Consumption and Nutrition: Age – Intake Profiles for Czechoslovakia 1989-1992

Ruth Miquel[*], François Laisney[†],

[*] University of St Gallen, Switzerland
[†] University Louis Pasteur of Strasbourg, France and ZEW, Mannheim, Germany

September 2000

Abstract

This paper provides non-parametric estimates of the relation between nutrient intake and age for Czechoslovak individuals, as a function of characteristics of both the individual and the household she lives in, on the basis of household purchases. Results show no significant difference between the age - energy intake profiles of men and women. The decomposition of this intake between carbohydrates, lipids (i.e. fats) and proteins shows a lack of balance in the diet in Czechoslovakia, but significant progress toward a more balance diet has taken place over the period. Finally, household characteristics such as the woman's level of education, or household income, have at most a marginal impact on these profiles.

Correspondence: Ruth Miquel: University of St. Gallen, Swiss Institute for International Economic Relations and Applied Economic Research (SIAW), Dufourstr. 48 CH-9000 St.Gallen, Switzerland
e-mail: Ruth.Miquel@unisg.ch

François Laisney: BETA-THEME, Université Louis Pasteur, Strasbourg, 61 Avenue de la Forêt Noire, F67000 Strasbourg, France, and ZEW, Mannheim, FRG
e-mail: laisney@zew.de

Non-technical summary

While health receives increasing attention from the economic profession, one can argue that the empirical analysis of household budgets has been motivated at the outset by an important aspect of the study of health status, connected with nutrient intakes: this is the topic of the seminal study of Engel (1857). One way to model the relation between the economic and nutritional status of households is to posit a causal relationship between these two. The aim of studies using this representation is to evaluate the Engel curve for *calories*. Unfortunately, the conclusions of empirical studies of this type diverge starkly, mainly for the following reasons, discussed by Bouis and Haddad (1992): (i) the data come from different populations, (ii) different estimation methods are used, and (iii) the variables included in the relation between calories and income differ. In particular, choosing calorie availability as dependent variable, as opposed to calorie intake yields biased results.

However, it is difficult to obtain calorie intake data, because their construction is very demanding. Chesher (1997) proposes a model for the study of mean calorie intake as a function of sex and age, using only household purchases, and not the individual consumption. Since household budget data sets are much more common than individual consumption data sets, this make Chesher's approach widely applicable.

In this study, we use that approach to study the nutrient intakes of Czechoslovak individuals as a function of age and sex, the evolution of this quantity across time (1989-1992) and the influence of current income and educational level of the wife on these curves. The nutrient we study is the daily energy intake (Kcal/d). We obtain the following substantive results.

The age profiles of the energy intake of Czechoslovakians are similar for males and females. The overall energy intake increases with age until 55, is approximately constant between 55 and 70, and decreases afterwards. The contribution of carbohydrates to the energy intake is below the 50% recommended for a balanced diet, whereas the contribution of proteins is above the advisable 12% level. However, these imbalances have markedly decreased over the four-year period studied. The contribution of lipids to the total energy intake matches the advised level of 30%. In households based on a couple, the education level of the woman plays almost no role in the quantity of calories ingested, but a higher education level of the wife leads to a slight moderation of energy intake. The impact of income on the nutrient intakes is not consistently significant over time for proteins, lipids and carbohydrates. However, it is positive for beverages. Overall, the age profile of energy intake seems to follow the general variations of income over the life cycle (increasing during the working life, with income, decreasing for pensioners).

Beside this substantive contribution, the paper also provides a methodological contribution, as we adapt Chesher's approach to the particularities of the data at hand in a way that should prove useful to other researchers.

1 Introduction

Questions concerning health and the medical status of populations receive increasing attention from the economic profession, and especially from the econometricians, as documented by the recent special issue of the *Journal of Econometrics* (see Bhargava, 1997). Indeed, one could argue that empirical demand analysis has even been motivated at the outset by an important aspect of the study of health status, connected with nutrient intakes, as the seminal study of Engel (1857) shows.

One way to model the relation between the economic and nutritional status of the household is to suppose the existence of a causal relationship between these two, that is to say to postulate that the nutritional inflow is conditioned by the income and the demand for food. The aim of studies using this representation is to evaluate the Engel curve for calories. Effectively, a precise estimation of income elasticities for the demand for calories is needed for the guidance of nutritional policies. However, the conclusions of empirical studies of this type diverge starkly. On the basis of elasticities between 0.6 and 0.9 (Strauss, 1982, Pitt, 1983), the World Bank (1981) advocates an increase of poor people's income as the most efficient long term policy to suppress malnutrition. Yet Behrman and Wolfe (1984), and Subramanian and Deaton (1996) find elasticities which, although strictly positive, are as low as 0.12 for the first and 0.14 for the second. Deaton and Subramanian have also estimated the cost of calories in Maharashtra (India) and pointed out that the energy intake needed for daily activities can be obtained for about 4% of the daily wages, calling into question the *malnutrition trap* (see for example Dasgupta, 1997). Finally, Behrman and Deolalikar (1987), Bouis and Haddad (1992), and Bouis (1994) found surprising income elasticities for the demand for calories near zero, implying the inefficiency of the income increase policy to improve the nutritional diet.

Behrman and Deolalikar explain these differences in estimates by a problem of aggregation between food and nutriment. They compare *direct* estimation of the calories income elasticity, obtained by conversion of food groups into aggregate calories prior to estimation, and *indirect* estimation, involving the post-estimation computation of a weighted average of the expenditure elasticities by groups of food. For the same households, they find positive but insignificant direct elasticities and significant indirect elasticities around 1. They then conclude that level of the aggregation at which we make the conversion food-nutriment is of paramount importance and argue in favour of the direct method, which is more sensitive to the substitution between high nutritional food and low nutritional food for the same group of goods (some of them could be cheaper than the others).

Bouis and Haddad (1992) put forward three additional reasons which could explain these surprising results: (i) the data come from different populations, (ii) different estimation methods are used, and (iii) (the most important) the variables included in the relation between calories and income differ. They make a distinction between four possible variables: the current income or the total expenditures, and the calorie availability and the calorie intake. They find that choosing calorie availability as dependent variable yields upward bias estimations. In fact the problem rests with the available data: food surveys typically give calorie availability, whereas "24 hours recall" inquiries give calorie intake.[1]

However, it is difficult to obtain calorie intake data, because their construction is very demanding. Moreover, collected data suffer from measurement error as people find it difficult to remember exactly all the different kinds of foods they have eaten and all their quantities;

---

[1] As the name of the second type of data indicates, people report quantities eaten during the 24 last hours.

the respondents may also not want to tell the truth, i.e. they may not wish to reveal their eating habits. Chesher (1997) proposes a model for the study of mean calorie intakes (and other nutrients such as fat, calcium, etc.) as functions of sex and age, using only household purchases (and not the individual consumption).[2] Estimation is based on penalised least squares, with smoothing over age, or over both age and time. Chesher's results for Great Britain are that, when he controls for household characteristics such as income or location, the nutrient intakes decrease slightly for old people and for children, but the effect of household characteristics is weak and the profile of the curves does not change much with these. Typically, one observes a quick growth of energy intake with age, with a peak around 14 years, a decrease for people aged 25-30 and a new quick growth around 50-60 years of age. Chesher (1998) is interested in the evolution over time of this relation between nutrient intake and age, and shows an important decline of the average energy intake for men and women between 1975 and 1992. His model relates household food purchases during a survey period to the average nutrient intakes of different types of individuals. For each sex, age-intake relationships are estimated non-parametrically (i.e. without specifying a parametric form for the relation between age and energy intake).

In this study, we use a new data set with the methodology developed by Chesher to study the nutrient intakes of Czechoslovak individuals as a function of age and sex, the evolution of this quantity across time (1989-1992) and the influence of the current income and the educational level of the wife on these curves. Beside this substantive contribution, the paper also provides a methodological contribution, as we adapt Chesher's approach to the particularities of the data at hand in a way that should prove useful to other researchers.

The nutrient we study is the daily energy intake (Kcal/d). Ideally, in a balanced diet, this intake should have the following composition (alcohol not included): proteins 12%, lipids 30-35%, carbohydrates 50-55%.[3] The observed intake pattern is in most cases very different from the ideal one (see e.g. Chesher, 1997, for Great Britain, and Dupin and al., 1992, for France), with a marked excess of lipids. Moreover, the dominant sedentary lifestyle encourages weight excess and different kinds of metabolic troubles. A lot of people react by decreasing their consumption of foods of plant origin, full of starch, whereas they ought to first decrease their consumption of lipids. The excessive intake of lipids is suspected to contribute to increased heart disease. Another important nutritional imbalance, not often mentioned, is the insufficient contribution of the carbohydrates. The question is not to increase the proportion of carbohydrate with quick absorption, such as sucrose, but to increase the proportion of starch, the intestinal absorption of which is slow and gradual, and therefore beneficial for a metabolic use. Ideally, the carbohydrate contribution must come 90% from slow sugars and 10% from quick sugars. The composition of the diet is of public interest, as prevention is much cheaper than cure.

In our analysis we will focus on the decomposition of carbohydrate intake between the contribution of starchy food (slow sugars) and the contribution of sweet food, fruits included (quick sugars).[4]

Section 2 introduces the data, Section 3 presents the model used. The results are discussed in Section 4, and Section 5 concludes.

---

[2] This is important: since household budget data sets are much more common than individual consumption data sets, it make Chesher's approach widely applicable.

[3] Throughout the paper, lipids refer to fats.

[4] Starch: potatoes, bread, other bakery products, flour, other cereal products, rice and pasta, dried vegetables.
  Sugars: sugar, chocolate and fruits.

2 Data

The data used come from the Czechoslovak Budget Survey for the years 1989-1992. The data were collected by groups of 2 years, 1989-1990 and 1991-1992, with some differences between the two groups to which we shall return later. The data has a panel structure over each pair of years, but unfortunately the household identifier is discarded between years, and thus considerable effort would be necessary to recover the panel structure (at best with imprecision). This is unfortunate because it will put the potential efficiency gains and control over unobserved heterogeneity to be obtained from using the panel structure out of our reach. It will also make inferences which are not restricted to a single year questionable.

The data contain detailed socio-demographic information concerning the household head and his partner[5] (if the household head is married), such as sex, age, education level, and incomes from various sources, and also information on household characteristics such as number of household members. This concerns 5679 households for 1989, 5776 for 1990, 3948 for 1991 and 3336 for 1992. Table 1 provides information on the composition on these households. Important information such as sex or age is missing for certain members of the household. For example, if the household is composed of two sisters we only know that the household head lives with someone else, but we do not know that it is with a person of the same sex, and even that this person is her sister. This will lead us to make some assumptions on household composition. These will be presented in section 4.

Table 1: Household composition in each sample:

	year	min	Mean	max	st.dev.	Number of households in sample
	1989	1	2.89	8	1.32	5679
Number of	1990	1	2.88	8	1.32	5777
members	1991	1	2.77	7	1.25	3948
	1992	1	2.74	7	1.26	3336
	1989	0	1.04	6	1.08	5679
Number of	1990	0	1.03	6	1.08	5777
children	1991	0	.92	5	1.00	3948
	1992	0	.90	5	1.00	3336

We define the current income of the household as earned income, plus benefits and pensions, minus taxes, in nominal terms. For a small proportion of the population ($< 0.5\%$) this income is negative and the households involved are deleted from the sample for the part of the analysis that uses household income. We will also consider level of education of the wife. The data present 8 different levels, from "uncompleted primary school, no education" to "post-graduate education or scientific degree".

The data report the purchased quantities for a period of 12 months. This is a great advantage over household surveys which only record expenditures or quantities purchased over a short period, as the discrepancy between quantities purchased and consumption over a given period diminishes starkly with the length of the period considered. However, some households have not provided answers over the full 12 months, and we have decided to discard those in order

---

[5] When the household is a couple, the man is always the household head.

to avoid the problems connected with modelling infrequent purchases. Another advantage of this choice, which is especially interesting in the case of food consumption, is that keeping households observed over a part of the year only would introduce seasonal heterogeneity, which we completely eliminate here. Restricting attention to households who have reported quantities for all 12 months leads to the loss of 335 observations for 1989, 502 for 1990, 480 for 1991 and 421 for 1992. Food is disaggregated into 37 categories including alcohol, beverages, and sweets.[6] The quantities reported are not limited to purchases but contain goods available "in kind", mainly produced by the households themselves.

Although the numbers of meals taken outside the households are reported (restaurants, canteens,…), we do not know the quantities of the different kinds of goods ingested on those occasions. Thus, in order to convert the food ingested into nutrients, we use only the quantities of food purchased and the "in kind" quantities, and do not take the nutrient intakes from outside meals into account. We use the conversion factors proposed in the "Tables de composition des aliments" of the Institut Scientifique d'Hygiène Alimentaire (1990) and in Noël (1988). These conversion factors are more detailed than our food groups, and thus we use the arithmetic mean of the different factors included in each, (or the mean of the factors of different pieces when the food is some meat).

