

Household food consumption, individual caloric intake and obesity in France

Céline Bonnet, Pierre Dubois & Valérie Orozco

Empirical Economics
Journal of the Institute for Advanced
Studies, Vienna, Austria

ISSN 0377-7332
Volume 46
Number 3

Empir Econ (2014) 46:1143-1166
DOI 10.1007/s00181-013-0698-1

Volume 46 · Number 3 · May 2014

Empirical Economics

A Journal of the Institute for Advanced Studies, Vienna, Austria

EDITORS
Hedder M. Anderson
Badi H. Paltal
Robert M. Kling
Subal C. Kumbhakar
Michael Lochner
EDITORIAL BOARD
G. Arbia
M. Blewett
E. Bunn
B. Candelon
D. Checchi
G. Doppelhofer
J. P. Elhorst
B. Fitzroberger
M. Franzese
M. Frölich
D. Giannone
D. I. Harvey
N. Hautsch
D. A. Jäger
O. Jordà
A. Kanas
C. Kunschitz
T. Klein
K. Laliri
M.-J. Lee
S. E. Leher
B. Liesenfeld
H. Liersfeld
J. A. E. Machado
D. L. Millimet
J. Moeller
E. Moscone
C. J. O'Donnell
D. Papell
B. Perron
E. Pesavento
M. Pfaffermayr
P. Schmidt
R. C. Sickles
T. Stengos
M. G. Tsionas
A. Ullah
G. Uysal
F. Valid
M. Wagner
H.-J. Wang
J. Westerland
B. A. Wilke
F. Windmeijer
R. Winkel-Egger
A. Woodland
C. Wütsch
HONORARY BOARD
B. Fildes
W. Franz
L. R. Klein
W. Fildes
B. Raj

Why does monetary policy respond to the real exchange rate in small open economies? A Bayesian perspective
C.J. Garcia · W.D. Gonzalez 789

Evaluating changes in the monetary transmission mechanism in the Czech Republic
M. Franta · R. Horvath · M. Rusnak 827

Measuring the effects of monetary policy in Pakistan: a factor-augmented vector autoregressive approach
K. Munir · A. Qayyum 843

Zero lower bound, ECB interest rate policy and the financial crisis
S. Gerlach · J. Lewis 865

The predictive power of yield spread: evidence from wavelet analysis
A.B. Dar · A. Samantaraya · F.A. Shah 887

Sticky information and inflation persistence: evidence from the U.S. data
B. Molinari 903

Comparative fiscal illusion: a fiscal illusion index for the European Union
R. Dell'Anno · B.E. Dallery 937

A revisit on the tax burden distribution and GDP growth: fresh evidence using a consistent nonparametric test for causality for the USA
A.K. Tiwari · M. Mutascu 961

On the power of bootstrap tests for stationarity: a Monte Carlo comparison
S.G. Gulesserian · M. Kejriwal 973

 Springer *continued on back cover*

Your article is protected by copyright and all rights are held exclusively by Springer-Verlag Berlin Heidelberg. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Household food consumption, individual caloric intake and obesity in France

Céline Bonnet · Pierre Dubois · Valérie Orozco

Received: 27 June 2011 / Accepted: 6 February 2013 / Published online: 10 May 2013
© Springer-Verlag Berlin Heidelberg 2013

Abstract Using 2 years of home scanned data on household food purchases in France, we show how to infer the profile of average individual caloric intakes according to gender, age, and body mass index of household members. We also infer the individual consumption of macronutrients as carbohydrates, lipids or proteins. We then provide a descriptive analysis of eating in France over a long period of consumption that we compare to nutritional recommendations given by the French National Health and Nutrition Program. The results suggest that French people eat too much fats and proteins and not enough carbohydrates with respect to recommendations. They also show that obese or overweight individuals consume more calories at all ages and their consumption of fat is 20 % higher than normal individuals, meaning that public policies should aim at reducing fat consumption. We also find that the obesity of teenage boys cannot be explained by different food consumption as their dietary intake profiles are similar to other teenagers. Promoting physical activity for boys rather than public policies reducing food consumption could be more efficient.

Keywords Body Mass Index · Obesity · Nutrients · Energy · Individual food consumption

C. Bonnet (✉) · V. Orozco
Toulouse School of Economics (GREMAQ, INRA), 21 Allée de Brienne, 31000 Toulouse, France
e-mail: cbonnet@toulouse.inra.fr

V. Orozco
e-mail: valerie.orozco@toulouse.inra.fr

P. Dubois
Toulouse School of Economics (GREMAQ, INRA, IDEI), 21 Allée de Brienne,
31000 Toulouse, France
e-mail: pierre.dubois@tse-fr.eu

JEL Classification I18 · D12**1 Introduction**

According to the World Health Organization (WHO), obesity and overweight have become an international problem: more than 30 % of persons in the United States of America are obese, 24 % in Mexico, 23 % in United Kingdom, 21 % in Australia, 14 % in Spain and Germany and 10 % in France¹. Obesity has been linked to various medical conditions such as hypertension, high cholesterol, coronary heart diseases, type 2 diabetes, psychological disorders such as depression, and various types of cancer. The costs associated to the obesity are then huge. In the US, obesity costs more in medical care expenditure than cigarette smoking—around \$75 billion in 2003—because of the long and costly treatments for its complications (Grossman and Rashad 2004). Including the indirect costs such as lost days of work and reduced productivity in addition to direct costs such as personal health care, hospital care, physician services, and medications, Wolf and Colditz (2006) estimate the total cost of obesity in the US in 1995 at a total of \$99.2 billion. In France, Levy et al. (1995) find a conservative estimate of direct and indirect costs of obesity for France of more than € 1.8 billion for direct costs and € 0.1 billion for indirect costs in 1992. Controlling for age, gender, professional categories, and alcohol and cigarettes consumptions, Emery et al. (2007) find that obesity would imply an additional cost in health expenditure in France of between € 2.1 and € 6.2 billion per year. Using French data from the Decennial Health Survey, Paraponaris et al. (2005) show that the overweight and obesity status reduces the employability of workers. The obesity prevalence has then become a public health concern of governments. Actually, obesity and overweight are the results of energy imbalance between calories in and calories out. Food consumption is one of the variable that policies may try to influence. Moreover, policies aiming at modifying calories out may also depend on the level of food consumption as the extent of energy imbalance depends on calories consumption.

There is a growing interest in the analyses of the causes of obesity. Economists have thus proposed explanations of obesity looking at technological changes, changes in taste and consumer habits, and at changes in the social environment. According to Lakdawalla and Philipson (2009), declines in the real prices of grocery food items caused a surge in calories intake that can account for as much as 40 % of the increase in the body mass index of adults since 1980. Technological advances in agriculture caused grocery prices to fall and these declines caused consumers to demand more groceries (Philipson and Posner 2003). Technological changes in the home kitchen seem to have also fostered more calories intake because of new tools responsible for reduction in the time spent preparing meals at home (Cutler et al. 2003; Cutler and Glaeser 2005). Microwaveable meals and other foods that are easy to cook are desirable because they are quicker to prepare but they also have generally higher contents in calories and fats. Other factors have contributed to the growth in obesity like the decline in physical activity since 1980. Chou et al. (2004) find that the per

¹ Source: OECD Health Data 2008.

capita number of fast-food and full-service restaurants, the prices of a meal in each type of restaurant, food consumed at home, cigarettes, and alcohol, and clean indoor air laws, explain a substantial amount of the trend of obesity in the US since the 70s. [Classen and Hokayem \(2005\)](#) show that demographic factors, mothers' obesity status and education also affect youth obesity (in the US). For example, obese mothers are at least 23 % more likely to have an overweight youth than their peers with a Body Mass Index (BMI)² in the acceptable range. This result can come from either genetic transmission of obesity or from a technological explanation related to the food production process within the household. [Anderson et al. \(2003\)](#) find that the rise in average hours worked by mothers can account for as much as one-third of the growth in obesity among children in certain families. In part, the rise in obesity seems to have been an unintended consequence of encouraging women to become more active in the workforce.

In this context, the objective of this paper is to provide a descriptive analysis of individual food consumptions according to some demographics. We want to compare this global snapshot to nutritional recommendations of the national programs for Health and Nutrition.³ This is particularly of interest since the governments want to target obese people only to reduce the cost of public interventions. This paper focuses on the French case where obesity has actually been increasing since the 1990s. According to the 2003 Decennial Health Survey by the National Institute of Statistics and Economic Studies (INSEE) ([Paraponaris et al. 2005](#)), the percentage of overweight increased from 32.9 to 37.5 % and obesity from 6.3 to 9.9 % between 1980 and 2003.

