat{g}_h(X_i) - g(X_i))w(X_i), \\ &\text{where } \varepsilon_i = Y_i - g(X_i).\end{aligned}$$

This estimator is even asymptotically biased and minimising (9) with respect to h will result in a tendency to choose too narrow bandwidth. However the asymptotic bias of (9) is due only to the fact that the expectation of the last term is not asymptotically negligible in comparison with the expectation of *ASE*. There are several possible ways to construct an asymptotically unbiased estimator of *ASE*. One is by introducing a correcting term into (9). This leads to the construction of *penalising functions* such as the Akaike's information criterion and (9) adjusted by some penalising function is called a *penalising function selector*. There exist several penalising functions in the literature, all of them with the same asymptotic properties. Another approach is based on the fact that only the diagonal error terms ε_i have non-zero contributions to the expected value of the last term in (9). This observation leads to the idea of *cross-validation*. The cross-validation employs the estimates $\hat{g}_{h,i}(X_i)$ of $g(X_i)$ in which information contained in the i:th observation (Y_i, X_i) is not used. Thus (9) is replaced by :

(10) $$CV(h) = n^{-1}\sum_{i=1}^{n}(Y_i - \hat{g}_{h,i}(X_i))^2 w(X_i).$$

Härdle and Marron (1985a ,1985b) have proved, that under certain conditions choosing h to minimise (10) leads asymptotically to optimal bandwidth selection. That is:

(11) $$\frac{CV(\hat{h})}{\inf_{h\in H_n} ASE(h)} \xrightarrow{\text{a.s.}} 1,\ n \to \infty,$$

where $\hat{h}$ is the bandwidth minimising *CV(h)* and H_n can be described as a set of "reasonable" bandwidths. Cross validation criterion leads asymptotically to choosing the bandwidth which is optimal in the sense of (11) also with respect to other quadratic distance measures mentioned above. Same result holds for the penalising function selectors. However, the computational burden involved in applying these bandwidth selecting criteria can be very large.

Since under fairly general conditions the distribution of the Nadaraya-Watson estimator at k different points converges weakly to the distribution of a multinormal random vector, assuming negligibly small bias compared to the estimator's variance makes it possible to construct approximate pointwise confidence intervals for the true value of $g(x)$ at $x=x_1.,x_2, \ldots, x_k$. The (100-2α) % - confidence interval is

(12) $$\hat{g}(x) \pm c_\alpha \left[\int K^2(u)du\right]^{1/2} \hat{\sigma}(x) / (nh\hat{f}(x))^{1/2} .$$

In the above formula c_α is the (100-α) quantile point of the standard normal distribution, $\hat{\sigma}(x)$ is given by:

(13) $$\hat{\sigma}^2(x) = (nh)^{-1} \sum_{i=1}^{n} \frac{K(\frac{x-X_i}{h})}{\hat{f}(x)} (Y_i - \hat{g}(x))^2 ,$$

and

(14) $$\hat{f}(x) = (nh)^{-1} \sum_{i=1}^{n} K(\frac{x-X_i}{h}) .$$

The confidence intervals constructed in this way ignore bias and will not perform well around "small peaks" for example, where the bias tends to be large. In such cases it will be preferable, of course, to shift the estimated confidence interval by a bias estimate. More advanced techniques for constructing confidence intervals are based on bootstrap simulations.

3 DATA AND ESTIMATION

3.1 THE FINNISH HOUSEHOLD EXPENDITURE SURVEYS

The Finnish household expenditure surveys conducted by Statistics Finland (SF) were used in the estimation of the age profiles. The data are collected in the years 1971, 1976, 1981, 1985 and 1990. The data used have been adjusted by SF, so that the definitions and content of the variables in the different cross-sections are comparable. However, it should be kept in mind that the sampling methods differ somewhat in each survey. The sample units of the survey are households. In 1971 the sample unit was so-called "dwelling unit", which includes all individuals sharing the same dwelling and thus is a broader concept than household. In general the 1971 survey seems to be in many respects different from the subsequent surveys and the results obtained for the year 1971 should be interpreted with care and possibly disregarded if some development or trends over time are investigated . The surveys do not cover individuals living in institutions. This fact is of importance when estimating for example the age profiles of public transfers in kind such as social services, since social services to elderly people living in institutions are not accounted for.

An important feature of the data is that the probability of being included in the sample is not equal for all households. A major reason for this is that naturally there exists no household register in Finland and the sample households were formed around persons selected from the population register. This means that the probability of being selected is proportional to the household size and is larger for larger household. Apart from this, there is significant non random nonresponse among the primarily selected households. For these reasons each household in the sample is given a weight, which is the inverse of the estimated probability of being included in the sample (see Laaksonen, 1988).

3.2 THE ESTIMATION PROCEDURES

The age and cohort profiles are estimated with respect to the age of the head of the household. The head of the household is the person who is primarily responsible for the support of the household (whose income is largest among the household's members). First the profiles for different categories of income, taxes and transfers were estimated in nominal terms. These profiles were converted into real terms (1981 Finnish Markkas) using a price index computed from the data[1].

Welfare comparisons among households must take account of differences in the size and composition of households. For this purpose the age and cohort profiles of the different categories were estimated also per number of consumption units (equivalent adults) in the household. The number of consumption units in the household is computed is computed following a rough equivalence scale recommended by OECD. According to it the first adult in the household comprises 1 consumption unit, and each additional adult comprises 0.7 consumption units. Each child counts for 0.5 consumption units.

The Nadaraya-Watson estimator described in the previous chapter was used for estimating the age profiles, but the observations were weighted by the inverse of the estimated probability of being selected in the sample. The *quartic kernel* was chosen for the estimator. It is given by

$$K(u) = \frac{15}{16}(1-u^2)^2 I(|u| \leq 1) \ . \tag{15}$$

The bandwidth is 4 for all cross-sections. It was selected after fitting the age profile of gross income with bandwidths of 1, 3, 4 and 5. It seems to preserve the essential features of the data without too much "noise" both in the largest and the smallest cross-section. This despite the fact that the sample sizes varies from 2986 in 1971 to 8258 in 1990. The reason is that the age of the head of the household is measured only with one year precision, so difference of bandwidths makes difference in estimation only if the difference is larger than one. Formal bandwidth selection criteria were not used. The calculations were programmed by the author using GAUSS, Aptech Systems, Inc., with details available upon request.

---

[1] The index is computed by weighting the price indices in nine broad categories of consumption goods. The weights are estimated for each household as to take into consideration differences in the relative share of the category in the consumption of different household types. For deflating the estimated age profiles, an average value of the index across households was taken.

4 SUMMARY OF THE RESULTS

The Appendix presents the estimated density function of the age of the head of the household and the age profiles of the following variables:

- household's size in terms of OECD consumption units
- factor income
- disposable income
- disposable income plus consumption of public services
- household's expenditure
- income transfers
- direct taxes and social security payments
- consumption of public services (receipts of transfers in kind)
- net receipts of public transfers.

In the figures the curves are presented only for households with head aged between 20 and 80, since the fraction of households, whose head's age does not fall in this interval is very small and the results obtained for them are quite unreliable. The cohort profiles for the households aged 20, 30, 40, 50 and 60 in 1971 are plotted on each diagram.

The mode of the estimated density function of the household age moves systematically to the right when the results for the 1976-1990 surveys are compared. This clearly reflects the ageing of the baby-boom generation born in Finland after World War II. The results for 1971 are quite different, which, on one hand, is due to the fact that the persons belonging to the baby-boom generation were still young at that time and on the other hand, as already mentioned, the sampling unit was somewhat different in this survey.

The age profiles of the household's size in terms of OECD consumption units suggest, that the average household size has been decreasing in all age groups over time. Exceptions are those households, whose head is under 25 years old: their average size has somewhat increased in 1990 in comparison with the two previous surveys. The estimated age profiles also suggest that a large drop in the average household size has occurred between 1976 and 1981. In fact this observation is most likely to be caused by a statistical discrepancy: there are reasons to believe that in the 1971 and 1976 expenditure surveys the number of households is significantly underestimated, which of course leads to overestimating the average household size.

The age profiles for the different categories of income and for household's expenditure are concave, as expected, when the dependent variable is income per household, with peak around the age of 44. The shape of the age profiles of taxes have, naturally, quite the same characteristics as the age profiles of income. The average disposable income varies with age much less than to factor income. So direct taxes and money transfers make the average income more evenly distributed among households of different age. The shape of the age profiles of disposable income plus consumption of public services (which is equivalent to receipts of transfers in kind) is very much like the shape of the age profiles of disposable income. It seems that transfers in kind do not considerably redistribute average income in the household age dimension. Naturally, the age profiles of the variables per OECD unit are more "even" in comparison with the corresponding age profiles per household, since on there is a positive relationship between income, taxes and expenditure on one hand and household size on the other hand.