3 Model

The model we use to estimate the energy intake – age profiles was proposed by Chesher (1997). For a household with N individuals, two types of characteristics must be taken into account: the individual characteristics of person p (age, sex,…) at time t, x_{pt}, and the household characteristics (income, region of residence, etc.), Z_t. We are interested in the expected individual consumption of nutrient at time t, c_{pt}, conditional on these two variables, Z_t and $x_t = (x_{1t}, x_{2t}, …, x_{Nt})$:

$$(1) \qquad E\left(c_{pt} \mid x_t, Z_t\right) = f\left(x_{pt}, x_{-pt}, Z_t\right),$$

where $x_{-pt} = (x_{1t}, x_{2t}, …, x_{p-1t}, x_{p+1t}, …, x_{Nt})$.

The nutrient consumption of the household, c_t, is the sum of the individual consumptions of its members. The conditional expectation of this consumption given x_t and Z_t is then

$$(2) \qquad E\left(c_t \mid x_t, Z_t\right) = \sum_{p=1}^{N} f(x_{pt}, x_{-pt}, Z_t)$$

But with the data used, as with most data sets available, only total household consumption is available. Thus an indirect method is needed to estimate the function $f(x_{pt}, x_{-pt}, Z_t)$. We use the conditional moments

$$(3) \qquad E\left[\left\{c_t - \sum_{p=1}^{N} f\left(x_{pt}, x_{-pt}, Z_t\right)\right\} g(x_t, Z_t)\right] = 0,$$

where $g(x_t, Z_t)$ is an arbitrary square integrable function of x_t and Z_t.

---

[6] We did not include alcohol and beverages in the estimation, except for the last section (4.3) where we take into account the households' characteristics.

The nutrient consumption of the household, c_t, is obtained from the household purchases. The purchased quantities of different goods are translated into nutrient quantities. q^c_{it} represents the quantity of good i consumed by a household and α_i the quantity of target nutrient included in one unit of good i. The quantity of nutrient consumed by the household is $c_t = \sum_i \alpha_i q^c_{it}$.

All this will apply directly if we observe the quantity consumed q^c_{it}. However, we only have q_{it} the purchased quantity for all kinds of goods. These two quantities can differ because consumption and purchases are not realised at the same time. But as the period covered by each interview is long enough (12 months) to allow the periods where the stocks increase (before celebrations for example) to be compensated by the periods where the stocks decrease (the purchases are less important during these periods), we can assume that the expected quantity of each good in the purchases is equal to the expected consumption of these goods, for each x_t and Z_t. Under this assumption, the quantity of nutrient present in the goods purchased, $y_t = \sum_i \alpha_i q_{it}$ is such that

$$(4) \qquad E\left(y_t \mid x_t, Z_t\right) = E\left(c_t \mid x_t, Z_t\right) = \sum_{p=1}^{N} f(x_{pt}, x_{-pt}, Z_t) .$$

The moment conditions (3) are then equivalent to

$$(5) \qquad E\left[\left\{ y_t - \sum_{p=1}^{N} f\left(x_{pt}, x_{-pt}, Z_t\right) \right\} g(x_t, Z_t) \right] = 0 .$$

For simplification, Chesher assumes that the consumption function can be expressed as the product of a function of individual characteristics and a function of household characteristics:

$$(6) \qquad f\left(x_{pt}, x_{-pt}, Z_t\right) = \tilde{f}\left(x_{pt}\right) h(Z_t),$$

where $h(Z_t)$ can include x_{-pt}. This assumption implies that the ratios of individual consumptions, $f(x_{pt}, x_{-pt}, Z_t)/f(x_{p't}, x_{-p't}, Z_t)$, with p and p' denoting two different individuals in the household, are independent of household composition. It thus precludes any influence of the age of children on parents' consumption. In support of this hypothesis, Chesher refers to a paper by Mills and Tyler (1992) and a report of the Department of Health (1989).

The individual characteristics taken into account are age (a_{pt} for person p at time t) and sex ($G_p = 1$ if person p is a male and 0 if p is a female). It would seem arbitrary to assume a priori that the diets of men and women are identical (and indeed Chesher, 1997, and Mills and Tyler, 1992, do report significant differences), so that we estimate different nutrient intakes-age functions for men and women:

$$(7) \qquad \tilde{f}\left(x_{pt}\right) = G_P f_M\left(a_{pt}\right) + (1 - G_p) f_F\left(a_{pt}\right)$$

where $f_M(.)$ and $f_F(.)$ are the nutrient intake-age functions for men and women, respectively.

The nutrient intake is related in a complex way to the different periods of life. It varies for example with the level of activity, with tastes, etc. This is why Chesher (1997 and 1998) uses a non-parametric approach to estimate this relation.

4 Estimation

The unknown functions, $f_M (.)$ and $f_F (.)$, are replaced by discrete approximations, more precisely by step functions potentially changing for each age group (possibly corresponding to each year of age).[7]

Let $w_{pt,a} =1$ if $a_{pt} =a$, that is if the age of person p at time t is a, and 0 otherwise. The functions $f(.)$ are then replaced by $f_G (a_{pt}) = w_{pt,a} \beta^G_{a,t}$. So the objective (4) becomes

$$E (y_t |x_t, Z_t) = (\beta^0_t + \sum_{p=1}^{N_t} \sum_{a=0}^{T} \{G_p\, w_{pt,a}\, \beta^M_{a,t} + (1-G_p)\, w_{pt,a}\, \beta^F_{a,t}\})\, h(Z_t)$$

(8)
$$= [\, \beta^0_t + \sum_{a=0}^{T} n^M_{a,t}\, \beta^M_{a,t} + n^F_{a,t}\, \beta^F_{a,t})\,]\, h(Z_t)$$

where n^G_{at} is the number of household members of sex G, age a at time t and T denotes the maximal age considered. The constant β^0_t is included to capture the flow of nutrient in the household which is not consumed by any of its members (food eaten by domestic animals but bought for the humans, spoiled food, etc.).

As we have mentioned before, the sex of a person is only available for the household head (and for his wife, if the head is married) in the Czechoslovak data. Moreover we have no information on the ages of the children and of the additional adults (i.e. adults present in the household other than the head and his spouse), except that we know the age group of each child (0-5, 6-9, 10-14), and that the additional adults are 15 years old or more. These missing data oblige us to make some simplifying, and admittedly largely ad hoc, assumptions.[8] The first one is obviously that for children we will use the age groups and not each year of age separately. Moreover:

H1: For the children aged between 0 and 14 years of age, we make no distinction between boys and girls.

This assumption of no differentiation of nutrient intake by sex can be justified by the results of Bhalotra and Attfield (1998). They find no significant differences due to sex in children's consumption of food for Pakistan.

The following assumptions concern only the households with additional adults. These households represent 7.66 % of the sample (435 households) for 1989, 6.84 % (395 households) for 1990, 6.18 % (244 households) for 1991 and 7.31 % (244 households) for 1992.

H2: We group the young people aged between 15 and 18 into one class, without making a distinction between sexes, as for the children. We make the same assumption for people with age above 80.

---

[7] The adjustment needed for our data will be given precisely in the next section.
[8] The imputations are simply based on common sense considerations, and not on extraneous data. However, we do conduct a limited amount of sensitivity analysis, presenting some results for the sub-sample of households without additional adult members (see Figures 6a-b and Tables 3a-b below).

H3: We assume that if the additional adult in the household is young, she is at most 25. In others words, we assume the existence of an age ceiling for the young adults staying at home.

H4: For old people, we assume that there exists an age floor (65) when these people are additional adults.

H5: We replace the dummy $w_{pt,a}$ by its expectation conditional on the age of the household head and the presence of additional members in the household:

$$E\ (w_{pt,a} \mid cfyear) = P[\ w_{pt,a} = 1\ /\ cfyear\ ]\ nbsup$$
$$= \pi\ (a)\ nbsup$$

where *cfyear* is the age of the household head, *nbsup* is the number of additional adults in the household. [9] We also assume $a \in [15, 25]$ if *cfyear* $\leq$ 55, that is to say that the additional adults are young if the household head is at most 55 years old, and $a \in [65, 81+]$ if *cfyear* > 55: the additional adults are old people if the household head is over 55.

H6: We impose some restrictions on $\pi\ (a)$:

- $\int_{\underline{a_j}}^{\overline{a_j}} \pi_j(a)\ da = 1$ if *cfyear* $\leq$ 55, 0 otherwise, $\int_{\underline{a_v}}^{\overline{a_v}} \pi_v(a)\ da = 1$ if *cfyear* > 55, 0 otherwise, $[\underline{a_j}\ ,\ \overline{a_j}\ ] = [15\ ,\ 25]$ $[\underline{a_j}\ ,\ \overline{a_v}\ ] = [65\ ,\ 81+]$.
- The conditional probability for each additional person to be in the interval [$15\ ,\ 18$] is set at 1/2 if *cfyear* $\leq$ 55.
- If $a \in [19\ ,\ 25]$, $\pi_j(a)$ is linearly decreasing. Given the first two points above, $\int_{19}^{25} \pi_j(a)\ da = 1/2$ and then we have $\pi(a) = (25/42) - (1/42)a$.[10]
- The conditional probability for each additional person to be over 81 is set at 1/2 if *cfyear* > 55.
- If $a \in [65\ ,\ 80]$ $\pi_v(a)$ increases linearly. With the previous restrictions we get $\pi(a) = (1/240)a - (65/240)$.

For estimation, Chesher uses a "penalised" estimation method (Silverman (1986)), as proposed by Engle et al. (1986) in their estimation of the relation between the daily consumption of electricity and the daily temperature. This spline smoothing approach quantifies the trade-off between the aim to produce a good fit to the data and the aim to produce a curve without too much local variation. To obtain a function $r(x_i)$ which is the "best" approximation of the data Y_i, this method minimises the sum

---

[9] We define the number of adult members other than the household head and his spouse as the difference between the number of household members and the number of children plus 2 in the household formed by one couple (number of members – number of children – 2 (for the household head and his spouse), in the household without couple as the difference between the number of household members and the number of children plus 1. For some households (29 for 1991 for example), this number is set to 0 because one of the two spouses is not present in the household during part of the year: the application of the formula would give a negative number for the additional adults.

[10] $\pi(a) = \pi(\underline{a}) - [(\pi(\underline{a}) - \pi(\overline{a}\ ))/\ (\overline{a} - \underline{a}\ )](a - \underline{a}\ )$.

(9)
$$S_\lambda = \sum_{i=1}^{n} (Y_i - r(x_i))^2 + \lambda \int_{\Re} [r''(x)]^2 \, dx \, ,$$

where λ is a smoothing parameter and $\int_{\Re} [r''(x)]^2 \, dx$ denotes the penalty used against local variation.

4.1 Cross-sectional estimation

In this sub-section we ignore the household characteristics, Z_t, to simplify.

4.1.A Separating Men and Women

For this study as we differentiate the individuals with respect to their sex, we need an extra assumption for the households having additional adults. (Recall that the data do not report the sex of children and additional adults).

H7: The probability for an additional member to be a female equals 1/2.

All these assumptions lead us to replace $G_p \, w_{p,a}$ and *(1-G_p) $w_{p,a}$* by ½ *π(a) nbsup* in the households with „additional members". Equation (8) becomes

(10)
$$E(y \mid x) = \beta^0 + \sum_{a=19}^{80} \left(n_a^M \beta_a^M + n_a^F \beta_a^F \right) + \sum_{e=1}^{4} n_e \beta_e^E + n_{81} \beta_{81} \, ,$$

where e stands for the age classes: (0-5 years), (6-9 years), (10-14 years), (15-18 years).

To the objective $\sum_{i=1}^{n} [y_i - E(y_i \mid x_i)]^2$ we add the penalty term

(11)
$$P = \lambda_M^2 \int (\frac{d^2 f^M(a)}{da^2})^2 \, da + \lambda_F^2 \int (\frac{d^2 f^F(a)}{da^2})^2 \, da.$$

The integration is made on age and the constants λ_G, $G = F, M$, control for the smoothing level, playing a role similar to the bandwidth parameter in kernel regression. In this section, we are only interested in the comparison between men and women, which is why we apply no penalty (smoothing parameter equals 0) to the children and to the old persons above 81. For instance, there is no reason to impose, via the penalty, that children's consumption should be related to the consumption of young men rather than to the consumption of young women, or vice versa. The discrete form of (11) uses the sum of squares of second differences of the β_a^S instead of the integrals:

(12)
$$\tilde{P} = \lambda_M^2 \beta^{M'} A' A \beta^M + \lambda_M^2 \beta^{F'} A' A \beta^F$$

where A is a 60x62 differencing matrix of the type:

$$
(13) \qquad A = \begin{vmatrix} 1 & -2 & 1 & 0 & . & . & . & . & 0 \\ 0 & 1 & -2 & 1 & 0 & . & . & . & . \\ . & 0 & 1 & -2 & 1 & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . & . \\ . & . & . & . & 0 & 1 & -2 & 1 & 0 \\ 0 & . & . & . & . & 0 & 1 & -2 & 1 \end{vmatrix}.
$$

Again following Chesher (1997), we do not differentiate the smoothing parameters for men and women. Thus we have to estimate a linear model (and simply use Ordinary Least Squares estimation) regressing the first column of D on the others:

$$
(14) \qquad D = \begin{bmatrix} y & \iota_n & N^M & N^F & N^E & N^{81+} \\ 0 & 0 & \lambda A & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda A & 0 & 0 \end{bmatrix} = \begin{bmatrix} y & X \\ 0 & \lambda W \end{bmatrix}
$$

where ι_n is a vector of ones; N^M and N^F are matrices with one line per household, reporting respectively the number of males and females of each age a between 19 and 80; N^E is a matrix of the same type but the members belong to the 4 age-classes between 0 and 18 years. N^{81+} is a vector containing one line for each household and indicating the number of household members above 80 years of age.