To describe the French individual food consumption, we use a long period of 2 years of observations of all food purchases at the household level and we infer the profile of average individual food intakes according to gender, age, and BMI. The dataset we use records household food purchases covering 354 product categories in France over the years 2001 and 2002 (KANTAR Worldpanel). As this dataset does not contain nutritional information of purchases, we collected them and matched nutrients composition (calories, proteins, lipids, and carbohydrates) to each product that the household buys. To recover individual level estimates of food consumption from household food purchases, we use a method introduced by [Chesher \(1997, 1998\)](#) that allows to get individual level of consumption according to age and gender. As our evidence shows that there is an important source of heterogeneity explaining daily food intakes according to the BMI, we also use this information to get the individual food demands. We also infer the individual consumption of macronutrients (carbohydrates, lipids, or proteins) because most of recommendations of the French National Health and Nutrition Program focus on particular consumption of nutrients ([ENNS 2006](#)). For example, they advocate a reduction in added sugar consumption by 25 % or a reduction in lipids to less than 35 % of the total food calories intake. Our results suggest that French

² The Body Mass Index (BMI) is the most commonly measure to assess whether an individual is obese or not. It corresponds to the weight divided by height squared and an individual is considered as obese if his BMI is larger than 30.

³ For examples: the Food Standard Agency in United Kingdom, the French National Health and Nutrition Program, the Centers for Disease Control and Prevention in the United States.

people eat too much fats and proteins and not enough carbohydrates with respect to recommendations. They also show that obese or overweight individuals consume more calories at all ages (except for teenage boys) and their consumption of fat is 20 % higher than normal individuals, meaning that public policies should aim at reducing fat consumption. We also find that the obesity of boys could not be explained by different food consumption as their dietary intake profiles do not depend on BMI. Promoting physical activity for boys rather than public policies reducing food consumption could be more efficient.

As far as we know, few empirical studies describe the daily intake across individuals in France. The nutritional survey realized by the French National Agency on Health Safety from January 2006 to February 2007, called INCA2, is one example.⁴ This survey provides information on food intakes and health outcomes of individuals but usually over a relatively short period of time (seven days). Even if food intakes are exhaustively observed on these seven days, it is likely that variations over time of food intakes due to activity levels, occupation or seasonality will affect the precision of the estimated average long-term food consumption of individuals. Moreover, this type of survey consists in food frequency questionnaires that do not provide precise information on quantities and that are subject to major approximation errors by respondents. Finally, this survey only contains information of more than 4,000 individuals. We will see later that our analysis allows to describe food consumption of more than 22,000 individuals. Another study of food consumption in France is [Nichèle et al. \(2008\)](#). They analyze the long-term evolution of nutrition in France and its link to obesity but assume an equal division of food among household members. Our method differs and aims at finding how the division of food within the household is done according to observed individual characteristics without preventing an equal division of food between household's members.

Section 2 presents the data and some descriptive statistics. Section 3 shows how to use household level data to obtain estimates of individual level consumption and presents results of such a method on French data. Section 4 concludes. Appendices are in Section 5.

2 Data and descriptive statistics

2.1 Data sources

We use home scan data from the KANTAR Worldpanel company, providing information on household purchases covering 354 product categories over 2 years (2001–2002) in France for more than 8,000 French households. The data on food purchases are collected by home scan, which means that during the whole period in which households participate, they collect information on their purchases after each shopping trip at home. The collection of the information is done at home with a scanner which transmits the information to KANTAR Worldpanel. Households are chosen such that

⁴ Other national surveys exist: the National Health and Nutrition Examination Survey in the US for example. This survey collects the information over a 24-h dietary recall ([CDC 2012](#)).

they represent the French population according to some observable demographic and geographical characteristics.

Concerning household purchases, we observe the quantity purchased, the price, as well as a large set of characteristics of purchased goods (identified by their bar code). Our data also provide a detailed set of demographic characteristics of the household for each year, such as the number of persons and the number of children, the household income, their employment category, their region of residence and the type of residence (rental or owned), the town size, the qualifications and nationality of the person of reference. At the individual level, we have information on age and gender and also weight and height which allow us to compute the BMI of each individual.⁵

As the nutritional information for each purchased good is not recorded in this KANTAR Worldpanel database, we collected ourself data on the nutritional component of all products purchased in the database. In order to get an exhaustive information of all food items that French households bought, we considered several sources of information on the nutritional content.⁶ We then obtain the amount of energy in kilocalories (kcal) and in grams of proteins, fats, and carbohydrates per 100g for each of the 2,073 products. This nutrient information depends on the product characteristics. For example, we are able to differentiate plain yogurts according to their fat content. For snacks, we could go as far as brand level differentiation of nutritional content. Matching this information with household purchases, we obtain the total amount of nutrients or energy purchased. The nutritional composition of these 2,073 items was matched using characteristics of the products. Note that we are not able to consider time variation in the composition of these food items because it is difficult to have any time varying information on the nutritional composition of goods and we believe that the true variation must be relatively small. For many goods, the nutritional composition does not vary over a 2-year period of time (for example, soft drinks, yogurts. . .). For goods such as fruits and vegetables for which we can expect some variation, we need to use anyway some average approximate nutritional composition. For example, we use average nutritional composition of apples and strawberries without being able to observe if actual apples or strawberries purchased are more or less mature. This type of variation is impossible to observe with the available data.

Note that the available data on food consumption at the household level concern all food categories. These food items are classified into three categories corresponding to products with bar codes, to fruits and vegetables without bar code and to meat and

⁵ We removed 2,711 households that stopped participating in the survey at the end of 2001 and dropped 3,126 households because of missing information on age, gender, height, or weight for some individuals in the household. We finally used information of 4,166 households on the 2-year period that represent 11,237 individuals. This reduced sample has approximately the same characteristics than the initial one in terms of observables and thus attrition or missing values seem not systematically related to observable characteristics.

⁶ The different sources that allow us to build the dataset are: the Regal Micro Table, [Cohen and Sérog \(2004\)](#), nutritional web sites (www.i-dietetique.com, www.tabledescalories.com . . .) and food industry companies web sites (Picard, Carrefour, Telemarket, Unifrais, Bridelice, Andros, Florette, Bonduelle, McCain, Nestlé, Avico).

fish without bar code. For each household, all food purchases are collected except those in one of two categories without bar code (fruits and vegetables, or meat and fish) and purchases of products with bar codes are always collected.⁷ To overcome this problem of missing data of some food categories for some households, we implement a procedure of imputation at the household level. The method consists in using the full set of observed household characteristics to impute the unobserved value (quantity and expenditure of the unobserved food category) with the average value observed in households with the same set of characteristics. It has to be noted that the missing category is not systematically related to some household consumption behavior and concerns a small percentage of household consumption on average. Indeed, only 9.7% of the total food consumption is estimated from the imputation method.

To summarize this imputation procedure, let's define y_{it}^k the household i consumption for category $k = 1, 2, 3$ at period t and let us define $S_{it}^k \in \{0, 1\}$ equal to 1 only if y_{it}^k is observed. We denote W_{it} the large set of observed demographic variables and define $\omega_{it}^k = y_{it}^k - E(y_{it}^k | W_{it})$. We then assume that whether category k consumption is observed or not is independent of ω_{it}^k given all the observed covariates W_{it}

$$\omega_{it}^k \perp S_{it}^k | W_{it} \quad (1)$$

This independence implies the mean independence of y_{it}^k given W_{it} with the observation of y_{it}^k :

$$E(y_{it}^k | W_{it}, S_{it}^k = 1) = E(y_{it}^k | W_{it}, S_{it}^k = 0)$$

This implies that the conditional mean of household food consumption of category k is the same whatever it is observed or not. Households with characteristics W_{it} will thus have the same average consumption of category k on the sample for observed consumption and the sample of unobserved consumption. Conditioning on a lot of observed characteristics W_{it} is likely to explain a lot of variations across households and thus provides a way to impute the consumption of unobserved food categories of some households with the observed consumption of "similar" households. In practice, after some specification tests, the characteristics W_{it} include the declared household income, the household size, the age class of the household head, the gender and activity status of the individual making most food purchases in the household, indicators of socioeconomic class divided into 28 categories, indicators of the geographic region, 8 indicators of the level of qualification of the reference person, indicators of the citizenship of the reference person, the number of children under 16, the number of children under 6, 7 dummy variables for the type of housing, 8 dummy variables for

⁷ We observe food purchases for products with bar codes and meat and fish without bar codes for 60% of the sample, for products with bar codes and fruits and vegetables without bar codes for 38% of the sample and only for products with bar codes for 2% of the sample.

urban, rural and municipality population size. Details of the method used in practice are given in appendix 5.1.

