UNIVERSITÀ DEGLI STUDI DI PISA & SCUOLA SUPERIORE SANT'ANNA

DEPARTMENT OF ECONOMICS AND MANAGEMENT

Master of Science in Economics

Let Them Eat Meat? Projections of Animal Source Food Consumption in the USA

Candidate:
Marta TALEVI

Supervisor:
Prof. Davide FIASCHI

Academic Year 2013-2014

Abstract

The objective of the present work is to obtain projections for the consumption of animal source foods in the United States, based on the distributions of the main predictors over the population and their forecasted evolution up to 2030. In order to identify the key drivers of consumption and the effects they entail at the invididividual level, a semi-parametric Generalized Additive Model is estimated for each food group using micro-data from the US National Health And Nutrition Survey (NHANES). The estimated model is then fitted to a dataset of forecasted individual characteristics to obtain the projections. A particular focus is put on the role of income and age and on the non-linearities they bring into the model.

"The errors which arise from the absence of facts are far more numerous and more durable than those which result from unsound reasoning respecting true data."

Charles Babbage, 1832.

Contents

List of Figures

List of Tables

1 Introduction

In recent time we have witnessed a dramatic increase in the demand for animal-source foods, mainly due to the incessant economic development of vast regions of the developing world, but in part also to the invariably positive growth of the consumption in advanced countries. Given the consequences that 'abuses' in the demand and consumption of these foods have on health and on the environment, we felt compelled to investigate what are the mechanisms that drive such demand at the individual level. For this reason we set to analyse how the quantity consumed of different animal-origin food groups relate to economic, socio-demographic and cultural characteristics of the individuals, using survey micro-data from the United States. To this purpose we choose a data-driven approach based on the use of semi-parametric econometric modeling. Once these models have been estimated, we will proceed to elaborate some projections of the average and overall demand in 2020 and 2030. Importantly, we want to ground our projections on micro-data and more specifically on the evolution of the distribution of the predictors. Given the structure of the data at our disposal – that are not panels but repeated cross-sections on representative but different samples – traditional forecasting techniques cannot be applied to the individual data. To overcome this hurdle we will put in place some expedients, estimating the joint density of the predictors in each year we have access to, and then forecasting its evolution in the next decades. A similar work has been carried out by Lin et al. (2003), which examined the quantities consumed in the US of various types of foods in the years 1994-1998 and then projected them to 2020. Our main addition to these kinds of analysis is to take into account the distribution of individual characteristics at each moment in time rather than just using the trends of the population average, and to let the data 'free to speak' by using semi-parametric techniques rather than more limiting parametric ones, as done by Gozalo (1997) and by Blundell et al. (1998). A similar analysis of demand for foods using non-parametric techniques has been recently done for China (Du et al., 2004).

Before starting our analysis, we would like to answer some methodological questions. Why the choice to study demand as a quantity rather than as expenditures, as usually done in microeconomics? While a vast amount of literature already exists on the expenditure side, less has been done on quantity. Moreover, due to the recent concerns about climate change and environmental sustainability of our everyday life, and with a view to the emission reduction agreements that are being discussed internationally, quantity reduction in the livestock sector appears to be a field in which there is much to gain, on more than one side. In fact, if the health of our planet is not enough of a compelling reason to push change – given its public-good and externality-rich nature – concerns about individual health may do the trick. Since these are the main motivations of our analysis, quantity consumed seems to us to be a much more insightful unit of measure than expenditure. Why projections based on micro-data and not just a macro-data time series analysis? Because together with Cirera and Masset (2010), we believe that distribution matters. In fact, rather than focusing on the average 'representative' individual, we want to understand how different individuals choose on the ground of their characteristics, and then forecast how these characteristics (or at least the main ones) evolve, not just on average, but in their distribution across the population. This is especially crucial in the presence of non-linearities in the elasticity structure of demand, as consumption would similarly show a non-linear evolution, despite linear changes in the regressors. Why a semi-parametric model? We choose to use a semi-parametric model for various reasons, but mainly because we are interested in capturing non-linearities. To this aim, we want to allow for a more complex elasticity structure, letting it change as a function of the characteristics rather than being a fixed coefficient (Hastie and Tibshirani, 1986). Why a focus on the United States? Although we were originally interested in studying the evolution of the elasticity structure in the developing world, according to economic development, urbanization, and westernization of tastes, the lack of available data for such countries – but also for developed countries (Hawkesworth et al., 2010; Kearney, 2010) – was an insurmountable hurdle for a quantitative rather than just qualitative analysis. On this same line, although several developed countries do have survey data on individual consumption, many of them have not been collected for multiple waves, or the collection methodology and institution in charge have changed, making it complicated to undertake any comparison, and therefore constraining our choice even further. The US have systematically carried out the National Health And Nutrition Examination Survey (NHANES) in two-year waves since 2001 and other data exists for the years 1994, 1995, 1996. A part from the data argument, the US are actually an interesting case-study under other points of view. In fact, although the increase in per capita animal-source food consumption in the developed world has slowed down recently, it does not seem to have inverted its direction nor to stop. This despite the achievement of an almost-saturation level and despite the chronic over-consumption and the damages linked to it (Alexandratos and Bruinsma, 2012). While this phenomenon is widespread in the developed world, it is especially true for the US (FAO, 2009).