The *penalised OLS-estimator* is

$$
(15) \qquad \hat{\theta} = \left(X'X + \lambda^2 W'W \right)^{-1} X'y ,
$$

with variance

$$
(16) \qquad V(\hat{\theta}) = (X'X + \lambda^2 W'W)^{-1} X' \Sigma\, X\, (X'X + \lambda^2 W'W)^{-1} ,
$$

where $\Sigma = diag(V(\varepsilon / X))$, ε denoting the disturbance in the regression equation of y on X.

The variance estimations are made with the heteroskedasticity-consistent estimator $HC_2 = \dfrac{\hat{\varepsilon}_t^2}{1 - \hat{h}_t}$, with $\hat{h}_t$ the tth diagonal element of the projection matrix on the columns of D (dropping the first one). See for example Davidson and MacKinnon (1993, p552-556).

Figure 1 (in Appendix) presents the daily energy intake estimated with respect to the sex of people aged between 19 and 80 years of age, for 1991, for different values of the smoothing parameter, λ. The graphs also show (point-wise) confidence intervals at 95%, based upon asymptotic normality: $\hat{\beta}_a = \pm 1.96\,\hat{\sigma}_a$.

Remark: We include in the estimation the observations for children and senior people, but we do not smooth the estimation for these groups ($\lambda_E = \lambda_{81+} = 0$). We think their consumption can represent an important part of the household's consumption, and should be take into account.

For $\lambda = 5$, there clearly is excessive yearly variation for the estimated intake. Moreover the estimated energy intake for 19 year old men is not well determined (t=1.1). It appears desirable to increase the degree of smoothing. For λ near 30, the curves appear smooth enough and all the estimations are significant. So we choose to use a smoothing parameter of 30 for the following estimations, a decision based solely on the observation of the graphs; an alternative would have been cross-validation. The average levels found for energy intake (ranging between some 3620 and some 3880 kilocalories per day) are large in comparison with the results of Chesher for Great Britain.[11] But the OECD (1994) reports an average daily energy intake of 3233 kcal/d and of 3239 kcal/d for 1989 for Slovakia and for the Czech Republic respectively and of 3333 kcal/d and of 3301 kcal/d for 1990. Though these means are largely above the results for Great Britain and France they are of the same magnitude as the averages for Austria and Portugal . Nevertheless, the profiles for the evolution of energy intake with respect to age are similar to the profiles found by Chesher (1997). For men and women the energy intake increases with age until around 55. Only the young men profile presents a difference: between 20 and 30 years, the energy intake decreases in Chesher's results.

Figure 2 presents the estimated energy intake-age profiles for men and women on the same graph.[12] The second graph of the figure also shows confidence intervals at the 95% level. We can conclude that the distributions of the energy intake are not very different for men and women. This leads us to re-estimate the model without maintaining the distinction between the sexes and to smooth the consumption for the children and for the senior people, as described in the next sub-section.

When we study the decomposition of total energy intake between lipids, proteins and carbohydrates (beverages excluded), we note that, except for the lipids, the proportions differ from the recommended levels. The proportion of energy intake from carbohydrates is less than it is for a balanced diet (between 50 and 55% of total intake). This deficiency in carbohydrates is damaging to health. The energy intake from proteins (meats, fish, eggs, dairy products) is above the recommended 12% (this difference is more important for women). These results are presented in Figure 3 for 1991, where the diamond curves represent 12%, 30% and 50% of the total energy intake.

The decomposition of the carbohydrates' contribution to the energy intake between starch and "quick sugars" indicates an excess in the consumption of quick sugars and consequently a deficiency in starch (slow sugars). The difference between the contribution of (quick) sugars

---

[11] Those results exclude alcohol and confectionery, contrary to our results where only alcohol is excluded (for the estimations without household's characteristics). However this difference could not explain the different levels obtained.

[12] All the graphs for 1992 are available upon request (for these results and the ones discussed in the next sub-section). They are not presented here because they lead to similar conclusions.

and the advisable level is more important for women than for men. This gap is larger for women above 35 than for younger females. As regards men, the difference between ages is less pronounced and the gap between the intake and the recommended level increases significantly only after 45. All these results for 1991 are presented in Figure 4.

energy intake for men   ∘ energy intake for women

Kcal/d

6000

4000

2000

0

20   40   60   80

lambda = 30

Estimated energy intake in 1991 in Kcal/d

age

― energy intake for men   ∘ energy intake for women
── confidence interval for men   ─ ─ confidence interval for women

Kcal/d
6000

4000

2000

0

20   40   60   80

lambda = 30

Estimated energy intake in 1991 in Kcal/d

age
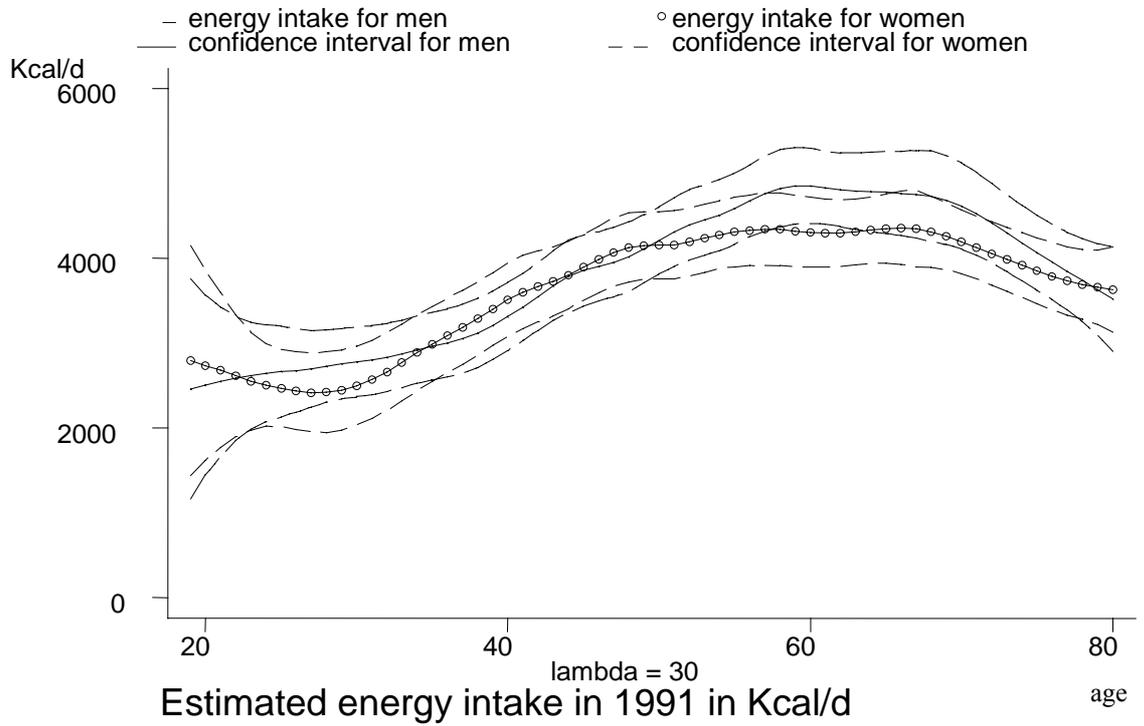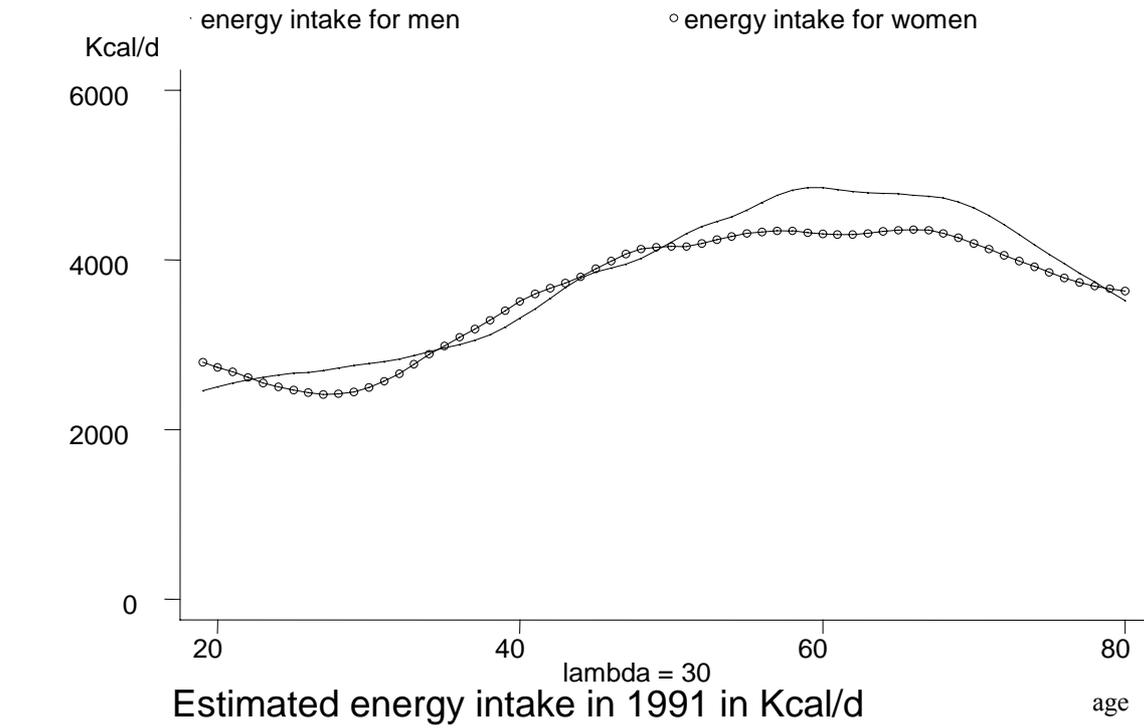
Figure 2: Estimated age profiles for total energy intake by sex for 1991, in kcal/day.