2.2 Descriptive statistics on obesity

We use the BMI to define obesity. This is a measure of corpulence status defined as the weight divided by the height squared (expressed in kilograms per square meter) and it is used by most nutritionists and epidemiologists. Although BMI is conveniently measured and is used by the WHO and the National Institute of Health (NIH) to track obesity over time, it is not necessarily the best measure of obesity. [Wada and Tekin \(2010\)](#) and [Burkhauser and Cawley \(2008\)](#) show that BMI can be differently affected by fat-free mass and body fat and that it would be important to distinguish between these two, while BMI does not. This must be kept in mind even if for simplicity we use BMI as many policy institutions. Adult individuals are considered as overweight if their BMI is between 25 and 30, and obese if their BMI is equal or greater than 30. For children, we use the definition from the international corpulence curves for boys and girls under 18 years old ([Cole et al. 2000](#)), which define thresholds according to gender and age.

In Table 1, we can see that the average BMI is 23 (kg/m²) and that 9 % of individuals in our survey can be considered as obese. This percentage of obese people is consistent with the national figures in France obtained from other studies ([Obépi-Roche 2009](#)). Our figures are also consistent with the national percentage of overweight people. Indeed, one-third of adults are overweight, that is, more than 20 million people in France. Obesity is particularly prevalent among people over 60 years old since it represents 15 and 14 % of the population of men and women, respectively. While there is no great and significant difference on average between males and females for obesity rates,⁸ the percentage of overweight is statistically higher for adult men than adult women. For children, there is no statistical difference between girls and boys.

Thanks to the exhaustive data on weight and height of all household members in the survey, it is also possible to look at the within-household correlation of BMIs. Actually, food consumption is largely a household activity and it is often argued that individual consumption is greatly influenced by the household. We therefore look at the within-household proximity of BMI status by defining for each individual his or her BMI deviation from the average BMI of individuals of the same age and gender. Looking at the within-household standard deviation of these excess BMIs (positive if above the mean and negative otherwise), we find that they are significantly positively correlated to the average BMI of the household, even after controlling for household size and demographics. There is apparently more heterogeneity of individual BMIs within more “obese” households. To deal with obesity and food consumption issues, it would be better to look at food demands at the individual level rather than at the household level to take into account this household heterogeneity in terms of corpulence.

⁸ The obesity rates are not significantly different at 1 % level.

Table 1 BMI, obesity and overweight

Sample means		<i>N</i>	BMI (Std. Dev.)	% Obese	% Overweight	
All		22,469	23.04 (4.93)	8.71	25.59	
Adults	All	18,311	24.39 (4.25)	10.00	29.12	
	Males	All	8,401	24.86 (3.82)	9.53	35.25
		16–60 years old	6,340	24.37 (3.81)	7.71	31.32
	Females	All	9,910	23.99 (4.54)	15.14	47.31
		16–60 years old	7,217	23.40 (4.53)	9.15	19.26
	Children less than 16	All	2,693	25.55 (4.19)	13.78	36.43
Males		4,158	17.13 (2.96)	3.01	10.03	
Females		2,067	17.14 (2.99)	3.43	10.45	
		2,091	17.12 (2.93)	2.58	9.61	

Table 2 Variance of average daily calories consumption of singles with different periods

		$\sqrt{\text{var}\left(\frac{1}{T} \sum_{t=1}^T y_{it} x_i\right)}$			
		<i>T</i>	1 week	1 month	1 year
Gender	Age				
	Male	20–30	1,749	1,233	874
		30–40	1,534	936	654
		40–50	2,231	1,403	835
		50–60	1,997	1,231	859
		60–70	2,390	1,376	922
		More than 70	1,554	990	690
Female	20–30	1,542	946	616	
	30–40	1,501	917	623	
	40–50	1,811	1,053	619	
	50–60	1,788	1,177	880	
	60–70	2,007	1,230	858	
	More than 70	1,552	1,118	939	

Before turning to the methodology used to estimate average individual consumption, it is useful to see with simple descriptive statistics the role of the duration of observation of food consumption over time. Let's consider for now only single households. Denoting y_{it} the consumption on day t , we can empirically see that $\text{var}\left(\frac{1}{T} \sum_{t=1}^T y_{it}|x_i\right)$ is decreasing in T for different sets of individual characteristics x_i . Indeed, Table 2 shows that the variance of the average con-

sumption of single households given some characteristics x_i is decreasing with the duration of measurement of consumption. This result strengthens our arguments that the long period survey (2 years) we use allows to get more precise information about the consumption behavior than short period surveys (1 week for example).

3 From household to individual consumption

We first present an econometric model allowing to estimate the average individual level food intakes using household food purchases. We then apply this method to our data and present the empirical results in France.

3.1 Method of identification and estimation

Using the household measure of food consumption, we first present conditions under which “average” individual consumptions can be identified and estimated, where the “average” must be understood as the mean conditional on a set of individual characteristics. These conditions rely on conditional moments.

3.1.1 Identification

Let us assume that for a person p in a household i at period t , the individual food consumption y_{ipt} is

$$y_{ipt} = \beta(x_{ipt}) + u_{ipt} \tag{2}$$

where x_{ipt} is a vector of individual characteristics of person p and u_{ipt} is a deviation for this person’s consumption. The different measures of food consumption will consist first of calories and then proteins, fats, and carbohydrates. Then, household consumption y_{it} is

$$y_{it} = \sum_{p=1}^{P(i)} y_{ipt} = \sum_{p=1}^{P(i)} \beta(x_{ipt}) + \varepsilon_{it} \tag{3}$$

where $\varepsilon_{it} = \sum_{p=1}^{P(i)} u_{ipt}$ and $P(i)$ is the number of individuals in the household i .

Assuming that $\forall p, i, t$

$$E(u_{ipt} | x_{i1t}, \dots, x_{iP(i)t}) = 0 \tag{4}$$

implies that

$$E(\varepsilon_{it} | x_{i1t}, \dots, x_{iP(i)t}) = 0$$

which allows us to identify β non-parametrically.

The assumption (4) implies that in equation (2) $\beta(x_{ipt})$ can be interpreted as the average consumption by individuals whose characteristics are equal to x_{ipt} and u_{ipt} is interpreted as the deviation from the mean of individual caloric intake of this person. This assumption (4) also implies that the function $\beta(\cdot)$ is overidentified by the natural additive structure between individual consumptions imposed on total household consumption: $E(y_{it}|x_{i1t}, \dots, x_{iP(i)t}) = \sum_{p=1}^{P(i)} \beta(x_{ipt})$.

3.1.2 Separability

Denoting $\mathbf{x}_{it} = (x_{i1t}, \dots, x_{iP(i)t})$, we can test whether $E(y_{it}|\mathbf{x}_{it})$ is separable across different individuals' characteristics. This can be done after estimating non-parametrically $E(y_{it}|\mathbf{x}_{it})$ and testing

$$\frac{\partial^2 E(y_{it}|\mathbf{x}_{it})}{\partial x_{ir} \partial x_{is}} = 0 \quad \text{for all } r \neq s \text{ from } \{1, \dots, P(i)\}.$$

However, estimating second derivatives of a non-parametric conditional mean regression leads to very imprecise estimates and we are thus never able to reject the null hypothesis of separability only because of the large standard errors of the estimates of these derivatives. But this is clearly because this test has low power.

The separability assumption of the conditional mean of household consumption depends on the crucial choice of covariates \mathbf{x} . If one defines different covariates \mathbf{x} and \mathbf{z} for each individual, denoting $\mathbf{x}_{it} = (x_{i1t}, \dots, x_{iP(i)t})$ and $\mathbf{z}_{it} = (z_{i1t}, \dots, z_{iP(i)t})$, the assumptions

$$E(y_{it}|\mathbf{x}_{it}) = \sum_{p=1}^{P(i)} \beta(x_{ipt}) \tag{A}$$

and

$$E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}) = \sum_{p=1}^{P(i)} \delta(x_{ipt}, z_{ipt}) \tag{B}$$

are not equivalent and neither is more general than the other.