The age profiles of income transfers rise quickly after the age of 45 and reach a peak between the age of 65 and 70. This is obviously due to the various pension systems. In the 1985 and 1990 data there is also a small peak approximately at the age of 30-35, which is probably due to the fact that young families with children have lately become recipients of considerable income transfers.

The age profiles of consumption of public services (equivalent to receipts of transfers in kind) clearly show, that the most important recipients are families with children. However, the results probably significantly underestimate the value of transfers in kind received on average by old persons, since people living in institutions are left outside the sample.

According to the estimated age profiles of net receipts of public transfers (money and in kind), the households, whose head is under 40 years old, have gradually become net recipients of income transfers. Thus in 1990 the amount of direct taxes paid exceeds the receipts of public transfers only households with head aged 40-55. One cannot however emphasise this fact too much, since indirect taxes, which are the major source of revenue for the public sector have not been examined.

REFERENCES:

Härdle, W. and Marron, J. S.(1985a), "Asymptotic Nonequivalence of Some Bandwidth Selectors in Nonparametric Regression", *Biometrika, 72, 481-4.*

- (1985b), "Optimal Bandwidth Selection in Nonparametric Regression Function Estimation, *Annals of Statistics, 13, 1465-81.*

Härdle, Wolfgang (1990), "Applied Nonparametric Regression", *Econometric Society Monographs No. 19,* Cambridge University Press.

- (1992), "Smoothing Techniques with Implementation in S", Springer-Verlag.

Laaksonen, Seppo (1988), "Katovirheen korjaus kotitalousaineistossa", Tilastokeskus, Tutkimuksia No 147.

Nadaraya, E. A. (1964), "On Estimating Regression", *Theory of Probability and its Applications, 10, 186-190.*

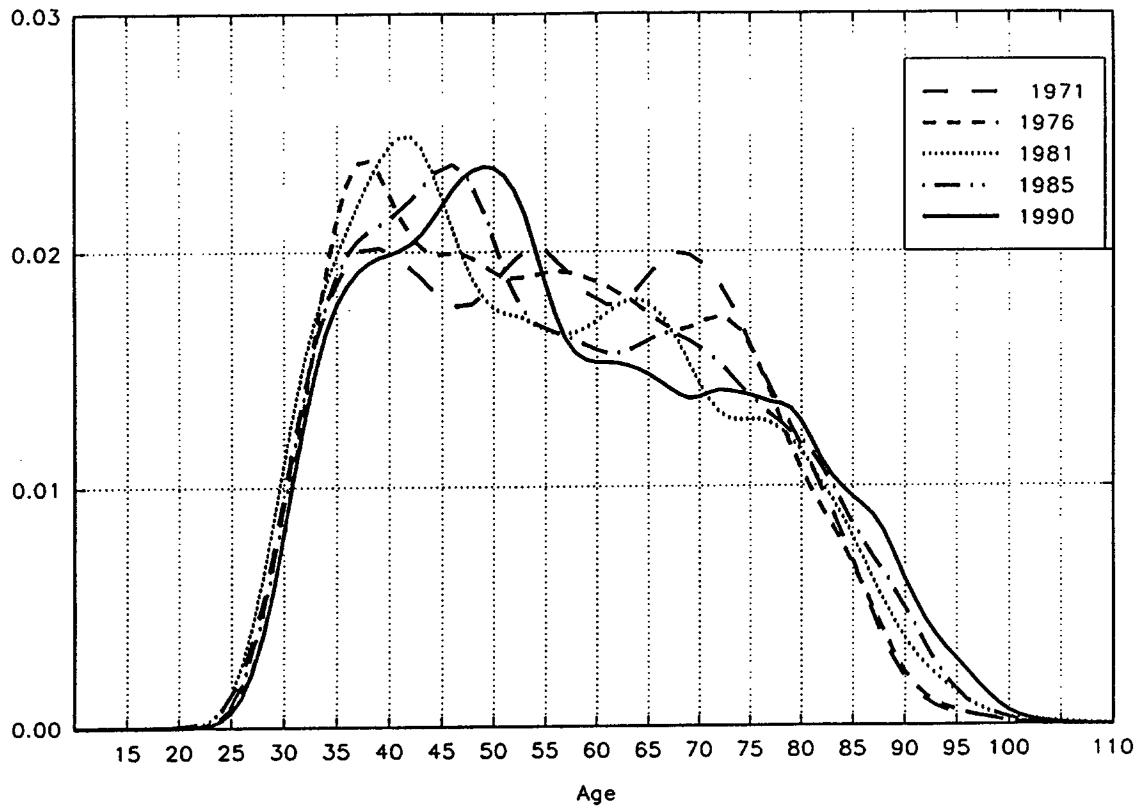
Sullström, Risto - Riihelä, Marja (1993), "Life-Cycle Profiles of Consumption and Income: An analysis of the Finnish household budget surveys in 1966-1990", VATT, Discussion papers No 50.

Parkkinen, Pekka - Mäki, Tuomo - Vanne, Rejo (1996), "Kehdosta hautaan - suomalainen hyvinvointiyhteiskuntansa hoivassa", VATT, Keskusteluaaloitteita No 119.

Watson, G. S. (1964), "Smooth Regression Analysis", *Sankya, Series A, 26, 359-372.*