In what follows we will provide an overview of the main motivations of this work, namely the consequences of animal-source food over-consumption on health and on the environment (Section 2). In Section 3 we will move to review the main macroeconomic trends worldwide, and then zoom in the different pictures displayed by developing and developed countries, with some illustrative examples from China, Brasil, India, Europe and the US. The spotlight will then be directed to US individuals: the NHANES survey and sample design will be presented in Section 4, while Section 5 will deal with the elaboration and estimation of our empirical model. Finally, Section 6 will illustrate the process used to forecast the main predictors and the projections obtained for the average and overall animal-foods consumption in 2020 and 2030. The last Section concludes.

2 The two sides of the coin: environment and health

Given the dramatic and continuous increase in animal-source food demand highlighted by Popkin (1994) in the developing world and the subsequent livestock revolution reported by Delgado et al. (1999, 2001) in place to satisfy it, it is important to realize what are the consequences linked to the new dietary patterns, and whether – and in what form – an intervention may be needed. In particular, as the Bloomberg School of Public Health (2013) brought to light, animal food production has worrisome implications in many fields, from individual and public health, to the ecological equilibrium of the production site, to climate change and food security. And production is only a part of the overall life-cycle. Together with sustainability of production systems, in fact, it is equally important to look at the sustainability of consumption, as highlighted by the 'demand constraint approach' (Garnett, 2013). Moreover, as stated by Hawkesworth et al. (2010) and Heller et al. (2013), food production – and even more so the livestock sector – is mainly driven by demand, a compelling argument for an investigation into the cause-effect relationships that animal-source food consumption triggers.

Livestock production in the developing world is projected to increase sharply in the next decades in order to catch up with global demand – estimates by Smil (2002) suggest that by 2050 global food production would need to be double that of the early 2000s – placing even more strain on already stressed resources as land and water. Taking together the land used for grazing and the cropland for animal feeding, the livestock sector is already the largest user of land – 80% of the total anthropogenic land use, according to Stehfest et al. (2009)– and an intense consumer of water. Globally, over one billion tons of cereals will be allocated to feed for the livestock sector in the year 2040 (Alexandratos and Bruinsma, 2012), equivalent to more than a third of the total projected production of cereals, and almost as much as the total amount of cereals that will be destined to human consumption. As for water, it has been estimated that in order to produce just one kilogram of beef, the need of water oscillates between 13,350 (Hoekstra and Chapagain, 2007) and 20.860 (Kreith and Davis, 1991) litres, making meat and dairy production alone responsible for 27% of the human water footprint (Hoekstra and Mekonnen, 2012). Such figures are likely to worsen progressively. In fact, growth in production will not come without constraints to face. Land and water available in per capita terms are reducing, not only quantitatively but also qualitatively, due to pollution, soil degradation, and similar problems. And efficiency gains in terms of input resources needed for unit of output comparable to those allowed by technological progress in other industrial sectors, are unlikely to occur in agricultural processes, and even less so in animals' metabolism (Hertwich, 2010). On the contrary, part of the pollution and degradation problems are a direct consequence of food consumption itself. Tukker et al. (2011) compared the impacts of different dietary patterns in Europe and dietary recommendations in order to quantify the differences in the environmental impact they cause, and conclude that when the life-cycle is taken into account, food consumption is one of the most impacting activities among all the categories of human consumptions, and meat and dairy are in turn the most impacting within this category. In the US, according to the US Greenhouse Gas Inventory compiled by the United States Environmental Protection Agency (EPA), the chain of producing, processing, distributing and retailing products of animal origin, is responsible alone for about 9% of total greenhouse gas emissions of the country, while worldwide estimates go up to more than 15% (Gerber et al., 2013; Steinfeld et al., 2006), of which more than 60% are associated with enteric fermentation and manure of the animals themselves, mechanisms that men cannot do much to reduce. With respect to a legume such as soy, pork and poultry are estimated to produce around 20 to 30 times the volume of noxious emissions per unit of protein, while cattle can go up to 150 times (Bailey et al., 2014; Nijdam et al., 2012; González et al., 2011). Beef and dairy production are therefore especially problematic in this sense, as cattle needs on average three times the feeding caloric intake of pigs and poultry (Smil, 2002), and ends up emitting about one third of the total greenhouse gas associated with dietary choices in the US, despite representing only about 4% of the retail food supply by weight (Heller and Keoleian, 2014). Nevertheless, given that consumption of pork and poultry is actually greater than that of beef almost anywhere in the world, and their consumption is on the rise, the total amount of emissions of these subgroups is also highly significant. In fact, IPCC reports (Metz et al., 2007; Edenhofer et al., 2014) qualify meat and dairy among the greatest contributors to global emissions together with power production, industry, and deforestation, and ahead of transports exhaust emissions, heating and cooling in buildings and disposal and treatment of waste. And yet this role of the livestock sector is hardly recognized by the public opinion. This perception gap has been recently highlighted by the survey carried out by Ipsos-MORI and Chatham House (Bailey et al., 2014) to explicitly explore the extent of public awareness on the links between everyday dietary choices and climate change. Only about 30% of the respondents stated that they did believe meat and dairy production to contribute "a lot" to climate change, as opposed to 70% for deforestation and more than 60% for power production, industry, exhaust emissions, and disposal and treatment of waste, despite the emissions of the latter sector is just about a third of those linked to the livestock sector.