Actually, it could be that (B) is true but not (A) or the contrary. First, by the law of iterated expectations, assumption (B) implies that $E(y_{it}|\mathbf{x}_{it}) = \sum_{p=1}^{P(i)} E(\delta(x_{ipt}, z_{ipt})|\mathbf{x}_{it})$ which is not necessarily separable between any x_{ipt} and $x_{ip't}$, for example if some z_{ipt} is correlated with $x_{ip't}$ given x_{ipt} , which shows that (A) is not true in this case. Second, (A) may be true and not (B). For example if (B) is not true because $E(y_{it}|\mathbf{x}_{it}, \mathbf{z}_{it}) = \sum_{p=1}^{P(i)} \phi(x_{ipt}, z_{i1t})$ where $\frac{\partial}{\partial z_{i1t}} \phi(x_{ipt}, z_{i1t}) \neq 0$. (A) will nevertheless be true if z_{i1t} is independent of x_{i1t} because then $E(\phi(x_{ipt}, z_{i1t})|\mathbf{x}_{it}) = \beta(x_{ipt})$. It is thus important to choose carefully the set of characteristics \mathbf{x}_{it} such that the separability assumption is satisfied.

Adding BMI to age and gender in the conditioning variables is thus not an “extension” of Chesher (1997) in the sense of nesting the specification choices but the choice of additional explanatory regressors may be important to allow more individual characteristics to affect mean individual consumption and to validate the separability assumption (B) needed for identification, as we will see below.

Remark that Chesher (1998) also uses region of residence and income but with an assumption of multiplicative separability between individual characteristics and household characteristics (region and income) in the conditional mean. For simplicity, we chose not to add these multiplicative separable household characteristics in our specification and we prefer to be more precise on individual characteristics by adding BMI to age and gender. The use of additional covariates such as income and region of residence could be done but conditionally on age, gender, and BMI, there is less scope for significance of these household variables than when BMI is not used. On the other hand, Chesher (1997) concludes that the introduction of region and income has only a small effect on estimation.

De Agostini (2005) and Miquel and Laisney (2001) applied Chesher (1998) to datasets from the UK and the Czech Republic. Allais and Tressou (2009) applied it to break down the consumption of seafood across individuals in France to determine their exposure to methylmercury, which involves health risks. In these studies, to obtain consistent estimates of average individual consumptions for a person of a given age and gender, the authors need to assume that household level deviations ε_{it} (which are the sum of individual level deviations u_{ipt}) are not correlated with the demographic composition of the household in terms of age and gender. In particular, if the obesity status of individuals is correlated with age and gender and also with individual food consumption (which is intuitively likely to be the case), then biased estimates of individual consumption by age and gender will be obtained. For example if in couples, the man's BMI is correlated with the woman's age given the man's age, then it is likely that the separability assumption of the conditional mean consumption of households will not be true when using only gender and age as covariates. To show such an example, we regressed the BMI of the man on the age of the woman controlling for the man's age with year dummies for his own age and we find significant correlations between the man's BMI and the woman's age, indicating that conditioning on gender and age of both partners will not provide consistent estimates of average individual consumptions. Moreover, conditioning on age and gender only does not allow us to distinguish the consumption of two individuals of the same age and gender but with different anthropometric measures. It is also likely that a lot of heterogeneity in individual consumption is related to the obesity status. We will thus prefer to use age (x_{ipt}^1), gender (x_{ipt}^2) and BMI (z_{ipt}).

Let's define the BMI measure z_{ipt} , individual and household consumptions can be written as:

$$y_{ipt} = \beta \left(x_{ipt}^1, x_{ipt}^2, z_{ipt} \right) + u_{ipt}$$

and

$$y_{it} = \sum_{p=1}^{P(i)} \beta \left(x_{ipt}^1, x_{ipt}^2, z_{ipt} \right) + \varepsilon_{it}$$

where x_{ipt}^1, x_{ipt}^2 are respectively age and gender of individual p in household i at time t .

Remark here that there is no problem of reverse causality as it is not the BMI level but the BMI variation between two instants which is indeed the result of the difference between energy in and out during this same time period. Thus, we can use initial BMI denoted z_{ipt} as a determinant of next period energy intakes y_{ipt} without having causality problems. It is certainly true that $z_{ipt+1} - z_{ipt}$ depends on y_{ipt} but this relationship is not the one studied in this paper.

3.1.3 Specification

As age and gender are discrete variables and BMI is continuous, after some specification search, we choose to specify the function β as follows:

$$\beta \left(x_{ipt}^1, x_{ipt}^2, z_{ipt} \right) = \sum_{a=1}^{100} \sum_{g=1}^2 1_{\{x_{ipt}^1=a, x_{ipt}^2=g\}} \beta_a^g \left[\delta_0^g + \delta^g \left(x_{ipt}^1 \right) \left(\frac{z_{ipt} - \bar{z}_{a,g}}{\sigma_{a,g}} \right) \right] \tag{5}$$

where $\delta^g \left(x_{ipt}^1 \right) = 1_{\{x_{ipt}^1 \leq 13\}} \delta_1^g + 1_{\{13 < x_{ipt}^1 < 20\}} \delta_2^g + 1_{\{x_{ipt}^1 \geq 20\}} \delta_3^g$, and $\bar{z}_{a,g}$ and $\sigma_{a,g}$ are respectively the mean and the standard deviation of the BMI for individuals of age a and gender g (100 years is the maximum age in the population).⁹ With this specification, the continuous part of the function β in BMI z is supposed to be an age and gender specific linear function of the standardized BMI by gender and age.

3.1.4 Estimation with smoothing

Concerning the estimation of the model parameters, we can obtain consistent estimates by ordinary least squares using the specification (5). However, because of the discrete observation of ages, we introduce a smoothing technique (Chesher 1998) penalizing the non smoothness of the function $\beta \left(x_{ipt}^1, x_{ipt}^2, z_{ipt} \right)$ with respect to the age variable x_{ipt}^1 . The method amounts to estimating β (the vector of β_a^g for $a = 1, \dots, 100$ and $g = 1, 2$) of (5) as follows

⁹ Note that a constant term β_0 can be added to the previous specification and interpreted as a waste. Empirically, we will prefer the previous specification except for the disaggregation of lipids where it matters for fats used for cooking. The constant is then associated with the presence of a woman in a household.

$$\widehat{\beta} = (x'x + \lambda^2 W'W)^{-1} x'y$$

where λ is a penalization parameter, $W = I_2 \otimes A$, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and the matrix A of

size 98×100 is $A = \begin{bmatrix} 1 & -2 & 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & -2 & 1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & 1 & -2 & 1 \end{bmatrix}$.

3.1.5 Measurement errors

Although these home scan data provide precise and reliable information on all food purchases, measurement errors due to the lack of observation of wasted food or food consumed by people outside the household must be considered. Assuming that all these errors, denoted ς_{it} , are uncorrelated with individual characteristics $x_{ipt}^1, x_{ipt}^2, z_{ipt}$ of household members, the same method of estimation can be applied using the observed household purchases $\widetilde{y}_{it} = y_{it} + \varsigma_{it}$ instead of y_{it} to obtain consistent estimates of $\beta_a^g, \delta_0^g, \delta_1^g, \delta_2^g$ and δ_3^g .

Note that, although for single households, the household consumption \widetilde{y}_{it} is a consistent measure of the individual consumption ($E(\widetilde{y}_{it} | P(i) = 1) = \beta(x_{ipt}^1, x_{ipt}^2, z_{ipt})$), it is not necessarily more precise than the estimated $\widehat{\beta}(x_{ipt}^1, x_{ipt}^2, z_{ipt})$ because of the measurement error ς_{it} . Actually, assuming that measurement errors are independent of household size (for simplicity but the same result is also obtainable in general), the variances of each estimator are $V(\widetilde{y}_{it} | x_{ipt}^1, x_{ipt}^2, z_{ipt}, P(i) = 1) = V(\varsigma_{it} | x_{ipt}^1, x_{ipt}^2, z_{ipt}, P(i) = 1) = V(\varsigma_{it})$ and $V(\widehat{\beta}(x_{ipt}^1, x_{ipt}^2, z_{ipt}) | P(i) = 1) = \frac{V(\varepsilon_{it} + \varsigma_{it})}{\text{card}\{i | x_{ipt}^1 = a, x_{ipt}^2 = g, z_{ipt} = c\}}$. The second will in general be lower than $V(\varsigma_{it})$ if the number of observations such that $x_{ipt}^1 = a, x_{ipt}^2 = g$ and $z_{ipt} = c$ is large enough.¹⁰

3.2 Empirical results

We apply the previous method to the total energy purchased by the household over a year which we have been able to construct thanks to the data on all food purchases matched with the collected nutritional information. We also apply it to macronutrients such as proteins, fats, and carbohydrates.