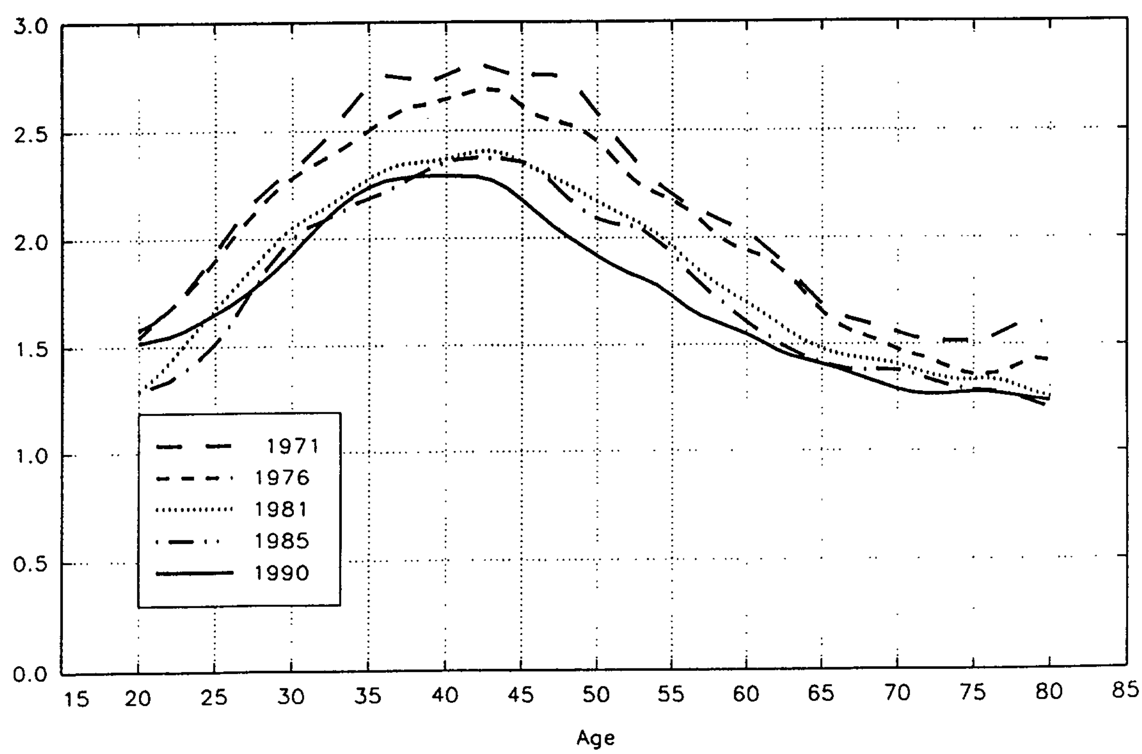
APPENDIX

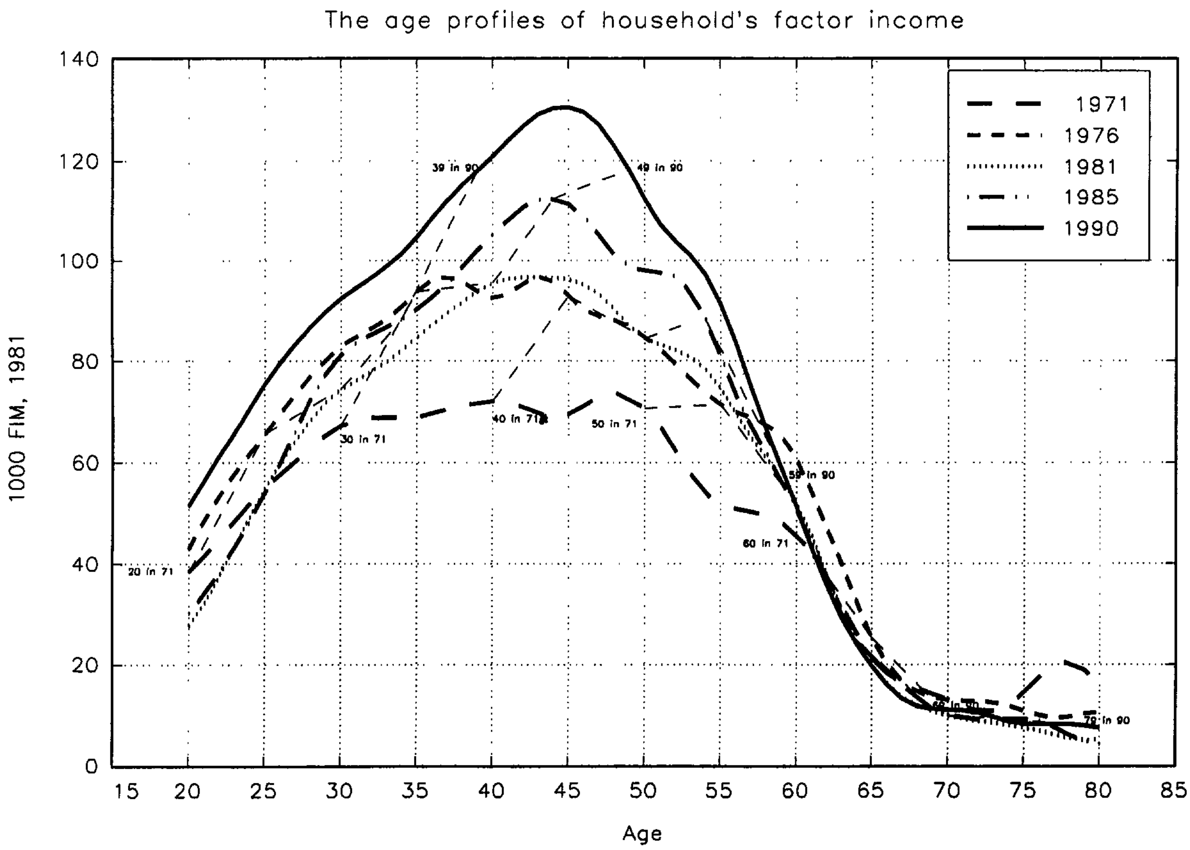
The density function of age of household's head