As we know, environmental externalities such as those we have described so far, are unlikely to trigger much reaction until internalization mechanisms are put in place to transform them in actual costs. Even less so if we consider environmental quality as a public good. Nevertheless, an important array of health-related consequences, both at the public and the private level, and the monetary costs associated with them may configure a convenient win-win situation. Although public health is 'just another public good', problems such as antibiotic resistance, feed additives and epidemics due to the concentrations of livestock and their closeness to inhabited and urban areas have by now reached some alarming peaks (Walker et al., 2005; Silbergeld et al., 2008). Even more compelling from an individual point of view, the connections between red and processed meat and animal-source food consumption in general, and the incidence of many diseases and cancers are by now hard facts. Just to cite some medical studies on the matter, Pan et al. (2012) showed that regular consumption of red meat increases by 13% the risk of dying from cardiovascular disease and by 12% the risk of diabetes for every additional 85g portion. Positive links were found with heart disease (Micha et al., 2010; Pan et al., 2012; Sinha et al., 2009), stroke (Kaluza et al., 2012), type 2 diabetes(Micha et al., 2010), obesity (Wang and Beydoun, 2009), certain cancers, such as colon and prostate cancers, (Pan et al., 2012; Marmot et al., 2007), and earlier death in general (Pan et al., 2012; Sinha et al., 2009). Although per capita poultry consumption has increased in the US[1], red meat still constitute the majority of meat consumed by the average Americans (Daniel et al., 2011).

The health and environmental issues we have reviewed so far also come with their economic costs. Barnard et al. (1995) compared the health care costs associated with vegetarians and omnivores – controlling for several other elements of lifestyle – and estimated that in 1992 the total direct medical costs attributable to meat consumption in the US was between 28.6 and 61.4 USD billion. According to the National Health Expenditure Accounts (NHEA) kept by the Centers for Medicare and Medicaid Services in that same year the total national health expenditure totaled 857.9 USD billion[2], meaning that direct medical costs attributable to meat consumption equal between 3 and 7% of the total national health expenditure. A more recent study by Finkelstein et al. (2009) did a similar analysis focused on obesity, one of the main health problem related to over-consumption of animal-source foods and fats affecting a substantial portion of the US population[3]. The estimated annual medical cost of obesity alone in the US in 2008 was estimated in 147 USD billion, a per capita expenditure 1,429 USD higher than that for normal weight individuals, equal to 6% of the total national health expenditure[4]. Diets poorer in animal-source foods, and especially in beef and dairy, would also reduce environmental impacts, resulting in important savings in mitigation expenses that would otherwise be needed. Stehfest et al. (2009) claim that worldwide adoption of the 'Harvard healthy diet', including no more than 70g beef, 70g pork, and 320g of chicken meat and eggs on average per week, could reduce mitigation costs to reach the required CO2 emission stabilisation target for 2050 by half, thanks also to the absorptive power of regrowing vegetation in grazing fields. Unfortunately, as we are about to see in the next section, although in many developed countries red meat is being slowly replaced by healthier and less environmentally-harmful poultry, population growth and the nutrition transition occurring in a vast array of developing countries mean that the worldwide trend for meat – and in particular beef – consumption growth is still a positive one.