¹⁰ Note that in our data the average size of these classes of individuals with the same age, gender, and obesity status is 10. Moreover, only 3.4% of the individuals belong to a class with no other individuals.

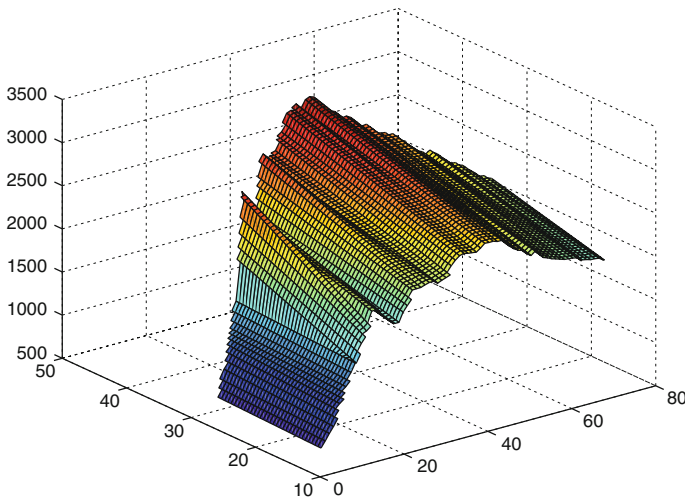


Fig. 1 Male individual energy consumption $\beta(\text{age}, \text{gender}, \text{BMI})$ (kcal/day)

As y_{it} is the total household yearly food consumption at home, the average individual food consumption $\beta(x_{ipt}, z_{ipt})$ corresponds to the average individual food consumption at home. To obtain an average daily food intake, we re-scaled the estimated individual food consumption using information of the number of meals away from home. As this information is not available in 2002 and 2003, we rather use a 2004 representative survey and apply the same disaggregation method described above for food consumption on the number of meals out of home to obtain an individual average conditional on his gender and age, and total household consumption. Hence, we can re-scale individual food purchases with an individual number of meals away from home. We denote n_{ipt} the number of meals taken out during a year t for person p in household i and n_{it} the total number of meals away from home for the household i during the year t . As $n_{it} = \sum_{p=1}^{P(i)} n_{ipt}$, we use that $E(n_{it}|x_{ipt}, y_{it}) = \sum_{p=1}^{P(i)} E(n_{ipt}|x_{ipt}, y_{it})$ to estimate $E(n_{ipt}|x_{ipt}, y_{it}) = \gamma(x_{ipt}, y_{it})$ using the 2004 survey. We specify $\gamma(\cdot)$ as a linear function of age and gender dummies (in x_{ipt}), controlling for the total food purchases of the household. Then, we obtain individual food intakes re-scaling $\beta(x_{ipt}, z_{ipt})$ using $\beta(x_{ipt}, z_{ipt}) \frac{n}{n - \gamma(x_{ipt}, y_{it})}$ where n is the maximum number of meals that the individual takes (that is 14 meals). For brevity, in the following, we denote as $\beta(\cdot)$ the re-scaled function.

Figures 1 and 2 present the graphs of the re-scaled estimated function $\beta(\cdot)$ (with penalization parameter¹¹ $\lambda = 500$). These graphs show that the individual energy consumption depends on the BMI of individuals and increases with BMI but with different slopes according to the age of the individual. The slope seems almost zero for young boys but larger and positive for adult men. For women, it seems that the caloric intake increases more clearly with BMI at all ages.

¹¹ Minimizing the sum of squared differences between the estimated consumption and the observed consumption for household with one individual, we find that the optimal penalization parameter is $\lambda = 500$.

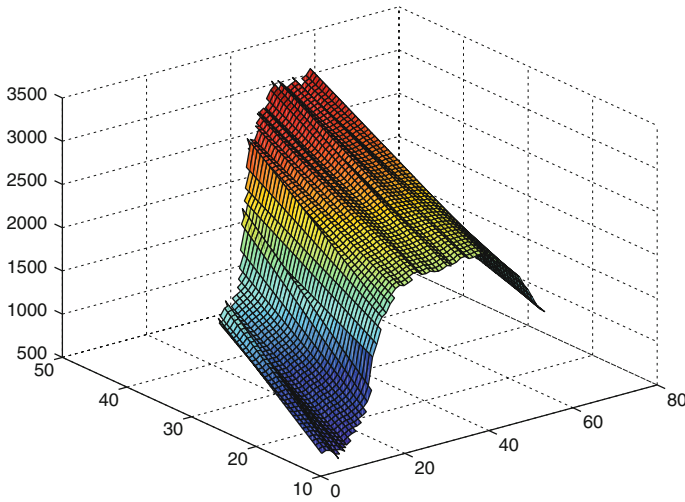


Fig. 2 Female individual energy consumption $\beta(\text{age}, \text{gender}, \text{BMI})$ (kcal/day)

Looking at the projection on the age hyperplane of the function β is also interesting to examine the age profile of individual energy consumption. Figure 3 presents such projections for three chosen categories of individuals defined as obese, overweight, and normal.¹² These graphs show mainly that the energy intake increases until 18 years old for both girls and boys (with a stagnation for boys between 8 and 11). Then, the energy consumption decreases until 25 and increases again until 70 for women and 55 for men. Chesher (1998) obtains the same shape of individual consumptions with respect to age with data from the UK. We also estimated these individual consumptions controlling separately for weight or height instead of BMI and we find the same shape of consumption with respect to age (see Figs. 5, 6 in appendix).

Finally, we can see that even if the age profiles of energy consumption have similar shapes across the three categories of individuals defined as normal, overweight and obese, the levels of energy consumption are clearly higher for obese than for overweight and for overweight than for normal. Moreover, it seems that overweight and obese people do consume more calories, especially during the periods of highest consumption, that is around 20 years. These curves also show that the differences between obese and non-obese people are greater for women than for men. Assuming that obesity comes from an excess caloric intake compared to calories spent through physical activity, the difference in obesity among women thus seems to be even more strongly related to differences in caloric intakes than for men, perhaps because of fewer differences in physical activity among women than among men (Barnekow-Bergkvist et al. 1996).¹³ Similarly, the striking feature that caloric intakes of obese, overweight and normal young boys does not seem to differ can be interpreted by the fact that

¹² In appendix, Fig. 4 presents individual calories consumption with 95 % confidence intervals.

¹³ See also http://www.science20.com/news_releases/women_exercise_less_men_study.