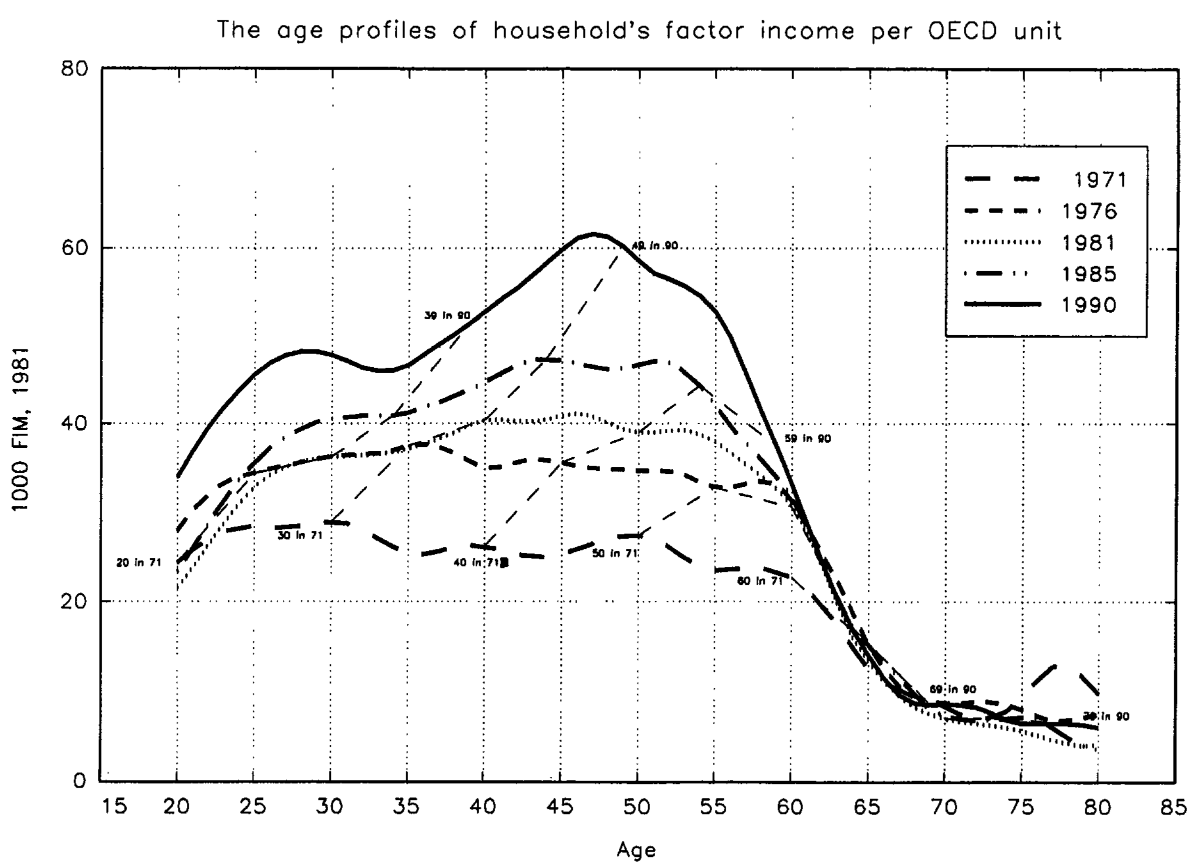


The age profiles of household's size in terms of

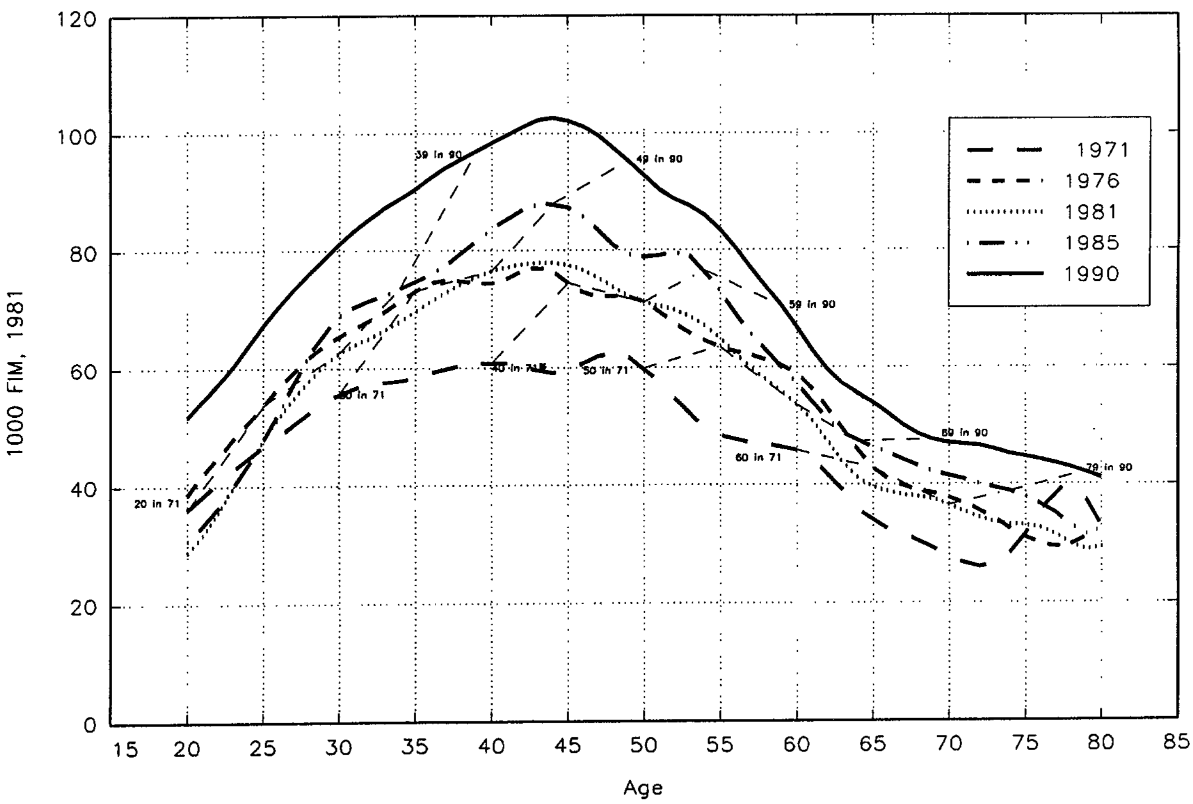
OECD consumption units



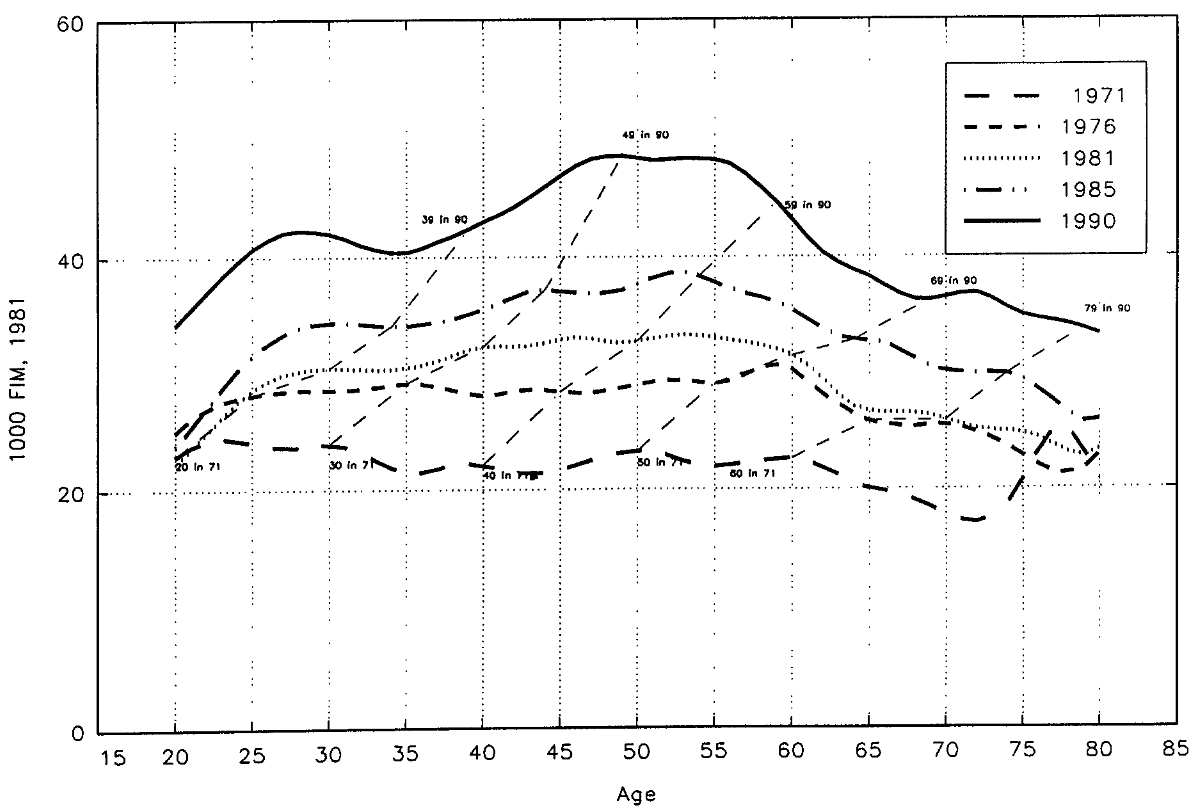

The age profiles of household's factor income
1971
1976
1981
1985
1990
1000 FIM, 1981
Age
39 in 90
49 in 90
30 in 71
40 in 71
50 in 71
59 in 90
60 in 71
20 in 71
69 in 90
79 in 90


The age profiles of household's factor income per OECD unit
1971
1976
1981
1985
1990
1000 FIM, 1981
Age
49 in 90
39 in 90
59 in 90
30 in 71
20 in 71
40 in 71
50 in 71
60 in 71
69 in 90
39 in 90