[1] Disaggregated trends of meat and other animal-source foods consumption are presented in the next section.

[2] National Health Expenditure Accounts (NHEA), Centers for Medicare and Medicaid Services.

[3] According to DGAC (2015) about two-thirds – or 155 million people – in the US are overweight, half of which have been qualified as obese by the US Centers for Disease Control and Prevention (CDC).

[4] According to the National Health Expenditure Accounts (NHEA) of the Centers for Medicare and Medicaid Services, the overall national health expenditure in 2008 was 2.414 billion USD.

3 Trends in animal food consumption: a demand-side story

Consumption quantity used in the analysis of macro-trends, are mostly estimated on the ground of 'food availability' in a country, i.e. netting domestic production by imports and exports and then adjusting for the various uses the commodity may have besides human consumption. Such method is used notably in the FAO 'food balance sheet', from which most of the data in the present section are taken. On the contrary, the micro-analysis that follows and that constitutes the main body of this work, is based on survey data actually collected from individuals on how much of each specific food they have consumed in a given span of time (24 hours). While the latter represents a reliable measure of consumed quantities, the former is systematically over-estimating the true quantities as the 'food availability' does not take into account retail- and consumer-level wastes and other elements that would be needed to get to the actual figures, so that this discrepancy should be taken into account when comparing figures from the different sections. Nevertheless, what we are interested to present in this section are the global trends in food consumptions and how they distribute in different area of the world. From Figure 1 it can be seen that while the developing world is still lagging behind but not too far away in the consumption of vegetal origin foods – namely cereals, roots and tubers, sugar and vegetal oils – and is actually ahead in the consumption of pulses, what really makes a difference in the dietary composition between the developed and the less-developed world is the consumption of animal-source foods. Currently, all the major meat

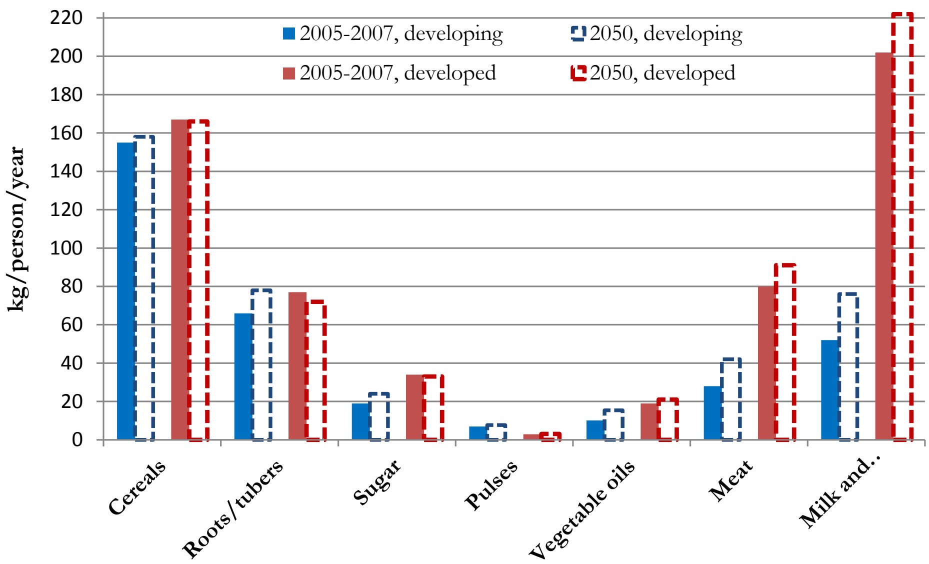


Figure 1: Food consumption per capita in the developing (blue) and developed (red) world for 2005-2007 and projected to 2050. Source: data and projections from Alexandratos and Bruinsma (2012, p.44). Meat includes bovine, ovine, poultry and pig-meat.