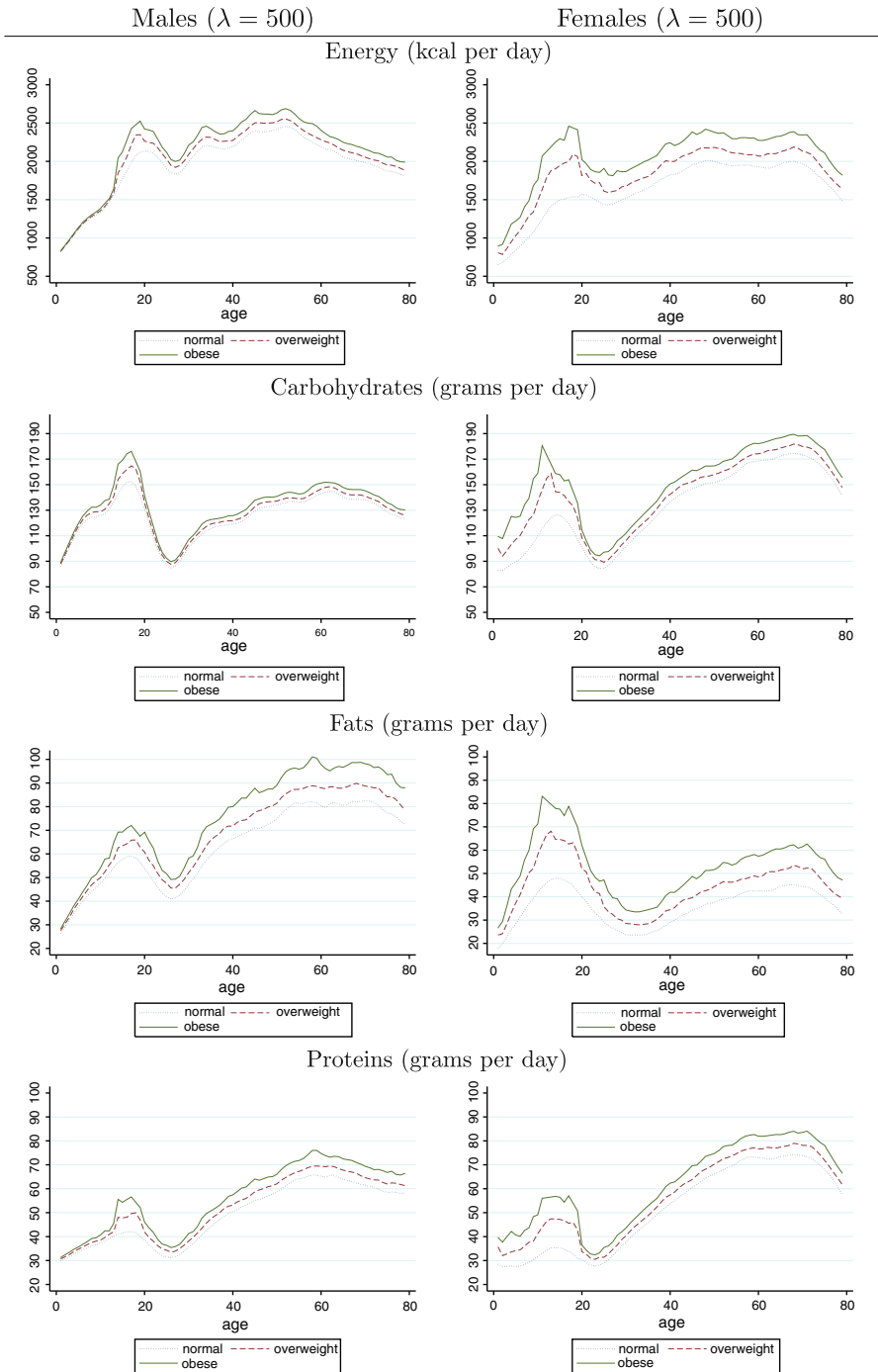


Fig. 3 Individual consumption of energy, carbohydrates, lipids and proteins

Table 3 Energy per household
($N = 4,166$)

	Share of energy from			
	Proteins	Fats	Carbohydrates	Other
<i>Note</i> $\alpha = 4$, $\theta = 9$, $\nu = 4$ Standard deviations are in parenthesis	0.147 (0.024)	0.430 (0.050)	0.379 (0.060)	0.042 (0.034)

physical activity (caloric expense) might be the sole source of variation that explains differences in BMI for these individuals.

Not surprisingly, this decomposition method provides patterns of individual food consumption that are very different than the one we would obtain if we had assumed an equal division of food across household members. Figure 7 in the appendix presents the individual caloric intake if we assume an equal division. We see that the food consumption for men and women is similar in that case whereas we find significant difference of the daily food intake across gender using the disaggregation method. Moreover, our disaggregation method allows to reveal more heterogeneity in food consumption across age than the equal division does.

Using the measures of proteins, carbohydrates, and fats, we estimate the corresponding individual quantities of nutrients consumed. Figure 3 also presents the age profiles of these estimated individual consumptions for normal, overweight, and obese people. It is interesting to see that the graphs of obese individuals are always above those of overweight which are above those of normal individuals, except for the carbohydrates consumption of men where all three graphs are very close. These graphs also show that boys under 10 have very similar consumption patterns for energy and all nutrient consumptions whether obese or not, which is not true for girls. More obese girls do consume more than less obese girls and this is true for all nutrient measures. The differences between obese, overweight, and normal people in terms of food intake is relatively the greatest for fats where, for example, after 35 years old, the obese eat on average more than 20 % more fat than normal weight individuals. They also eat more proteins and carbohydrates but the difference is relatively less. Finally, looking at the shapes of the age profile of carbohydrate, protein and fat consumptions, we observe that the increase in consumption lasts until 18 years old for boys and until 14 years old for girls whatever the nutrient. Also, the decrease in consumption at old age appears to be quite the same whatever the nutrient considered. This is clear after 60 years old for men and after 70 years old for women. We can also see that there are higher differences of the calories consumption between obesity status among women than men. This could be due to the consumption of carbohydrates where we find this result as well whereas we do not find it for proteins and fats.

Using the Atwater factors (α , θ , ν), used by nutritionists to calculate the energy content of food (Nichols 1994), the share of energy from proteins, fats or carbohydrates at the household and individual levels are presented in Tables 3 and 4. Table 3 shows that at the household level, proteins represent 15 % of energy, fats 43 %, and carbohydrates 38 %. Those results are far from nutritional recommendations (11–15 % for proteins, 30–35 % for fats, and 50–55 % for carbohydrates) and particularly for fats and carbohydrates. Table 4 also shows that the individual share of energy coming from fats is

Table 4 Estimated energy per person and shares of nutrients

		Energy	Share of energy from			
			Proteins	Fats	Carbohydrates	Other
All	(<i>N</i> = 11,237 indiv.)	1,881	0.164	0.361	0.428	0.047
	(<i>N</i> = 22,469 obs)	(425)	(0.029)	(0.077)	(0.064)	(0.023)
By gender	Males	2,035	0.145	0.426	0.378	0.051
	(<i>N</i> = 5,236)	(416)	(0.010)	(0.034)	(0.057)	(0.026)
	Females	1,747	0.181	0.304	0.472	0.042
	(<i>N</i> = 6,001)	(386)	(0.029)	(0.057)	(0.028)	(0.019)
Class of BMI	Gender					
	Normal	Males	1,917	0.141	0.410	0.398
(420)		(0.009)	(0.028)	(0.056)	(0.029)	
	Females	1,624	0.178	0.305	0.481	0.037
	(355)	(0.032)	(0.062)	(0.026)	(0.017)	
Overweight	Males	2,204	0.150	0.447	0.348	0.055
	(314)	(0.006)	(0.021)	(0.040)	(0.019)	
	Females	1,955	0.189	0.299	0.459	0.052
	(270)	(0.021)	(0.045)	(0.022)	(0.014)	
Obese	Males	2,283	0.154	0.468	0.341	0.037
	(390)	(0.005)	(0.029)	(0.045)	(0.020)	
	Females	2,194	0.189	0.310	0.440	0.061
	(296)	(0.013)	(0.032)	(0.020)	(0.014)	
Age	Gender					
	Adults	Males	2,209	0.146	0.438	0.357
(211)		(0.009)	(0.023)	(0.043)	(0.024)	
	Females	1,873	0.190	0.291	0.474	0.045
	(259)	(0.023)	(0.050)	(0.028)	(0.018)	
Children	Males	1,325	0.137	0.376	0.462	0.025
	(262)	(0.008)	(0.024)	(0.015)	(0.017)	
	Females	1,149	0.139	0.367	0.465	0.029
	(326)	(0.015)	(0.040)	(0.026)	(0.013)	

Note $\alpha = 4, \theta = 9, \nu = 4$

Standard deviations are in parenthesis

much lower and the share coming from carbohydrates is higher than at the household level. This sharing of energy from different macronutrients among individual members of the household is thus much more in line with dietary recommendations. It shows clearly that measuring the composition of food in terms of nutrients at the household level might be more alarming than the individual level estimates. It then highlights the importance of recovering individual food intakes. Moreover, our individual level estimates of nutrient shares are more similar to nutritional surveys. For example, the

INCA2 French study found 17 % for proteins, 44 % for carbohydrates and 39 % for fats.

Table 3 also shows that, at the individual level, the share of energy from carbohydrates is a little lower for more obese people who consume more of their energy in fats. This does not mean that obese people consume less carbohydrates as the absolute quantities consumed are higher but this means that the composition differs in proportion. Not only obese individuals eat too much but the composition of their eating is more distant to dietary recommendations. Ransley et al. (2003) similarly found in the UK that overweight households purchase a larger share of their energy in fat than lean households. Concerning children, they consume more of their energy from carbohydrates and less from proteins. In terms of gender comparison, women consume more of their energy in carbohydrates and less in fats than men.