Energy intake for men in 1991 in Kcal/d

Energy intake for women in 1991 in Kcal/d
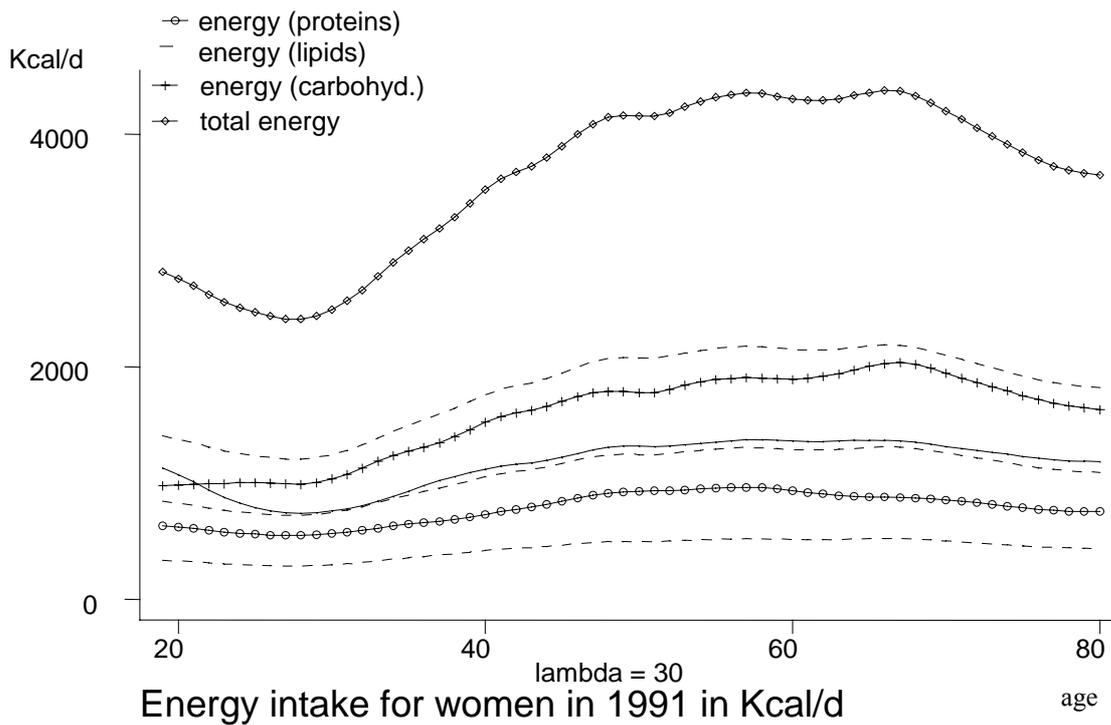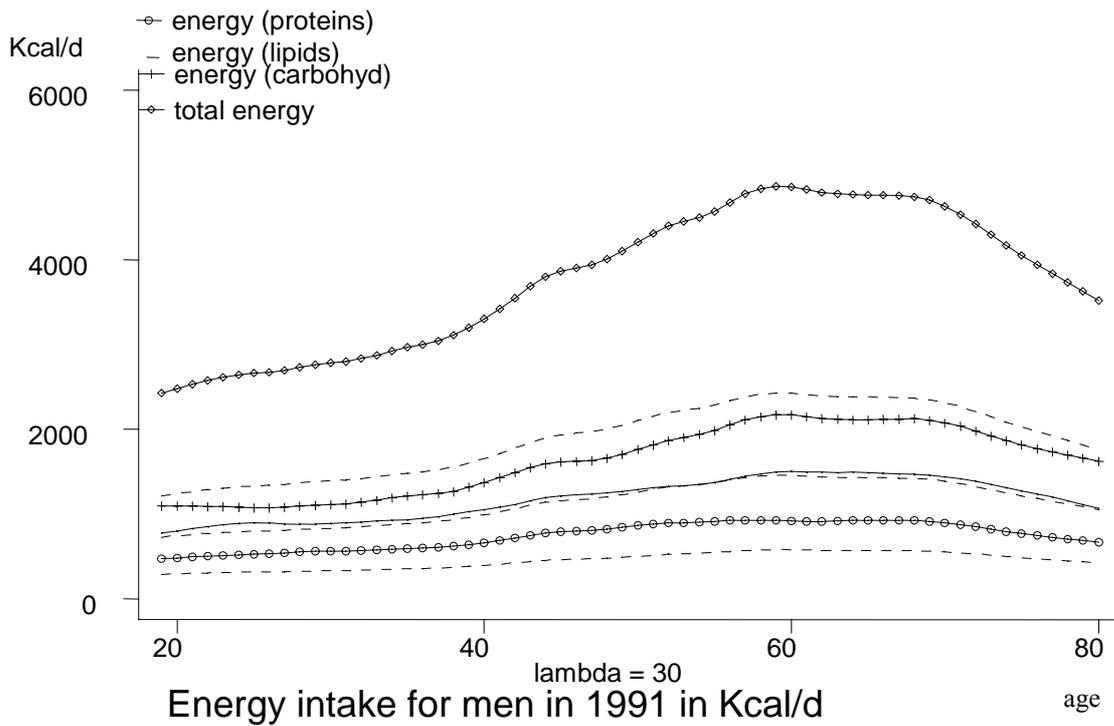
The dotted curves present 12%, 30% and 50% of total energy intake.

Figure 3: Estimated age profiles for total energy, proteins, lipids and carbohydrates intakes, in kcal/day, by sex.

Energy intake (carbohyd.) for men in 1991 in Kcal/d

Energy intake (carbohyd.) for women in 1991 in Kcal/d
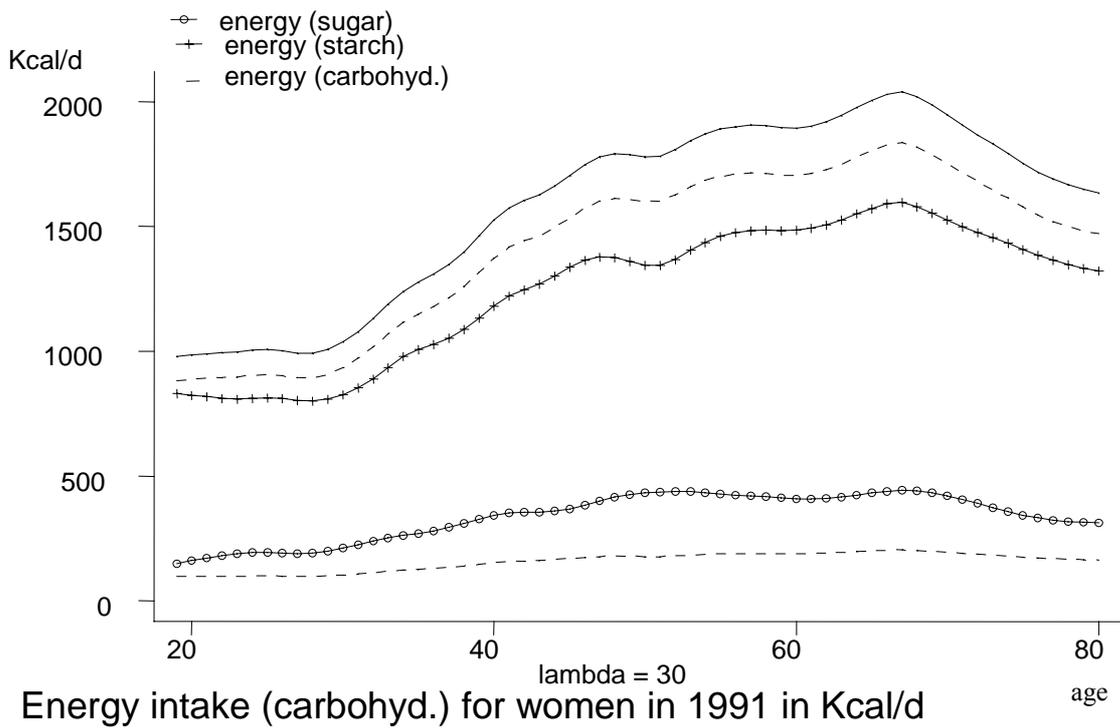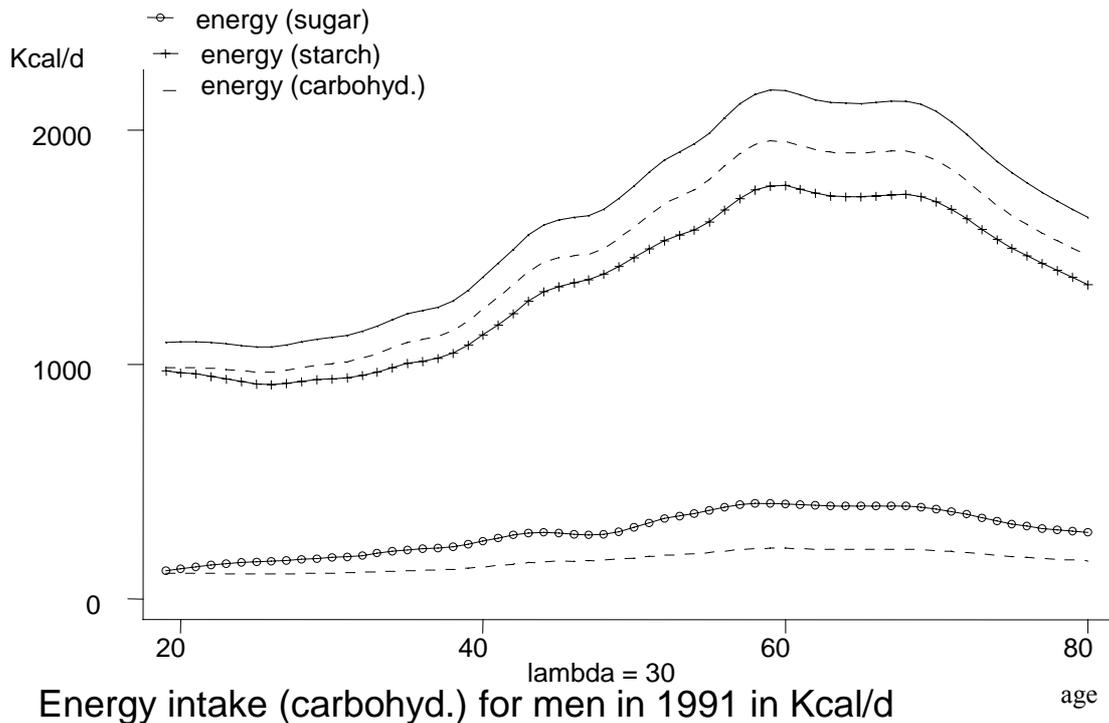
The dotted curves indicate 10% and 90% of total energy intake brought by carbohydrates.

Figure 4: Estimated age profiles, in kcal/day, 1991, decomposition of carbohydrates contribution to energy intake between quick sugars and starch, by sex.

4.1.B No separation by sex

As mentioned in the last section (comments to Figure 3), we re-estimate the energy intake – age profile without the distinction between men and women: we replace assumption H7 with

H8: We make no distinction between the individuals with respect to their sex.

Thus we modify equation (10) in the following way:

$$(17) \qquad E(y \mid x) = \beta^0 + \sum_{a=19}^{81} n_a \beta_a + \sum_{e=1}^{4} n_e \beta_e^E .$$

The penalty term now takes the simpler expression:

$$(18) \qquad P = \lambda^2 \beta' A' A \beta \qquad \text{where } A \text{ is now a 65x67 matrix.}$$

As we do not differentiate by sex, we can relate the energy intake of children to the energy intake of young adults. We use the same smoothing parameter for adults and children. This implies that the penalty terms of adults and children are contained in a unique matrix (the penalty matrix A is defined as before).[13] We do not interpret age and age – classes differently, i.e. we handle age – classes as age no matter the width of these classes.). Again the problem boils down to an OLS estimation with the matrix

$$(19) \qquad D = \begin{bmatrix} y & \iota_n & \tilde{N} \\ 0 & 0 & \lambda A \end{bmatrix}$$

where $\tilde{N} = N^E \mid N$, $\tilde{N}$ is a matrix with one row for each household and with columns indicating the number of household's members belonging to each age or age-class.[14]

Figure 5 (in the Appendix) presents the total estimated energy intake for years 1991 and 1992 with a smoothing parameter equal to 30.[15] The curves with non connected symbols represent 95% confidence intervals. The curve for energy intake is increasing for the young under 21, weakly decreasing until 30 years of age and after this age presents the same evolution as in Figure 2 (Section 4.1.A). The profile of energy intake is similar for the two years.

Figure 6a shows the decomposition into lipids, proteins and carbohydrates, and here the proportion of lipids in total energy intake almost exactly matches the recommended 30% of the total energy intake at all ages. However, the average level of lipids energy intake is very high, with approximately 920 kcal/day. This corresponds to the maximal lipids' energy intake for 60 years old in Great Britain (Chesher, 1997). The difference between the effective and advisable proportions of carbohydrates and proteins is still marked for adults above 20 years of age.

---

[13] Since the smoothing operates only on coefficients, the varying width of the age classes is of no concern.

[14] The sign | denotes concatenation of matrices.

[15] Estimates and standard errors are presented in the Appendix, in Table A5.
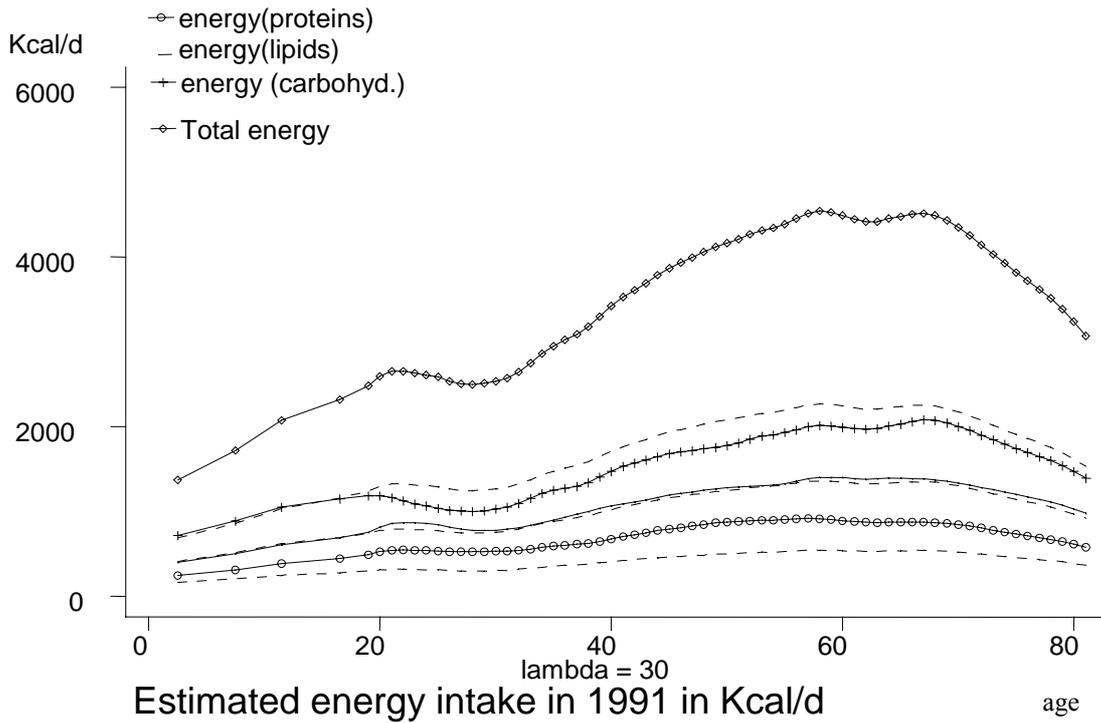
The dotted curves represent 12%, 30% et 50% of total energy intake.