Age profiles of disposable income

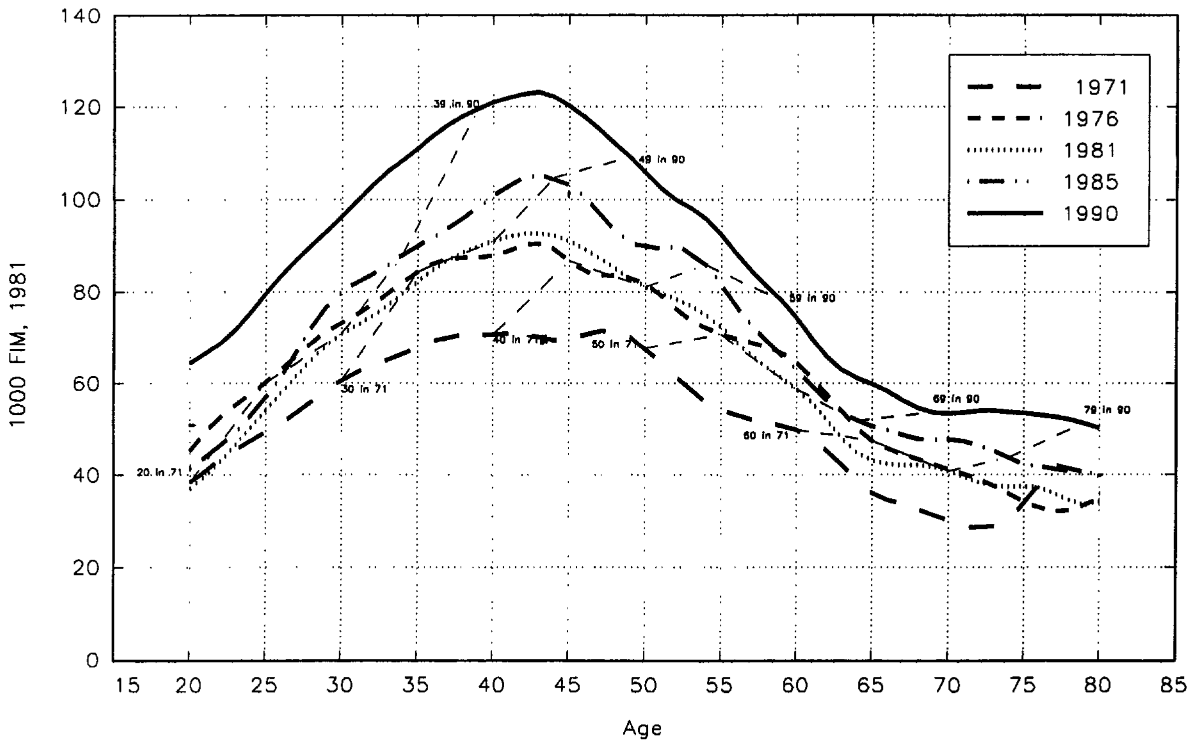


Age profiles of disposable income per OECD consumption unit



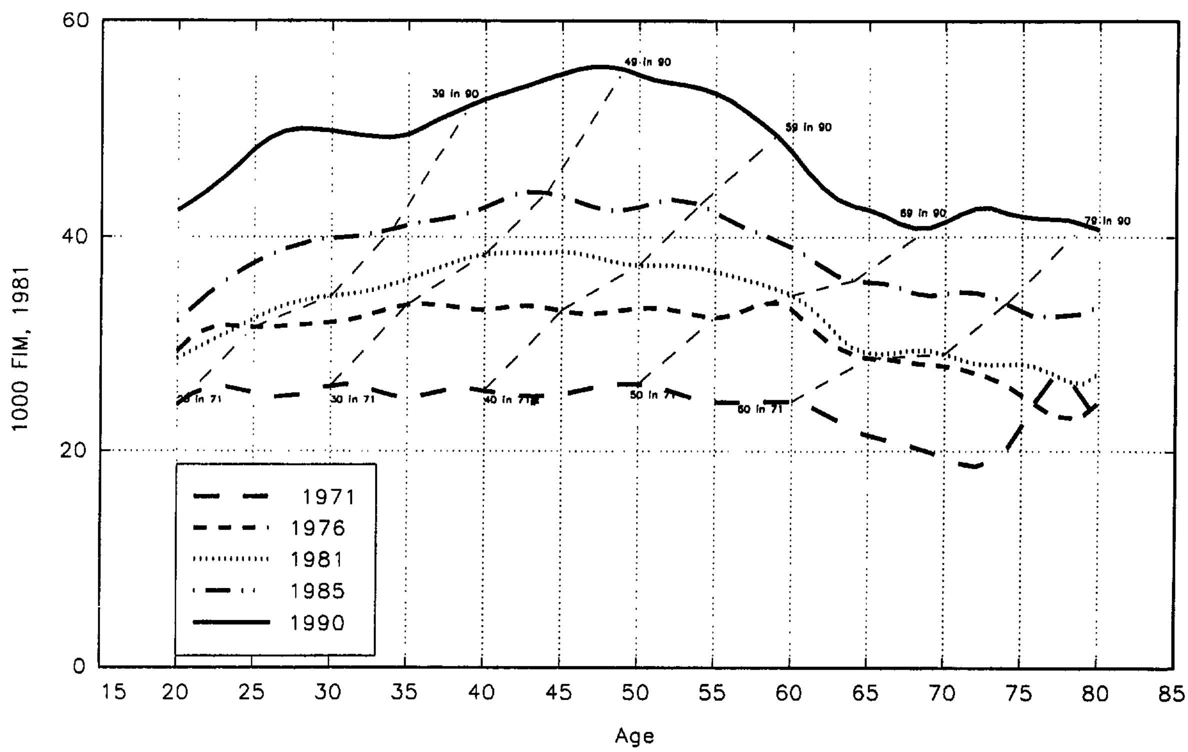
The age profiles of household's disposable income

and consumption of public services

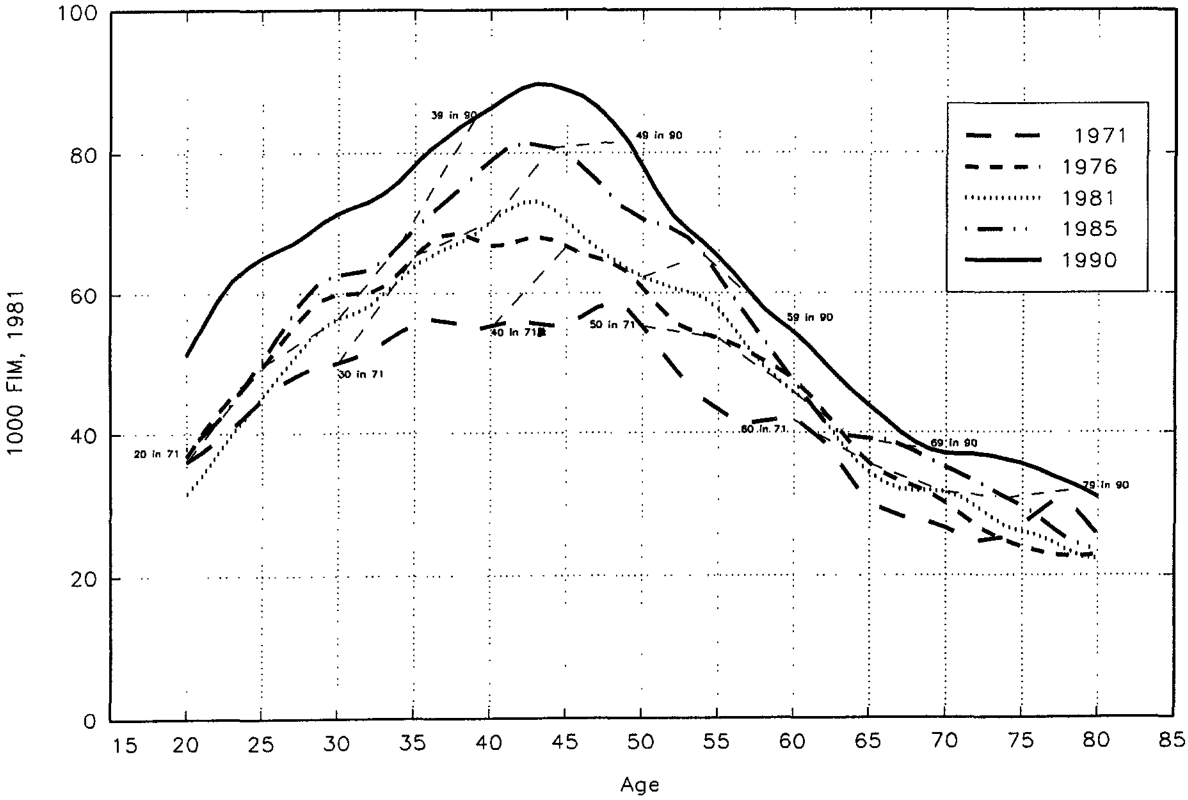


The age profiles of household's disposable income

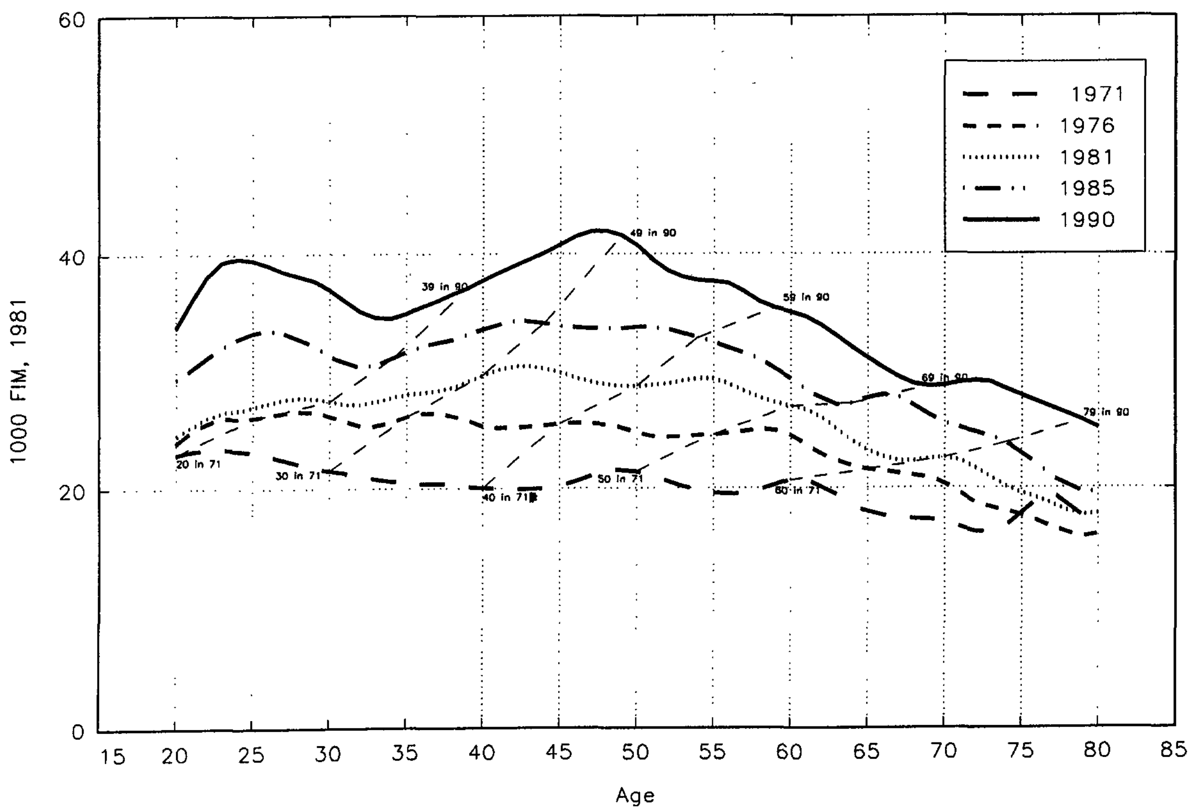
and consumption of public services per OECD unit