4 Conclusion

In this paper, we estimate average individual food consumptions, in particular the energy content of food products and the main macronutrients such as fats, proteins, and carbohydrates, from household food consumptions. To do that, we use the disaggregation method of Chesher (1998) taking into account more heterogeneity in individual food consumptions than in usual studies, adding anthropometric measures like the body mass index to age and gender. Adding these exogenous variables is neither more general nor less but changes the required assumption for identification.¹⁴ We find that individual food caloric intake is clearly positively correlated with the BMI of men and women at all ages. Such a result is consistent with the fact that eating more makes individuals heavier, unless physical activity is very negatively correlated with eating. Actually, if eating more decreases weight but physical activity is very negatively correlated with eating, then it could be that individuals eating more gain weight not because they eat more but because they exercise much less. We do not need to interpret the correlation found in terms of causal relationship but the evidence is consistent with larger caloric intakes causing weight gain unless there is a large negative correlation with physical exercise.

However, there is an exception on the positive correlation between energy intake and BMI and it concerns the obesity and overweight of boys under 10 years old which seem not correlated with their total caloric intake, suggesting that differences in caloric expenditures may explain variations in BMI. We also find that overweight and obese people consume more fats than normal people in terms of share of their food caloric intake, obese people eating 20 % more fats than normal weight individuals, going away from nutritional recommendations. We find that children have a higher consumption

¹⁴ Using more information as provided by the data on household and individual characteristics (such as household income, region of residence, employment category for each household member, type of residence, town size) could be interesting but is left for future research. Here, we assume that conditional on gender, age, and BMI, these other characteristics either do not matter much or are not correlated with the current characteristics used.

in carbohydrates than adults. However, their consumption of carbohydrates is not different across obesity status.

The age profile of carbohydrates, proteins, and fats consumptions shows that the consumption increases until adulthood, lasts a few years less for fats and proteins than for carbohydrates. Also, the decrease in consumption at old age appears to begin first with the decrease in proteins, then carbohydrates, and last fats for men.

These individual food consumption estimates can provide some background information to adapt public policies aiming at reducing the prevalence of obesity and overweight by identifying the sources of overconsumption and the population at risk. These results suggest some possible targeting of prevention policies according to the gender and age of the population at risk, with policies that should insist on the reduction of total caloric intake in different proportions according to age and that could also target different categories of nutrients (carbohydrates, fats, and proteins). For example, as the caloric intake does not seem to be different for normal and obese boys and the obesity largely results in the energy imbalance between calories in and calories out, public policy could focus on promoting physical activities for children and particularly for boys. An important result of this paper is also the 20% difference in fats consumption between obese and normal people. Hence, a public policy aiming at reducing the consumption of fat products would allow to well target obese individuals. Future research could also investigate the causes of such a result to guides public policies. Another interesting result is the differences of carbohydrates consumption between obesity status that are higher among women than men. Future research could explore whether the differences in carbohydrates consumption come from the consumption of added sugar or starch. In the first case, a public policy that aims at reducing added sugar consumption of women could help to reduce the female obesity prevalence.

Acknowledgments This paper is part of the “Politiques Nutritionnelles, Régulation des Filières Alimentaires et Consommation” (POLNUTRITION) project funded by the French National Program of Research, Grant Agreement No. ANR-05-PNRA-012. We thank Namanjeet Ahluwalia, Andrew Chesher, Nicole Darmon, Catherine Esnouf, Rachel Griffith, Vincent Réquillart, François-Charles Wolff for useful comments. We also thank the editor and two anonymous referees for helpful comments. All remaining errors are ours.

5 Appendix

5.1 Imputation methodology

To overcome the problem of missing data in one of the categories without bar code, we implement a procedure of imputation at the household level.

As said above, let y_{it}^k be the household consumption for category $k = 1, 2, 3$ and let us define $S_{it}^k \in \{0, 1\}$ equal to 1 only if y_{it}^k is observed. With a large set of demographic variables W_{it} , we define $\omega_{it}^k = y_{it}^k - E(y_{it}^k | W_{it})$. The conditional independence assumption (1) implies the mean independence of y_{it}^k given W_{it} with the observation of y_{it}^k :

$$E\left(y_{it}^k | W_{it}, S_{it}^k = 1\right) = E\left(y_{it}^k | W_{it}, S_{it}^k = 0\right)$$

However, taking into account household heterogeneity by conditioning on as many variables as possible will lead to a more difficult non-parametric identification of the conditional means $E\left(y_{it}^k | W_{it}, S_{it}^k = 1\right)$ due to the lack of sufficient observations given the large dimension of W_{it} .

But, one can also use the fact that (1) implies (Rosenbaum and Rubin 1983)

$$\omega_{it}^k \perp S_{it}^k | P\left(S_{it}^k = 1 | W_{it}\right)$$

and thus

$$E\left(y_{it}^k | P\left(S_{it}^k = 1 | W_{it}\right), S_{it}^k = 1\right) = E\left(y_{it}^k | P\left(S_{it}^k = 1 | W_{it}\right), S_{it}^k = 0\right)$$

where $P\left(S_{it}^k = 1 | W_{it}\right)$ is the propensity score of observing category k . One advantage of such an implication is that we can then condition on a unidimensional variable, the propensity score, and thus solve the dimensionality problem of conditioning on a large set of variables.

Moreover, we can apply this to any conditional mean for intervals of the propensity score, conditional on some W'_{it} , because we have

$$\begin{aligned} E\left(y_{it}^k | W'_{it}, S_{it}^k = 0, p \leq P\left(S_{it}^k = 1 | W_{it}\right) \leq p'\right) \\ = \int_p^{p'} E\left(y_{it}^k | W'_{it}, S_{it}^k = 0, P\left(S_{it}^k = 1 | W_{it}\right)\right) dF_{P(W_{it}) | W', S_{it}^k = 0} \\ = \int_p^{p'} E\left(y_{it}^k | W'_{it}, S_{it}^k = 1, P\left(S_{it}^k = 1 | W_{it}\right)\right) dF_{P(W_{it}) | W', S_{it}^k = 0} \end{aligned}$$

where $F_{P(W_{it}) | W', S_{it}^k = 0}$ denotes the conditional cumulative distribution function of the propensity score given W' and $S_{it}^k = 0$.

We thus first estimate the propensity score $P\left(S_{it}^k = 1 | W_{it}\right)$ and then impute the unobserved category k households consumption with propensity score p with the average observed household consumption of category k food products with the same propensity score.

The regressions allowing us to estimate the propensity score matching for observation of the “fruits and vegetables” or “meat and fish” categories are not presented for brevity, but were done using a probit model.