Figure 6a: Estimated age profiles for total energy, lipids, proteins and carbohydrates intakes, in kcal/day, without distinction by sex.
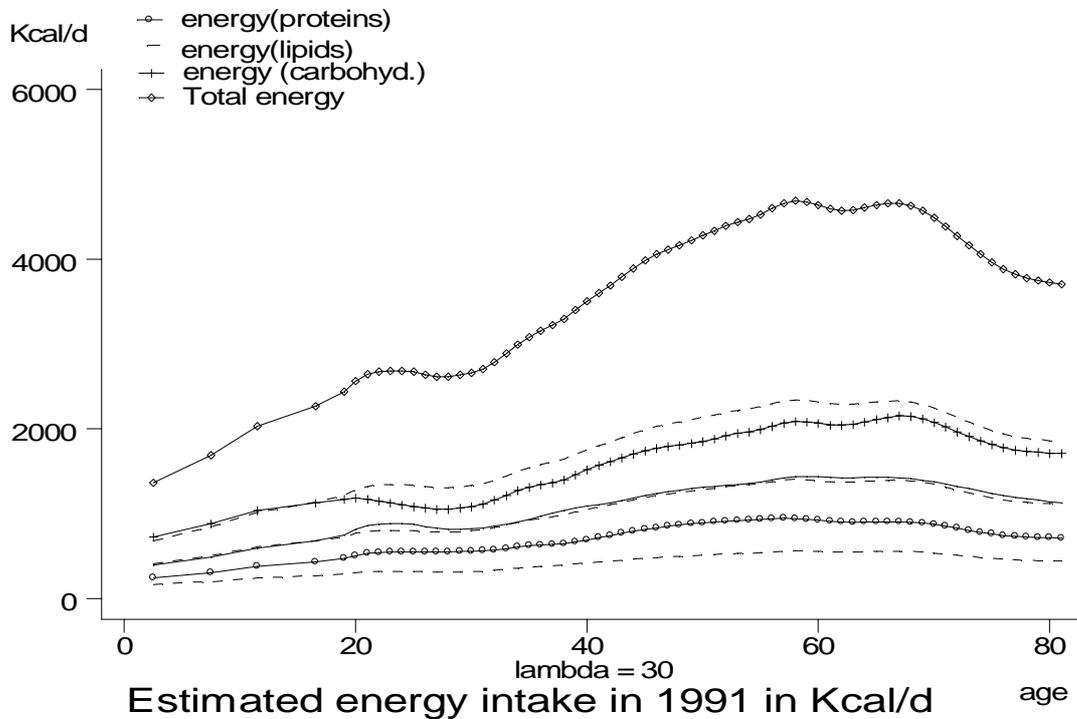

Figure 6b: Estimated age profiles for total energy, lipids, proteins and carbohydrates intakes, in kcal/day, without distinction by sex, for households without additional adult members.

In order to check the impact of the imputations made for the additional adults, we redo the estimation excluding the households with adults other than the household head and his wife. Figure 6b presents the equivalent of Figure 6a. We do not see noticeable changes in the estimated profile, except for old people. Their energy intake tends to become flat instead of decreasing. But this difference is not important enough to invalidate our assumptions.

As shows in Figure 7, the decomposition of carbohydrates energy intake between starches and sugars underlines the excess consumption of (quick) sugars. Interestingly, the children do not account for the most important abuses.

We also performed these estimations using 5-year age bands (as for children) instead of using each age separately. The results are similar to the previous results. The only difference is that the peak at 20 years old is less important, as was to be expected. These graphs are available upon request.
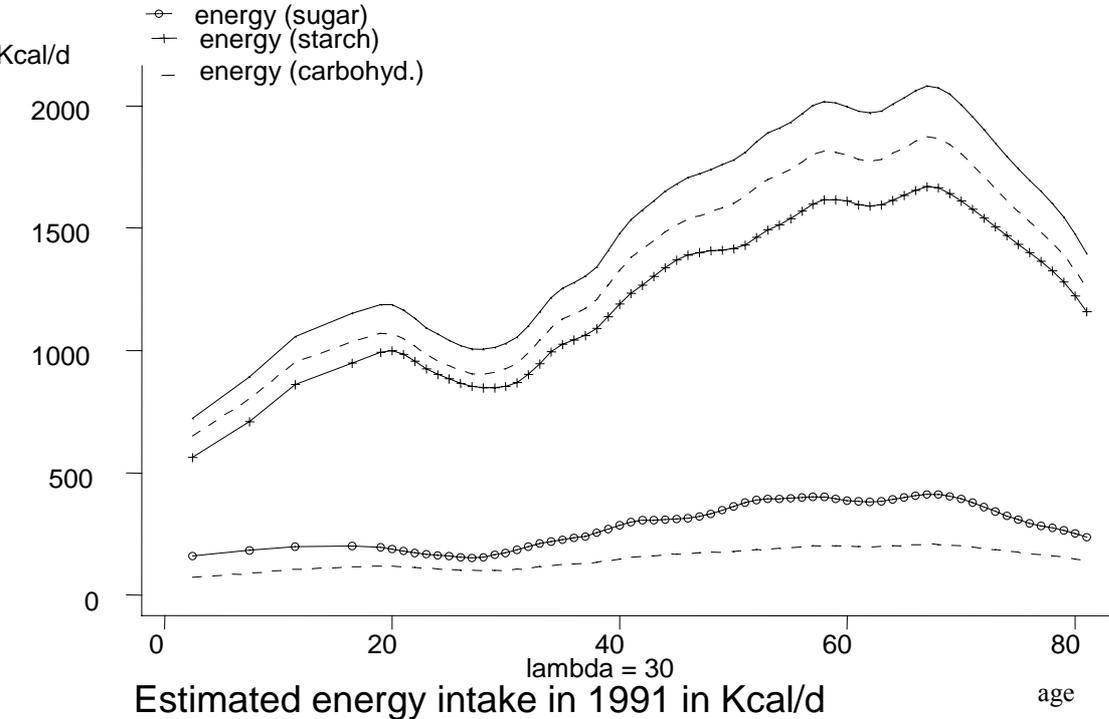
The dotted curves present 90% and 10% of total energy intake.

Figure 7: Estimated age profiles, in kcal/day, decomposition of carbohydrates' energy intake into sugars and starch, no distinction between men and women.

4.2 Estimation with smoothing across age and time

For the same reasons as in the last section we do not take into account the household characteristics, Z_t. We now estimate the relation nutrient intake–age for a period of 4 years (1989-1992). However the data from 1989 and 1990 do not record exactly the same information as the data for 1991 and 1992. For example, we could not retrieve the indication of the sex of the household head, even though this variable is mentioned in the description of the data files. The date of birth of the spouse is also missing. For 1990, the number of children

younger than 5 is equal to 0 for the whole sample, even though the difference between the number of children in the household and the children above 5 is non null for some families. So we consider age classes with a width of 5 years, except for the first class (width 4: 6-9) and the last class which corresponds again to people above 80 years of age. In addition, we use the assumptions H4, H5, H8 and

H2': We group the young people between 15 and 19 years of age in the same age class with no differentiation by sex. The same type of assumption is made for the senior people, aged of 80 years and more.

H3': If the additional adult in the household is young, she is 24 years old at most.

H6': $\pi(a)$ verifies:

- $P\{a \in [15, 19]\} = P\{a \in [20, 24]\} = 1/2$,

- $P\{a \in [65, 69]\} = P\{a \in [70, 74]\} = P\{a \in [75,79]\} = P\{a \in 80+\} = 1/4$.

H9: For a couple we assume that the age of the wife and the age of the husband belong to the same age class.
For example, denoting where age_H the husband's age and age_W the wife's age if age_H = 32, age_H $\in [30, 34 \text{ years old}]$ age_W $\in [30, 34 \text{ years old}]$.

When we smooth across age and time, we need three penalty terms: one to smooth across age, one to smooth across time and one to smooth across age and time together. Thus we add to the objective $(y_t - \beta_t^0 - \sum_{a=6}^{80} n_{at}\beta_{at})^2$ the following penalty term

(20)
$$P(r) = \alpha \int_{\Re^2} \left(r_{aa}^2 + 2r_{at}^2 + r_{tt}^2 \right) da\, dt,$$

where r_{aa}, r_{at}, r_{tt} are the second derivatives of r (r is defined in equation (9)) and α is a smoothing parameter.[16]

The discrete form of (20) leads to a penalty term of the following type[17]

(21)
$$P(r) = \lambda_{aa}^2 \sum_{a=10}^{79} \sum_{t=1}^{4} r_{aa}^2(a,t) + 2\lambda_{at}^2 \sum_{a=6}^{79} \sum_{t=2}^{4} r_{at}^2(a,t) + \lambda_{tt}^2 \sum_{a=6}^{80} \sum_{t=2}^{3} r_{tt}^2(a,t),$$

where the $r_{..}$ denote the second differences

(22) $r_{aa}(a,t) = \beta_{a+1,t} - 2\beta_{a,t} + \beta_{a-1,t}$
(23) $r_{at}(a,t) = \beta_{a+1,t} - \beta_{a,t} - \beta_{a+1,t-1} + \beta_{a,t-1}$
(24) $r_{tt}(a,t) = \beta_{a,t+1} - 2\beta_{a,t} + \beta_{a,t-1}$.

---

[16] a represents age classes: 6-9 years old, 10-14 years old...75-79 years old and 80-and more.
[17] The reason why the summation limits take these values can be read off equations (22-24).

After discretisation, the problem becomes the estimation by OLS of

(25) $$y = P_h M_h + \lambda_{aa}(P_{aa} M_{aa}) + 2\lambda_{at}(P_{at} M_{at}) + \lambda_{tt}(P_{tt} M_{tt}),$$

where the term $P_h M_h$ takes into account the observations for the households, and the three others terms the penalties.[18] $P_h, P_{aa}, P_{at}, P_{tt}$ are vectors of dummy variables which ensure that each penalty term is applied only to the suitable data. For the first n elements, P_h contains ones and the other $P_{..}$ contain zeros.[19]

(26) $$M_h = \sum_{t=1}^{4} d_t \left[\beta_t^0 + \sum_{a=6}^{80} n_{at} \beta_{at} \right],$$

where d_t is a dummy variable equal to 1 if the household was questioned during year t, 0 otherwise; n_{at} is the number of household's members belonging to age class a observed during year t. The 56 additional observations associated with the term $r_{aa}(a,t)$ are null for y, P_h, P_{at}, P_{tt} and equal to 1 for P_{aa}.

(27) $$M_{aa} = \sum_{t=1}^{4} d_t \sum_{a=9}^{79} (n_{at} \beta_{at})$$

where n_{at} and d_t belong to [N D],

(28) $$[N\,D] = \begin{bmatrix} A & Q_1 \\ \cdot & \cdot \\ A & Q_4 \end{bmatrix},$$

where A is the 14x16 matrix of second differences given in equation (13) and Q_t is a 14x4 matrix of zeroes except in column t which contains ones. For the elements corresponding to the second penalty term (90 observations) y, P_h, P_{aa}, P_{tt} are null vectors and P_{at} is a vector of ones.