The age profiles of household's expenditure

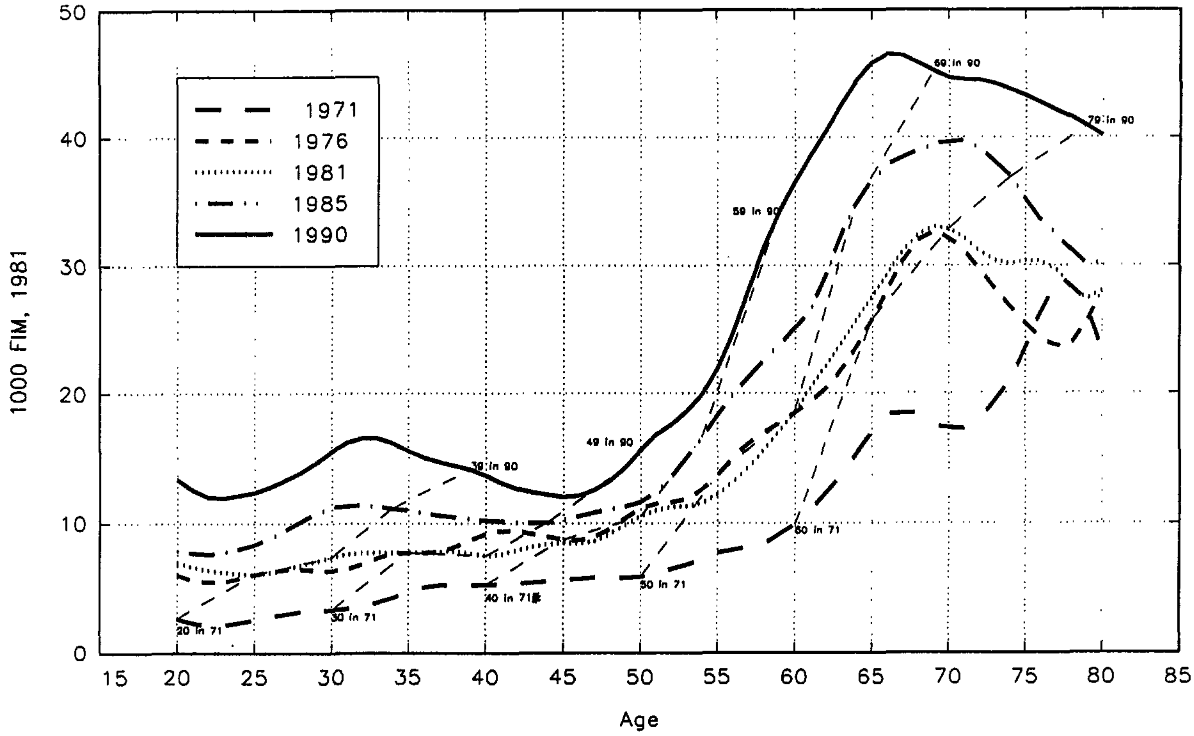


The age profiles of household's expenditure per OECD unit



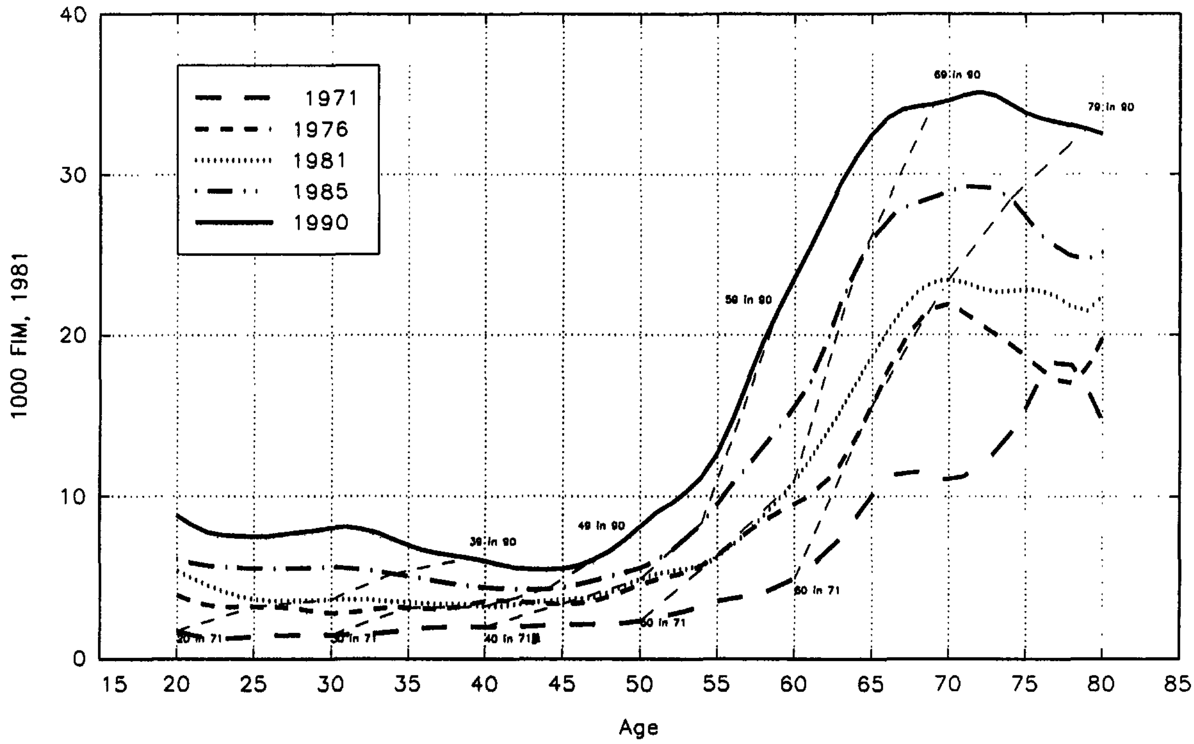
Age profiles of income transfers including transfers