5.2 Other Tables and Graphs

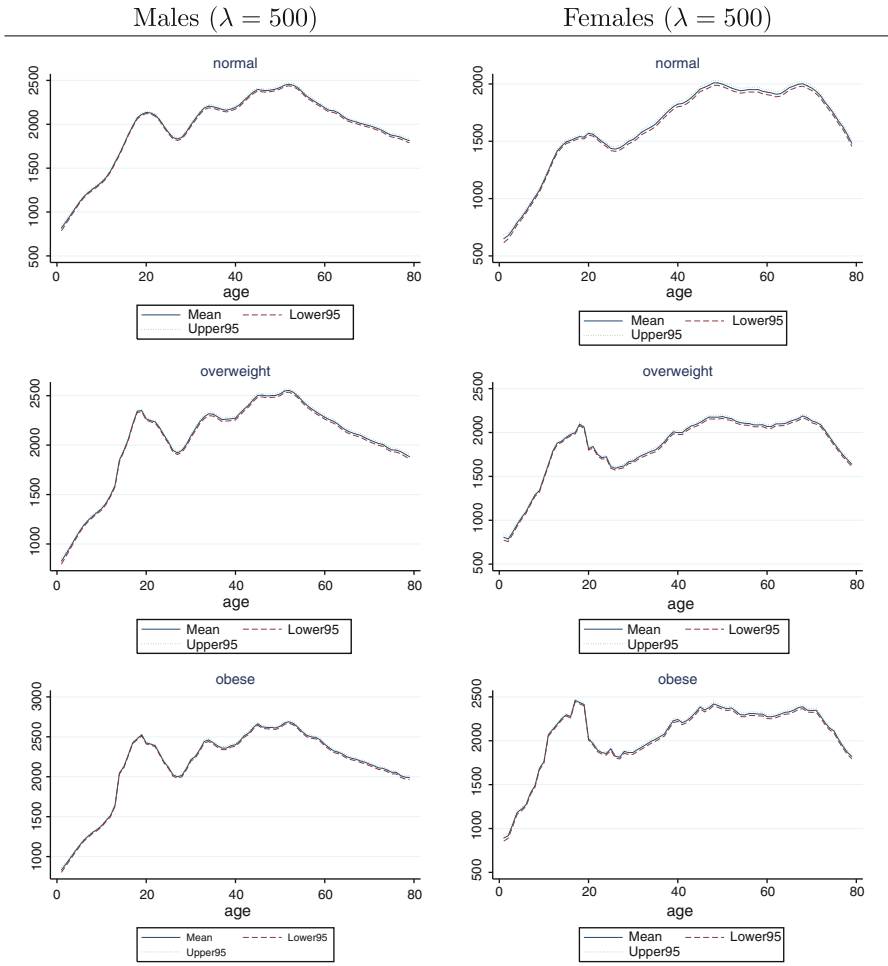


Fig. 4 Estimated individual calories consumption by obesity status per day

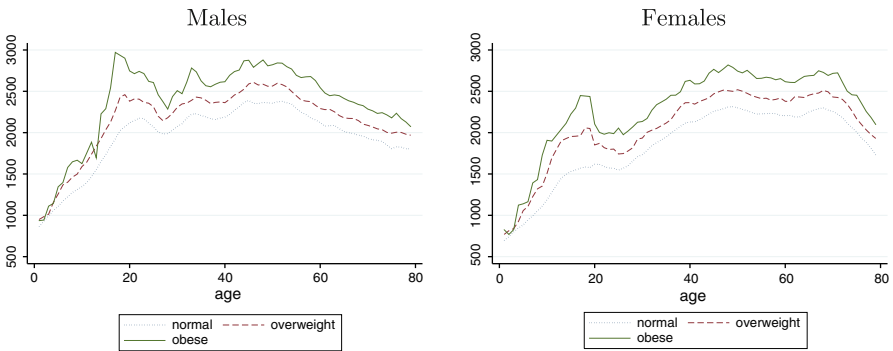


Fig. 5 Age profile of estimated calorie consumption per day with weight for z_{ipt}

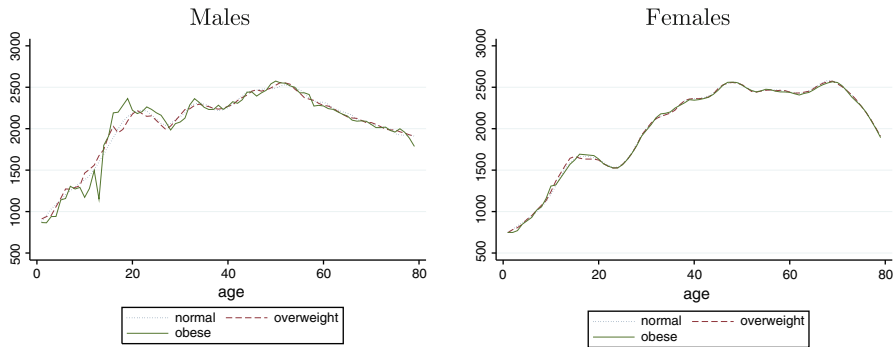


Fig. 6 Age profile of estimated calorie consumption per day with height for z_{ipt}

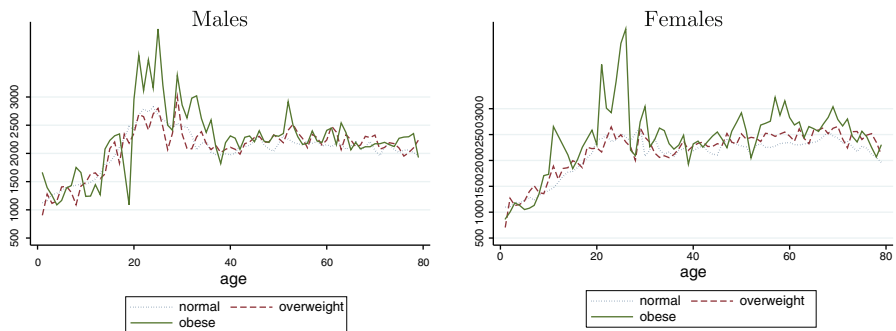


Fig. 7 Individual energy consumption if equal division within household (kcal/day)

References

- Allais O, Tressou J (2009) Using decomposed household food acquisitions as inputs of a Kinetic Dietary Exposure Model. *Stat Modell* 9(1):27–50
- Anderson PM, Butcher KF, Levine PB (2003) Maternal employment and overweight children. *J Health Econ* 22:477–504
- Barnekow-Bergkvist M, Hedberg G, Janlert U, Jansson E (1996) Physical activity pattern in men and women at the ages of 16 and 34 and development of physical activity from adolescence to adulthood. *Scand J Med Sci Sports* 6(6):359–370
- Burkhauser R, Cawley J (2008) Beyond BMI: the value of more accurate measures of fatness and obesity in social science research. *J Health Econ* 27(2):519–529
- CDC (2012) The National Health and Nutrition Examination Survey: Sample Design 1999–2006. Vital and Health Statistics Series 2 Number 155
- Chesher A (1997) Diet Revealed?: Semiparametric Estimation of Nutrient Intake - Age Relationships. *J R Stat Soc Ser A (Statistics in Society)* 160(3):389–428
- Chesher A (1998) Individual demands from household aggregates: time and Age variation in the composition of diet. *J Appl Econ* 13:505–524
- Chou SY, Grossman M, Saffer H (2004) An economic analysis of adult obesity: Results from the behavioral risk factor surveillance system. *J Health Econ* 23:565–587
- Classen T, Hokayem C (2005) Childhood influences on youth obesity. *Econ Hum Biol* 3:165–187
- Cohen JM, Sérog P (2004) *Savoir manger : Le guide des aliments*. Editions Flammarion
- Cole JT, Bellizzi MC, Flegal KM, Dietz WH (2000) Establishing a standard definition for child overweight and obesity worldwide: international survey. *Brit Med J* 320:1–6

- Cutler DM, Glaeser E (2005) What explains differences in smoking, drinking, and other health-related behaviors? *Am Econ Rev* 95(2):238–242
- Cutler DM, Glaeser E, Shapiro JM (2003) Why have Americans become more obese? *J Econ Perspectives* 17(3):93–118
- De Agostini P (2005) The relationship between food consumption and socioeconomic status: Evidence among British youths. ISER Working Paper 2005-21
- Emery C, Dinet J, Lafuma A, Sermet C, Khoshnood B, Fagnani F (2007) Cost of obesity in France. *La Presse Médicale* 36(6):832–40
- ENNS (2006) Etude Nationale Nutrition Santé 2006. Institut de Veille Sanitaire. http://www.invs.sante.fr/publications/2007/nutrition_enns/RAPP_INST_ENNS_Web.pdf
- Grossman M, Rashad I (2004) The economics of obesity. *Public Interest* 156:104–112
- Lakdawalla D, Philipson T (2009) The growth of obesity and technological change. *Econ Hum Biol* 7(3):283–293
- Levy E, Levy P, Le Pen C, Basdevant A (1995) The economic cost of obesity: the French situation. *Int J Obesity* 19(11):788–792
- Miquel R, Laisney F (2001) Consumption and nutrition: age-intake profile for Czechoslovakia 1989–1992. *Econ Transition* 9(1):115–151
- Nichèle V, Andrieu E, Boizot C, Caillavet F, Darmon N (2008) L'évolution des achats alimentaires: 30 ans d'enquête auprès des ménages en France. *Cahier de Nutrition et de Diététique* 43(3):123–130
- Nichols BL (1994) Atwater and USDA nutrition research and service: a prologue of the past century. *J Nutr* 124(9):1718–1727
- Obépi-Roche (2009) Enquête épidémiologique nationale sur le surpoids et l'obésité INSERM / TNS HEALTHCARE SOFRES / ROCHE http://www.roche.fr/fmfiles/re7199006/cms2_cahiers_obesite/AttachedFile_10160.pdf
- Paraponaris A, Saliba B, Ventelou B (2005) Obesity, weight status and employability: empirical evidence from a French national survey. *Econ Hum Biol* 3:241–258
- Philipson TJ, Posner RA (2003) The long-run growth in obesity as a function of technological change. *Perspect Biol Med* 46(3):87–108
- Ransley JK, Donnelly JK, Botham H, Khara TN, Greenwood DC, Cade JE (2003) Use of supermarket receipts to estimate energy and fat content of food purchased by lean and overweight families. *Appetite* 41:141–148
- Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika* 70(1):41–55
- Wada R, Tekin E (2010) Body composition and wages. *Econ Hum Biol* 8:242–254
- Wolf A, Colditz G (2006) Current estimates of the economic cost of obesity in the United States. *Obes Res* 6(2):97–106