(29) $$M_{at} = \sum_{t=2}^{4} d_t \sum_{a=6}^{79} (n_{at} \beta_{at})$$

with n_{at} and d_t belonging to

(30) $$[N\,D] = \begin{bmatrix} A^1 & Q_2 \\ A^1 & Q_3 \\ A^1 & Q_4 \\ -A^1 & Q_1 \\ -A^1 & Q_2 \\ -A^1 & Q_3 \end{bmatrix},$$

---

[18] The vector products are to be read element by element.
[19] Here n denotes the total number of households for the 4 years: n=17001.

where A^1 is the 15x16 matrix of first differences $\begin{bmatrix} 1 & -1 & 0 & \cdot \\ 0 & 1 & -1 & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$, and Q_t is a 15x4 matrix of

zeroes except in the column t which contains ones. Finally, the last 32 observations for the smoothing across time are null for y, P_h, P_{aa}, P_{at} and equal 1 for P_{tt}.

$$(31) \qquad M_{tt} = \sum_{t=2}^{3} d_t \sum_{a=6}^{80} (n_{at}\beta_{at})$$

with n_{at} and d_t belonging to

$$(32) \qquad [\text{N D}] = \begin{bmatrix} Q_1 & A \\ \cdot & \cdot \\ \cdot & \cdot \\ Q_{16} & A \end{bmatrix},$$

with A now equal to $\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix}$ and Q_a a 2x16 matrix with ones in the column corresponding to the age class a and zeroes elsewhere.
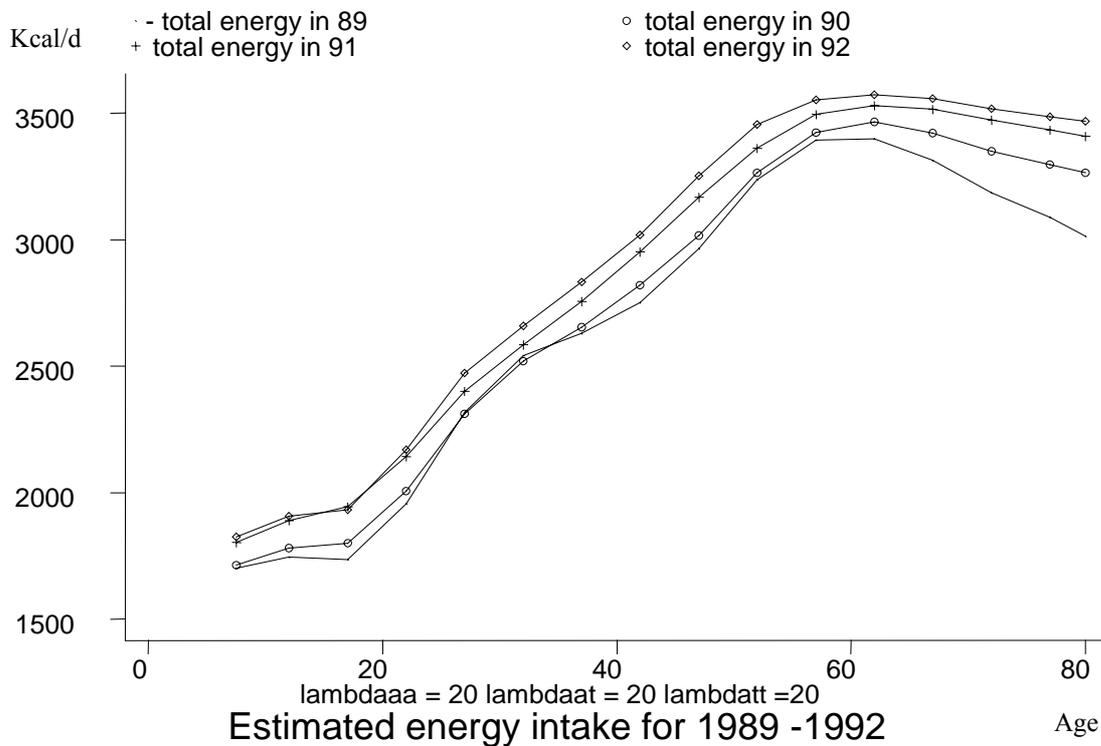

Figure 8: Estimated age profiles for total energy intake, smoothing across age and time, in Kcal/day.

Figure 8 presents the total estimated energy intake for smoothing parameters $\lambda_{aa} = \lambda_{at} = \lambda_{tt} = 20$. These curves show a strong increase of the energy intake with age after 20. The increase of this intake with time at all ages (an exception being ages around 30 for 1989 and 1990) is striking, the largest difference being between 1990 and 1991. Indeed, Fisher tests of equality of parameters β (energy intake) reject for all pairs of years (see Table 2).

Table 2: Fisher tests of equality of parameters β for all pairs of years: $F(16,17114)=1.644$

$F_{89\text{-}90}$ = 7.64	$F_{89\text{-}91}$ = 37.28	$F_{89\text{-}92}$ = 45.67
$F_{90\text{-}91}$ = 21.46	$F_{90\text{-}92}$ = 24.89	$F_{91\text{-}92}$ = 3.75

Figures 9 to 11 present the results for the energy intake decomposition between carbo-hydrates, proteins and lipids. The intake of carbohydrates follows the evolution of the total energy intake, that is to say it increases over time. The protein intake follows an opposite evolution, decreasing with time. Fisher tests reject the assumption of equality for all pairs of years. The lipids contribution increases with time. The assumption of equality is rejected for all pairs of years when all the age classes are included. However, we cannot reject it for 1989-1990 for people below 70 (*P-value* = 0,37 ; *F*(16,17114) = 1.08) and for 1991-1992 for people below 34 (*P-value* = 0.42 ; *F*(16,17114) = 1.01).

The average energy intake varies between 3699 (3868) kcal/d in 1991 (1992) when estimated under assumptions H1-H7 (male/female separation, smoothing across age only), and 2866 (2917) kcal/d in 1991 (1992) when estimated under assumptions H2', H3', H4, H5, H6', H8 and H9 (no male/female distinction, and smoothing across age and time). It is especially sensitive to the treatment of the age of adults, i.e. to whether we use age classes or age in years. When only the assumption on the smoothing differs, the average estimated energy intake is 2964 (2947) kcal/d in 1991 (1992) when smoothing occurs across time only, and 2866 (2917) kcal/d in 1991 (1992) when smoothing across age and time.

The contribution to energy intake as between carbohydrates, proteins and lipids intakes for one year evolves over time. The proteins contribution decreases in favour of both carbohydrates and lipids, with a more important increase for the carbohydrates. Even with this evolution, the lipids' share in total energy intake is always close to the recommended figure of 30% of total energy intake for a balanced diet. The proteins' share, decreasing with time, goes some way towards the recommended figure of 12%, but remains above it.[20] In summary, significant progress appears to have been made in these four years towards a more balanced diet.

---

[20] The graphs are presented in the Appendix of a fuller version available on the web www.siaw.unisg.ch/lechner/miquel, Figures A1-A4.
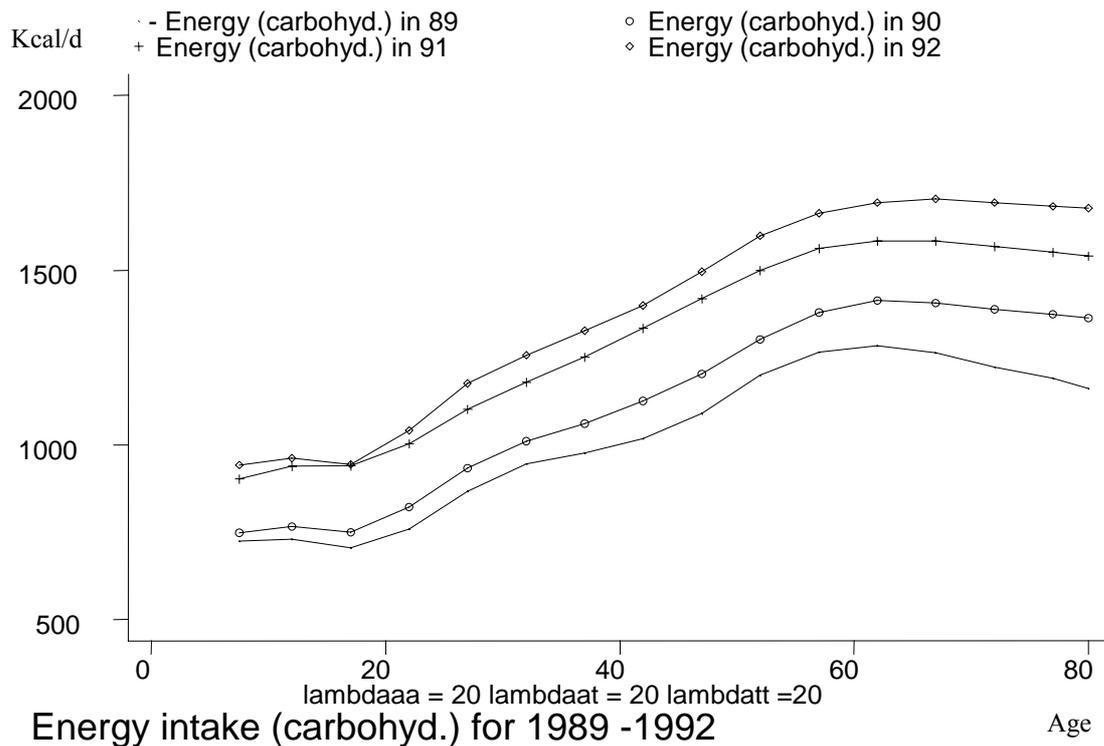
Figure 9: Energy intake: contribution of carbohydrates, smoothing across age and time.
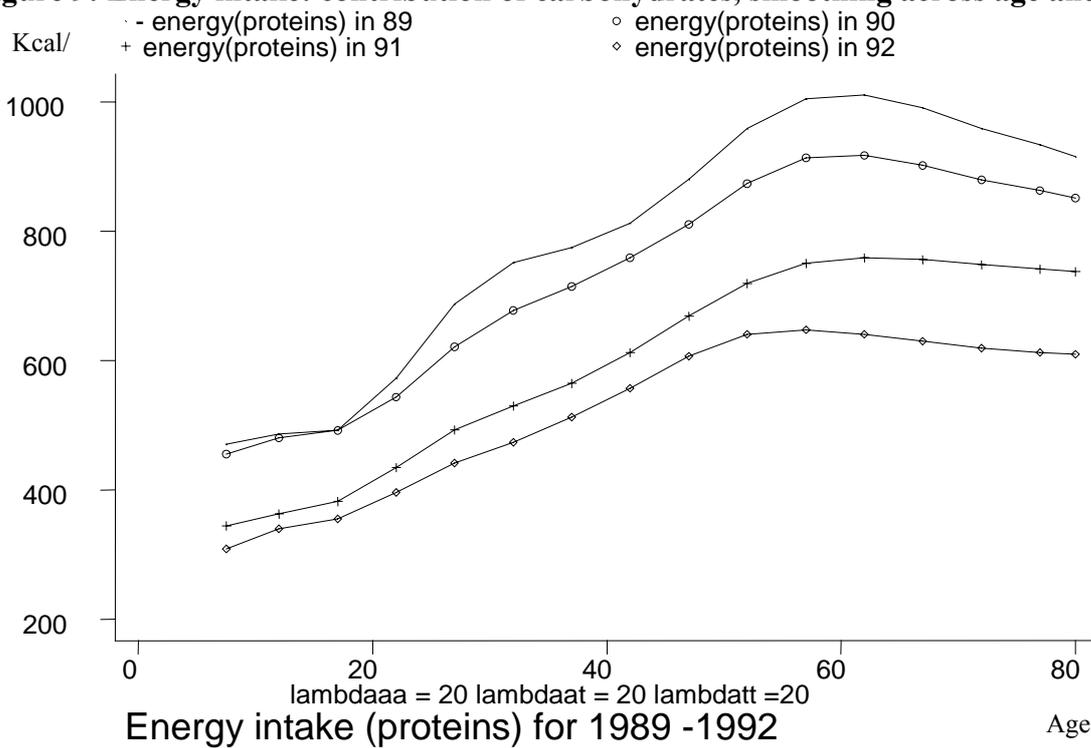

Figure 10: Energy intake: contribution of proteins, smoothing across age and time.
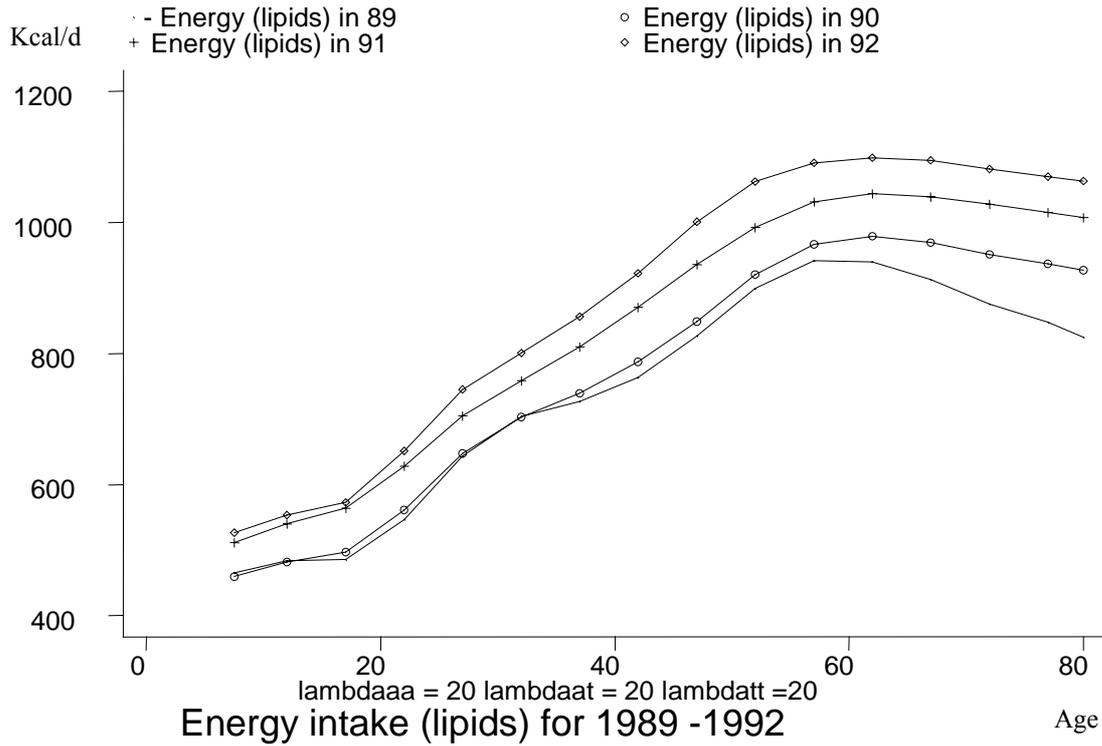
Figure 11: Energy intake: contribution of lipids, smoothing across age and time.