from other households

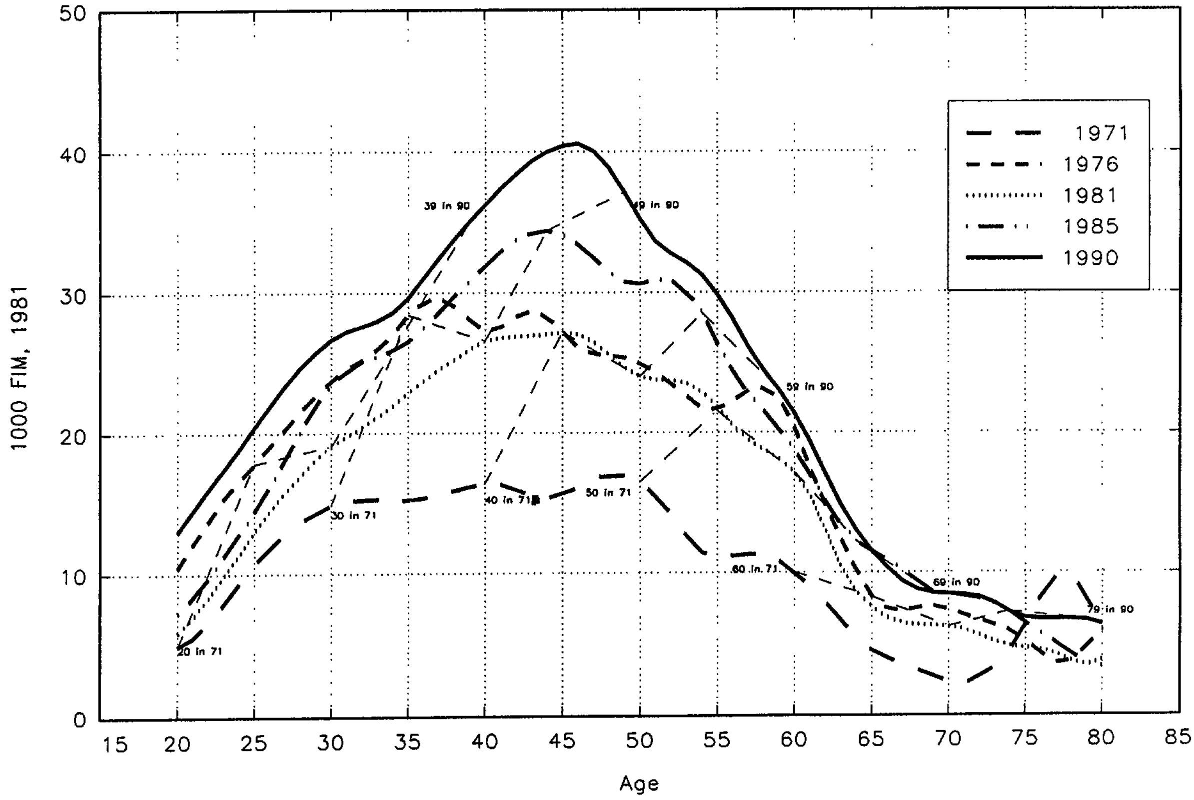


Age profiles of income transfers per OECD unit

including transfers from other households

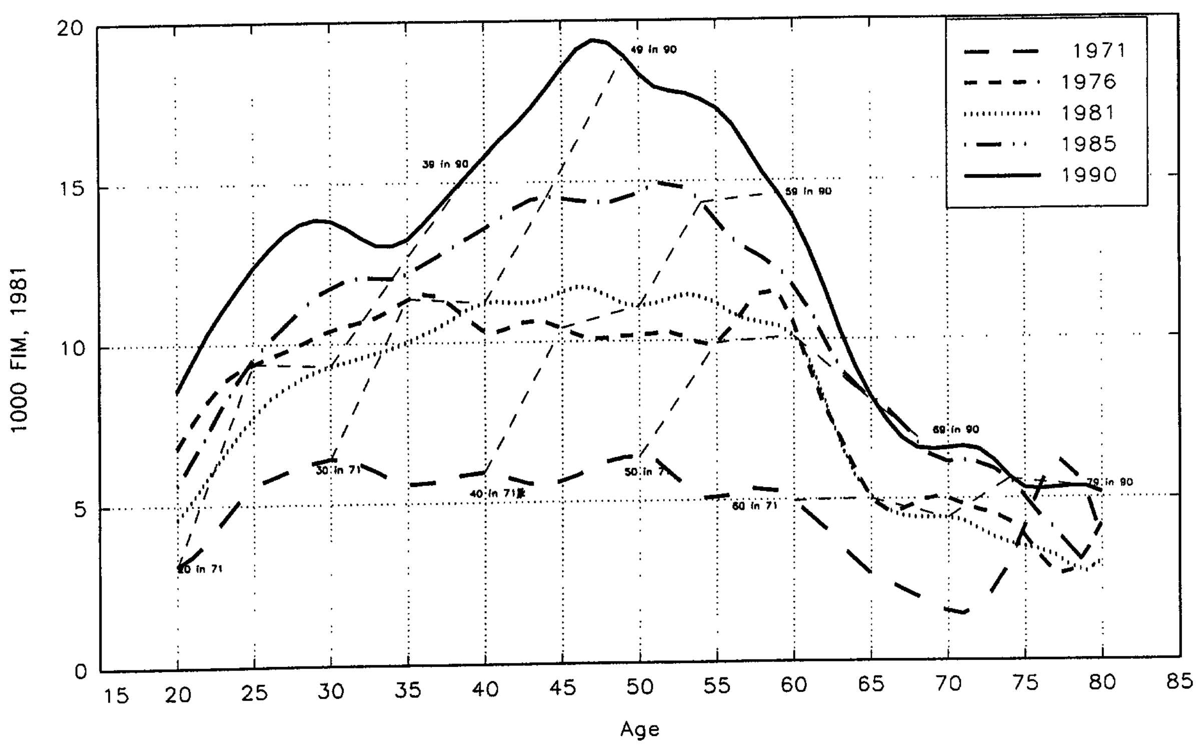


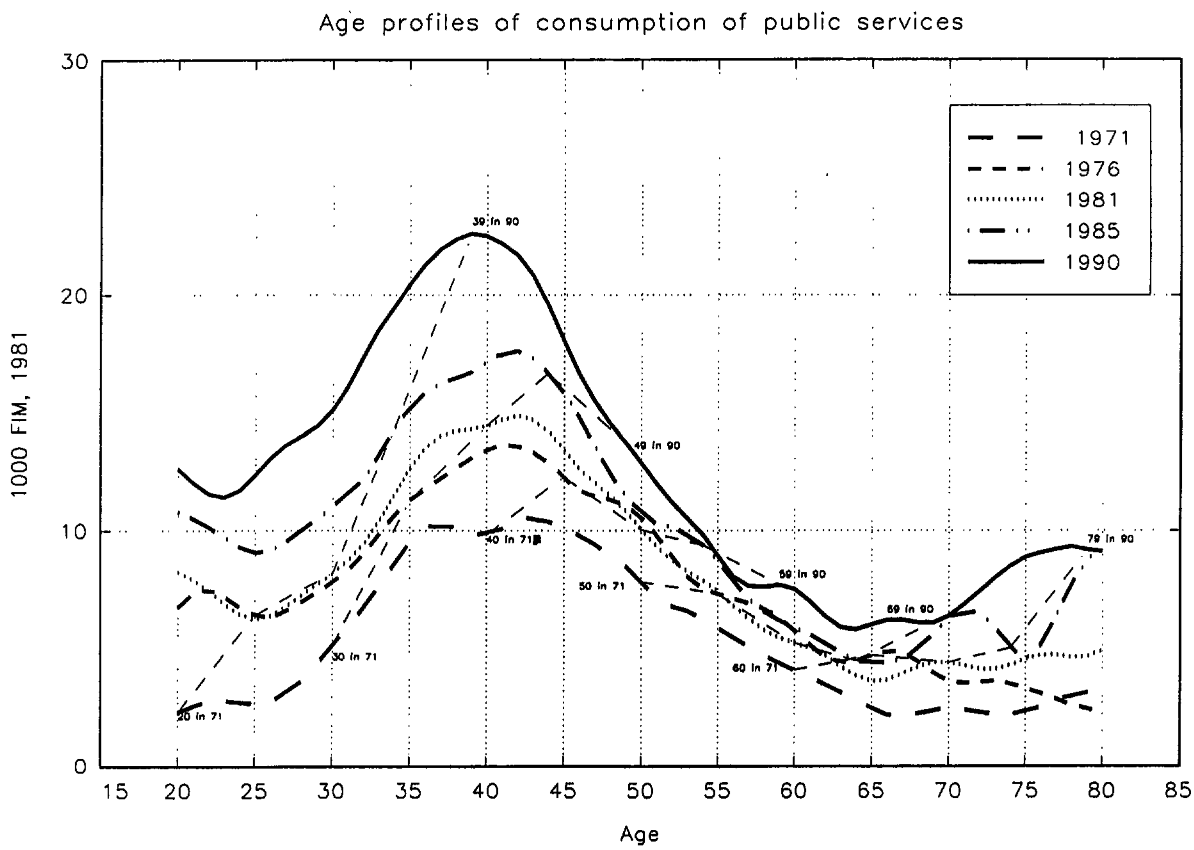
Age profiles of direct taxes and social security payments