4.3 Estimation with household characteristics

Finally, we take account of some household characteristics in the estimation procedure. More precisely, we want to study the influence of household income and of the wife's education level on individual energy intakes. The model we want to estimate,

$$E(y \mid x, Z) = \sum_{p=1}^{n} f(x_p) h(Z), \ E(y \ / x, Z) = \sum_{p=1}^{N} f(x_p) h(Z),$$

is now a non linear regression model. We specify the function $h(.)$ of household characteristics in the parametric form $h(Z) = \exp(Z' \delta)$. The non-linear least squares estimator used minimises

$$(33) \qquad R(\theta) = \sum_{i=1}^{n+n^*} (y_i - \exp(Z_i' \delta) x_i' \beta)^2.$$

The data corresponding to the total available energy intake in the household, y_i, and those corresponding to the household's characteristics, Z_i, are expanded with n^* rows of zeroes, one for each second difference appearing in the penalty term. The data corresponding to the number of persons by age, x_i, are expanded by the penalty terms λW. The estimator, $\hat{\theta}$, converges to the solution of the first order conditions of the limit problem

$$(34) \qquad E(n^{-1} \frac{\partial R(\overline{\theta})}{\partial \theta}) = p \lim_{n \to \infty} (n^{-1} \frac{\partial R(\overline{\theta})}{\partial \theta}) = 0,$$

25

the expectation being with respect to the true value, θ_0 of θ. Using a central limit theorem, $\sqrt{n}(\hat{\theta} - \bar{\theta})$ can be shown to be asymptotically normal. For more details see Chesher (1997).

The estimation is based on the assumptions H1, H2', H3', H4, H5, H6' and H8 (no distinction based on the sex of the individuals, age classes for adults). We regroup the different levels of education to obtain only two levels: a "lower" level of education (primary school, apprenticeship and high school) and a "higher" one. Thus the function $h(Z)$ is specified as $\exp[\delta_1 e_1 + \delta_2 \ln(y/Y)]$, where e_1 is a dummy variable taking the value 1 for women with the lower level of education, and of 0 for the others, and $\ln(y/Y)$ is the logarithm of the income normalised by its mean.[21] For the estimation, we exclude the households which are not based on a couple, and thus delete 137 observations for 1991 and 127 for 1992. The results for the estimated total energy intake and its decomposition in carbohydrates, lipids and proteins are presented in Figure 12, where the curves are drawn for the average income, i.e. for $\ln(y/Y) = 0$.

In the figure, the four groups of curves represent the energy intake in proteins, lipids and carbohydrates and the total energy intake respectively. The non - connected symbols represent the results when taking account of household characteristics, the lines represent the same results when ignoring these household characteristics. No marked differences emerge at mean income.
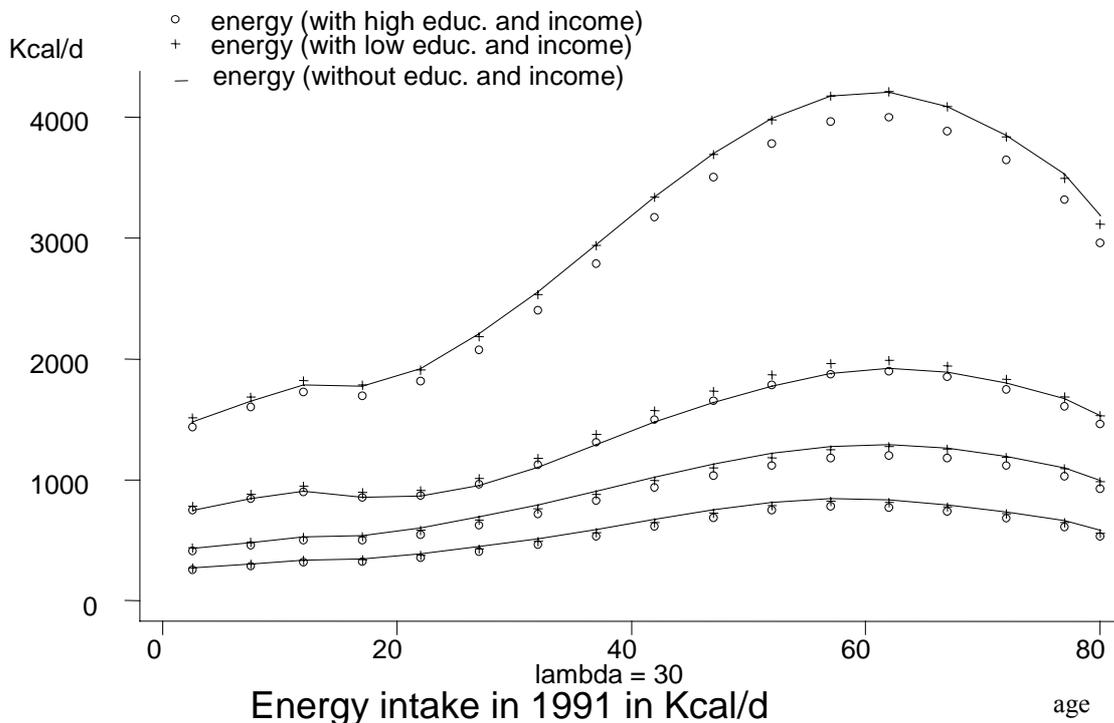
**Figure 12: Estimated age profiles for energy intake, with decomposition, taking account of income and education.**

---

[21] The high level of education does not appear in the objective because a constant term in $h(Z)$ would not be identified.

The estimates for the parameters of the wife's education and income are presented in Table 3a. The education level parameter, δ_1, though significantly positive for all years and all forms of energy intake, has a very small magnitude and thus only a marginal impact on energy intake. The income parameter, δ_2, is significantly positive for lipids and proteins in 1991 and significantly negative for carbohydrates in 1992. A Student test rejects the assumption of equality of the parameters between the two years only in the case of proteins ($\hat{t} = 2.67$). The influence of income on energy intake of these three nutrients is very weak, a conclusion that was already reached by Chesher for Great Britain.[22] For the intake of beverages (alcoholic and non–alcoholic beverages taken at home), the income coefficient is significantly positive, and much larger than for carbohydrates, lipids, and proteins.

Table 3a: Estimations for δ_1 and δ_2, the parameters of education level and income *(t values)*.

	Carbohydrates		Lipids		Proteins		Beverages	
	δ_1	δ_2	δ_1	δ_2	δ_1	δ_2	δ_1	δ_2
1991	**0.046**	-0.037	**0.058**	**0.057**	**0.052**	**0.056**	**0.073**	**0.361**
	(3. 6)	*(-1.5)*	*(4.7)*	*(2.4)*	*(4.4)*	*(2.4)*	*(3.4)*	*(8.9)*
1992	**0.028**	**-0.061**	**0.049**	0.021	**0.045**	-0.027	0.039	**0.350**
	(2.1)	*(-2.7)*	*(4.3)*	*(1.1)*	*(3.6)*	*(-1. 3)*	*(1.7)*	*(9.8)*

Table 3b: Estimations for δ_1 and δ_2, the parameters of education level and income *(t values)*, for households without additional adult members.

	Carbohydrates		Lipids		Proteins		Beverages	
	δ_1	δ_2	δ_1	δ_2	δ_1	δ_2	δ_1	δ_2
1991	**0.046**	-0.040	**0.054**	0.030	**0.048**	0.031	**0.066**	**0.261**
	(3. 4)	*(-1.4)*	*(3.9)*	*(1.1)*	*(3.8)*	*(1.2)*	*(2.9)*	*(5.6)*

Table 3b presents the results of the sensitivity analysis. The estimated coefficients are smaller in absolute value and less precise.[23] The conclusions concerning the education level parameter still stay the same. Except for beverages, the income parameter δ_2 is not significant.

The age profile of energy intake seems to follow the general variations of income over the life cycle. Until the age of 20 approximately, people have access to important financial resources because they depend on their parents and the energy intake corresponding to this period is increasing with age. Between 20 and 30 years of age, people leave their parents' house and so the young people's financial resources are limited; the energy intake is decreasing for this period. Then during the working life, their income rises again. We see the same evolution for the energy ingested. The last step corresponds to pension age, and the energy intake curve declines progressively.

---

[22] As pointed out by a referee, the presence of some negative coefficients (of which only one is significant) could possibly be related to strenuous work: if manual workers have lower incomes on average, we might expect a tendency towards seeing lower incomes associated with higher energy intakes.

[23] The loss in precision is not due only to the fairly small reduction of the sample size, but also to the higher homogeneity of the sub-sample without additional adult members.

Conclusions

The aim of this paper was to estimate the relationship between energy intake and age for Czechoslovak individuals and to study its evolution over the period 1989-1992 on the basis of the household purchases reported in the Budget Surveys. For this purpose we adapted the non-parametric approach of Chesher (1997, 1998) to the particularities of the available data, and obtained the following substantive results.

The age profiles of the energy intake of Czechoslovakians are similar for males and females. The overall energy intake increases with age until 55, is approximately constant between 55 and 70, and decreases afterwards. Chesher reports similar profiles for Great Britain, but the overall level of energy intake is much higher in Czechoslovakia than in Great Britain.

The contribution of carbohydrates to the energy intake is below the 50% recommended for a balanced diet, whereas the contribution of proteins is above the advisable 12% level. However, these imbalances have markedly decreased over the four-year period studied. The contribution of lipids to the total energy intake matches the advised level of 30%.

In households based on a couple, the education level of the woman plays almost no role in the quantity of calories ingested: although statistically significant, the effects are very small. However, a higher education level of the wife leads to a slight moderation of energy intake.

The impact of income on the nutrient intakes is not consistently significant over time for proteins, lipids and carbohydrates. However, it is positive for beverages (consumption outside the home is excluded).

Overall, the age profile of energy intake seems to follow the general variations of income over the life cycle (increasing during the working life, with income, decreasing for pensioners).

Finally, we should note that the period studied was a phase of important transformation in the Czechoslovak economy. In 1991, the Czechoslovak Republic abandoned planning and freed prices. It would be interesting to study the decomposition of the Czech and Slovak diets after this period, in order to check whether the diffusion of Western-European patterns of food consumption related to food imports has had an impact on diet quality. The role played by the changes in relative prices of food products on the nutriments' intakes should also be the subject of further research. The results presented in this paper provide an empirical starting point for such research.

Appendix
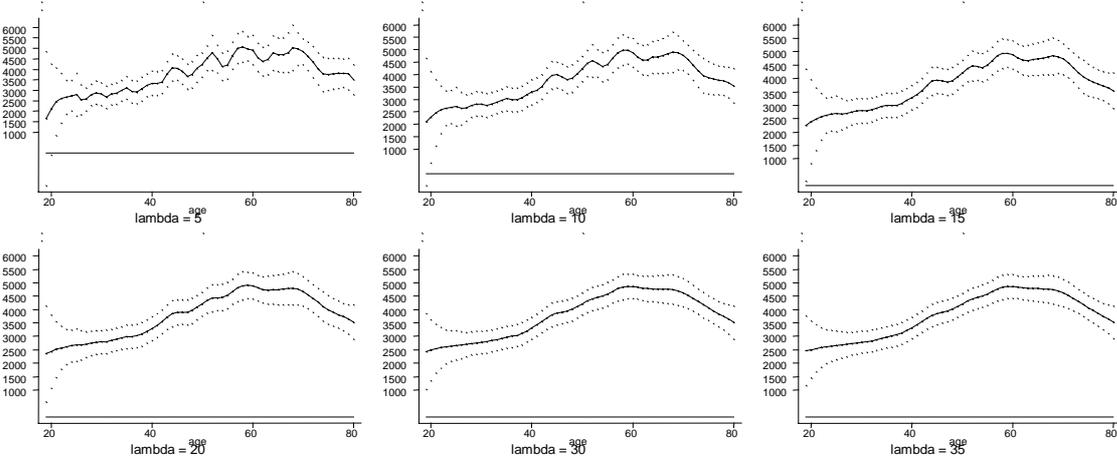
Total energy intake for men in 1991 in kcal/d
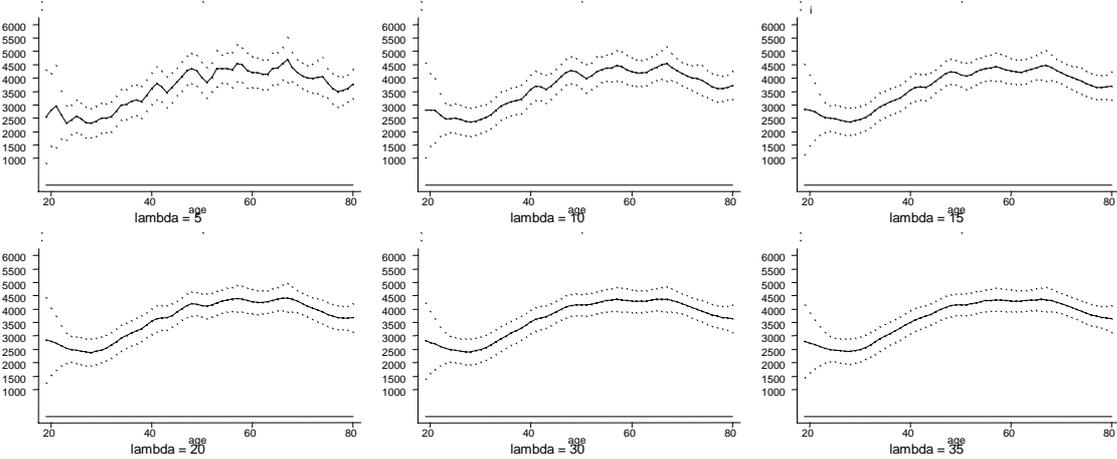
Total energy intake for women in 1991 in kcal/d

Figure 1: Estimated age profiles for total energy intake by sex, in kcal/day, 1991, different values of the smoothing parameter λ.