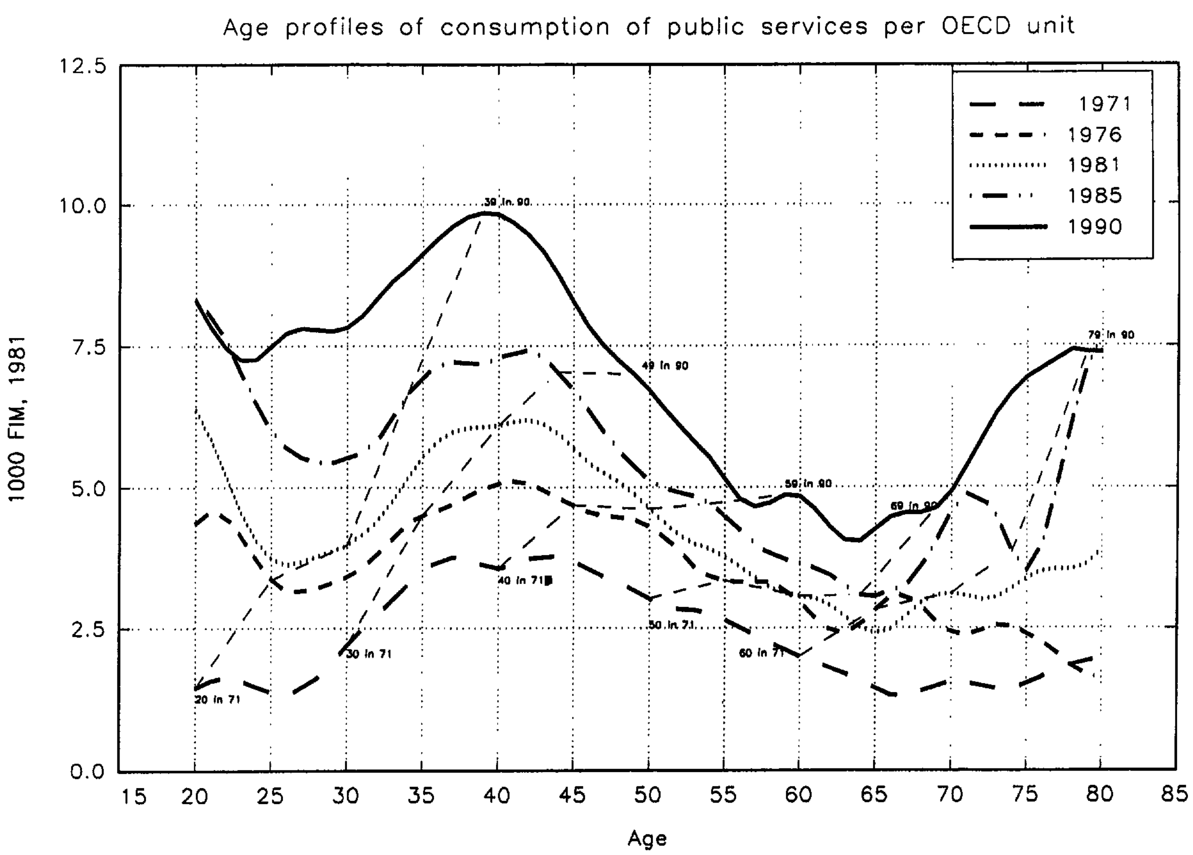


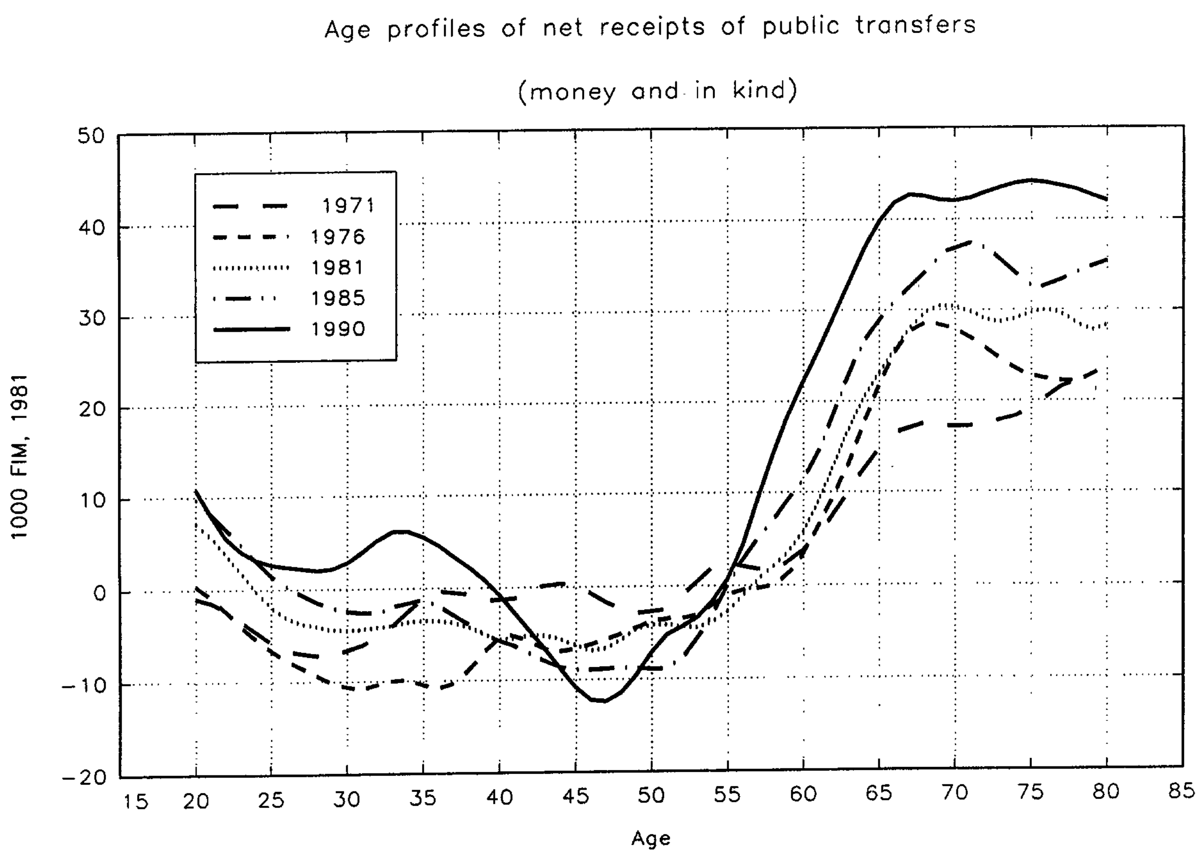
Age profiles of direct taxes and social security payments

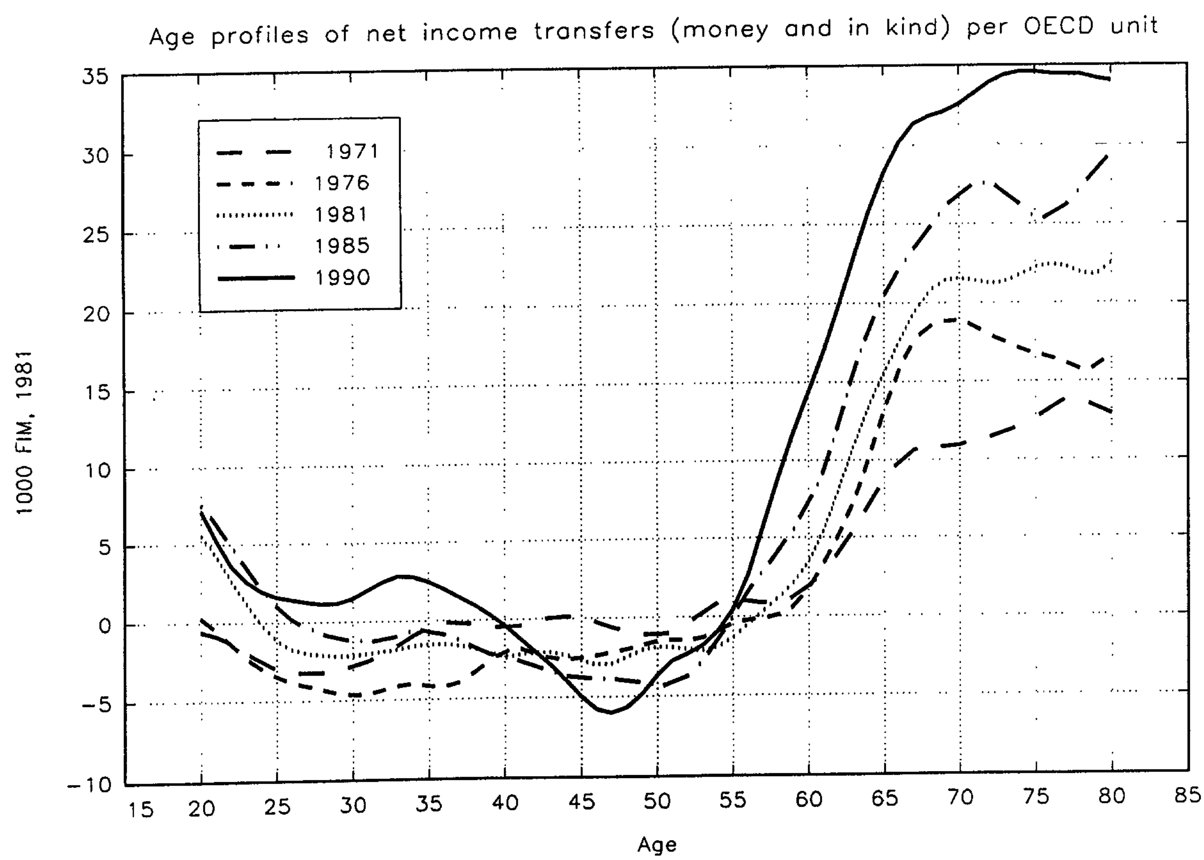
per OECD unit




Age profiles of consumption of public services
1000 FIM, 1981
Age
1971
1976
1981
1985
1990
39 in 90
49 in 90
59 in 90
69 in 90
79 in 90
20 in 71
30 in 71
40 in 71
50 in 71
60 in 71


Age profiles of consumption of public services per OECD unit
1000 FIM, 1981
Age
1971
1976
1981
1985
1990
39 in 90
49 in 90
59 in 90
69 in 90
79 in 90
20 in 71
30 in 71
40 in 71
50 in 71
60 in 71


Age profiles of net receipts of public transfers
(money and in kind)
1971
1976
1981
1985
1990
1000 FIM, 1981
Age
50
40
30
20
10
0
-10
-20
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85


Age profiles of net income transfers (money and in kind) per OECD unit
1971
1976
1981
1985
1990
1000 FIM, 1981
Age
35
30
25
20
15
10
5
0
-5
-10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85