Table A1: distribution of households by type

Household type	1989	1990	1991	1992
1	953	1010	801	673
2	50	42	29	26
3	3045	3075	2025	1641
4	272	242	132	153
5	249	256	154	144
6	83	77	64	52
7	30	34	19	13
8	167	180	150	133
9	830	860	574	501
Total	5679	5776	3948	3336

1 = complete family with housewife: two-member family without children, family with dependent children (no child is economically active) without other members (grand-mother, uncle, etc.) .
2 = complete family with housewife and additional member: family without children with another person (mother in law, grand-father, etc.), family with children, some of them economically active, family with children and one (or more) additional member(s).
3 = complete family with employed wife: same as category 1.
4 = complete family with employed wife and additional members: same as category 2.
5 = single parent family, the household head being this parent.
6 = single parent family, the household head being this parent, and additional members: same as category 2 (part with children).
7 = non family household.
8 = single man.
9 = single woman.

Table A2: proportion of households by number of children aged between 0 and 14.

	1989		1990		1991		1992	
# Children	# hh	%	# hh	%	# hh	%	# hh	%
0	2483	43.72	2554	44.22	1838	46.56	1613	48.35
0 to 2	1181	20.80	1174	20.33	1834	46.45	1488	44.60
2 to 3	1483	26.11	1495	25.88	250	6.33	220	6.59
3 to 4	467	8.23	494	8.55	25	0.63	14	0.42
4 and more	65	1.14	59	1.02	1	0.03	1	0.03

Table A3: proportion of households by number of members.

	1989		1990		1991		1992	
Members	# hh	%	# hh	%	# hh	%	# hh	%
0 to 2	1006	17.71	1049	18.16	1812	45.90	1605	48.11
2 to 3	1527	26.89	1561	27.03	736	18.64	574	17.21
3 to 4	1009	17.77	1015	17.57	1108	28.06	906	27.16
4 to 5	1560	27.47	1563	27.06	258	6.53	226	6.77
5 to 6	504	8.87	523	9.05	31	0.79	21	0.63
6 and more	73	1.29	65	1.13	3	0.08	4	0.12

Table A4: Mean income by number of household members, nominal, in current koruna.

# members	1989		1990		1991		1992	
	mean	s.d.	Mean	s.d.	mean	s.d.	mean	s.d.
1	2317	1039	2464	1006	37335	17348	42957	22720
2	5170	1937	5480	2013	81818	32026	93353	42536
3	6650	2188	6971	3181	96995	35670	119343	49770
4	7046	1782	7569	2158	108571	33574	129919	49119
5	7722	2313	8404	2526	113979	34633	127904	45114
6	7945	1795	9098	2168	128786	48566	147298	59463
7	8786	1730	8941	1370	139806	33547	208169	53001
Total	5749	2581	6119	2962	86625	40685	101056	53685
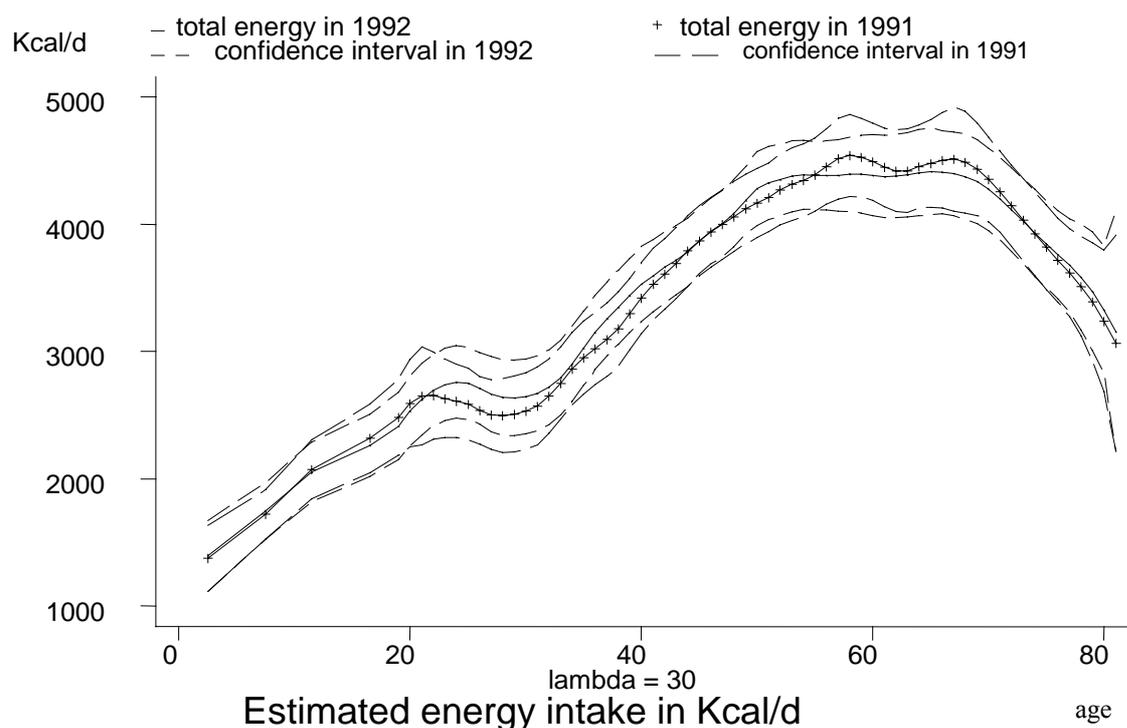
The confidence intervals are 95% confidence intervals.

Figure 5: Estimated age profiles for total energy intake, in kcal/day, without distinction by sex.

Table A5: Estimation results, no separation male/female, 1991 (see Figure 5)

variable	Coef.	HC2 s.e.	t	variable	Coef.	HC2 s.e.	t
Regression with robust standard errors					Number of obs =		3533
R-squared = 0.583				Dependent variable: household energy intake			
age5	1577	124	*12.7*	**age49**	4768	138	*34.5*
age9	1967	95	*20.7*	**age50**	4835	136	*35.6*
age14	2367	114	*20.7*	**age51**	4894	136	*36.1*
age18	2651	133	*19.9*	**age52**	4954	139	*35.7*
age19	2845	150	*19.0*	**age53**	4990	140	*35.7*
age20	2968	168	*17.7*	**age54**	5014	145	*34.7*
age21	3029	181	*16.8*	**age55**	5062	144	*35.1*
age22	3026	158	*19.1*	**age56**	5120	150	*34.2*
age23	2993	141	*21.3*	**age57**	5170	159	*32.5*
age24	2965	131	*22.7*	**age58**	5176	157	*32.9*
age25	2934	127	*23.1*	**age59**	5142	150	*34.2*
age26	2887	130	*22.2*	**age60**	5089	150	*33.9*
age27	2861	136	*21.1*	**age61**	5028	159	*31.7*
age28	2861	139	*20.6*	**age62**	4987	170	*29.3*
age29	2873	142	*20.2*	**age63**	4975	175	*28.4*
age30	2887	145	*19.9*	**age64**	5001	163	*30.7*
age31	2914	156	*18.7*	**age65**	5018	166	*30.3*
age32	2995	144	*20.7*	**age66**	5044	169	*29.8*
age33	3108	137	*22.6*	**age67**	5061	179	*28.3*
age34	3234	138	*23.5*	**age68**	5043	176	*28.6*
age35	3332	141	*23.7*	**age69**	4990	166	*30.1*
age36	3397	139	*24.4*	**age70**	4907	158	*31.0*
age37	3465	142	*24.5*	**age71**	4796	153	*31.3*
age38	3557	143	*25.0*	**age72**	4667	155	*30.2*
age39	3684	135	*27.2*	**age73**	4537	161	*28.3*
age40	3833	132	*29.1*	**age74**	4408	166	*26.6*
age41	3976	135	*29.5*	**age75**	4285	168	*25.5*
age42	4092	134	*30.6*	**age76**	4168	169	*24.6*
age43	4210	135	*31.2*	**age77**	4051	180	*22.5*
age44	4329	137	*31.7*	**age78**	3926	210	*18.7*
age45	4422	134	*33.1*	**age79**	3776	251	*15.1*
age46	4504	134	*33.6*	**age80**	3587	294	*12.2*
age47	4589	136	*33.8*	**age81**	3362	483	*7.0*
age48	4681	137	*34.2*	**Const.**	198	142	*1.4*

References

Behrman J., and Deolalikar, A. (1987): "Will Developing Country Nutrition Improve with Income? A Case Study for Rural South India", *Journal of Political Economy*, 95, 492-507.

Behrman, J., and Wolfe, B. (1984): "More Evidence on Nutrition Demand: Income Seems Overrated and Women's Schooling Under-emphasized", *Journal of Development Economics*, 14, 105-128.

Bhalotra, S., and Attfield, C. (1998): "Intra-household Resource Allocation in Rural Pakistan : a Semi-parametric Analysis", University of Bristol Department of Economics Discussion Paper 98 /447.

Bhargava, A. (1997): "Editor's introduction: Analysis of data on health", *Journal of Econometrics*, 77 (Annals of Econometrics, 1-295), 1-4.

Bouis, H. (1994): "The Effect of Income on Demand for Food in Poor Countries : Are Food Consumption Databases Giving Us Reliable Estimates?", *Journal of Development Economics*, 44, 199-226.

Bouis, H., and Haddad, L. (1992): "Are Estimates of Calorie-Income Elasticities Too High?", *Journal of Development Economics*, 39, 333-364.

Chesher, A. (1997): "Diet Revealed? Semi-parametric Estimation of Nutrient Intake-Age Relationships", *Journal of the Royal Statistical Society*, 160(3), p389-420.

Chesher, A. (1998): "Individual Demands from Household Aggregates: Time and Age Variation in the Composition of Diet", *Journal of Applied Econometrics, Special Issue: Application of Semiparametric Methods for Micro-Data*, Guest Editors: J. Horowitz, M.-J. Lee, B. Melenberg and A. van Soest, 13, 505-524.

Dasgupta, P. (1997): "Nutritional Status, The Capacity for Work, and Poverty Traps", *Journal of Econometrics*, 77, 5-37.

Davidson, R., and MacKinnon, J. (1993): *Estimation Inference in Econometrics*, Oxford University Press.

Department of Health (1989): "The Diets of British Schoolchildren", *Report on Health and Social Subjects* n°36, London: HMSO.

Dupin, H., Abraham, J., and Giachetti, I. (1992): *Apports Nutritionnels Conseillés pour la population française*, Tec.Doc. Lavoisier.

Engel, E. (1857): "Die Produktions- und Konsumptionsverhältnisse des Königreichs Sachsen", *Zeitschrift des Statistischen Bureaus des Königlichen Sächsischen Ministeriums des Innern,* Nr. 8-9.

Engle, R., Granger C. , Rice J., and Weiss A. (1986): Semi-parametric Estimates of the Relation between Weather and Electricity Sales, *Journal of the American Statistical Association,* 81, 310-320.

Institut Scientifique d'Hygiène Alimentaire (1990): *Tables de Composition des Aliments*, Editions Jacques Lanore.

Mills, A., and Tyler, H. (1992): *Food and Nutrient Intakes of British Infants Aged 6-12 Months*, London :HMSO.

Noël, A. (1988): *Tableau des Calories*, Editions S.A.E.P.

OCDE (1994): *Les Républiques Tchèque et Slovaque*. Études Économiques de l'OCDE. Paris.

Pitt, M. (1983): "Food Preferences and Nutrition in Rural Bangladesh", *The Review of Economics and Statistics*, 65, 105-114.

Silverman, B.W. (1986): *Density Estimation for Statistics and Data Analysis*, Monographs on Statistics and Applied Probability 26, Chapman and Hall, London.

Strauss, J. (1982): "Joint Determination of Food Consumption and Production in Sierra Leone: Estimates of a Household-Firm Model", *mimeo*, University of Virginia, Charlottesville, VA.

Subramanian, S., and Deaton, A. (1996): "The Demand for Food Calories", *Journal of Political Economy*, 104, 133-162.

World Bank (1981): *World Development Report*, Washington, World Bank.