consumer countries (over 90kg of meat per year per person, corresponding to 250g per day) are high income countries of the Western world, with the exception of Argentina. According to FAO food balance sheets and to Alexandratos and Bruinsma (2012), in 2005-2007 people in developed countries consumed three-times more meat and four-times more milk and dairy per capita than their counterparts in less-developed countries, but the picture for the developing world is a rather heterogeneous one, with more advanced regions such as Latin America and the Caribbean showing levels of animal-source food consumption comparable to those in the developed world, followed by the fastest growing economies in Asia, and then– at a distance –the least-developed countries (Figure 2). Several reasons can be called upon for this gap, from level of socio-economic development, to cultural proximity with the Western world, to large domestic production and increased availability of meat. FAO food balance sheets'

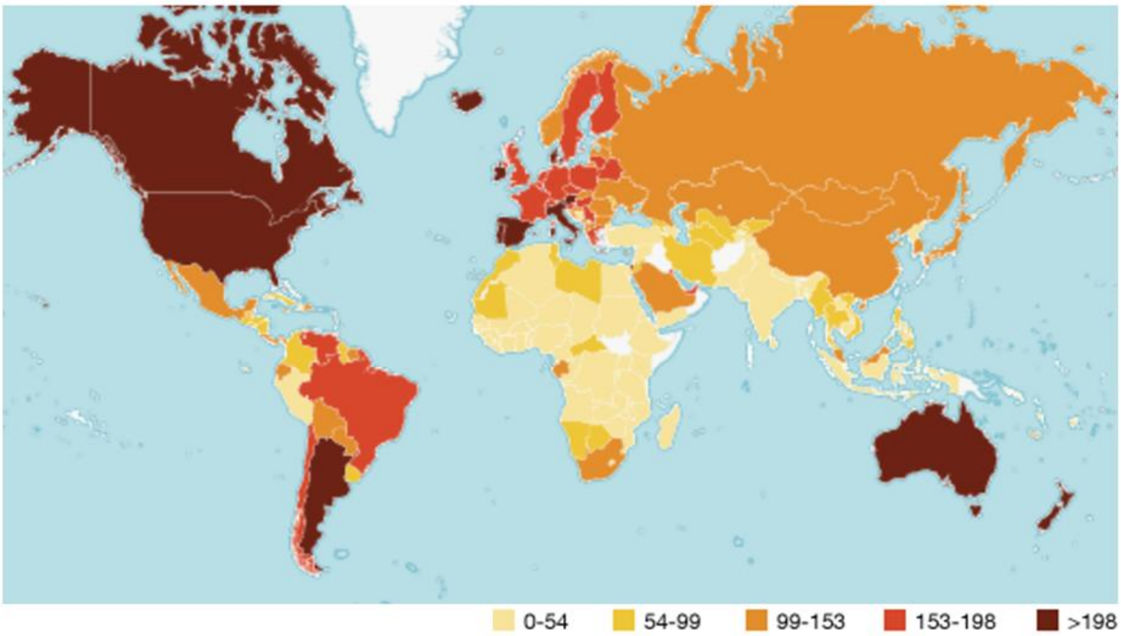


Figure 2: World average meat consumption per person in 2007. Measures are in pounds (1pound = 0.45kg). Source: Food and Agriculture Organization of the United Nations (FAO) 2010, Livestock and Fish Primary Equivalent, FAOSTAT on-line statistical service.

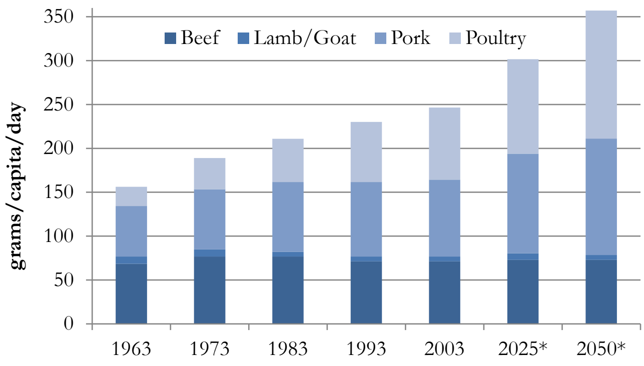
data and projections allow us to zoom in how the present levels of consumption have evolved – and are likely to evolve – over time in different regions. In Figure 3 and 4 we have presented the most relevant ones. The trends for overall meat consumption are all on the rise, although at different speeds. In the developed countries, growth is mainly led by increasing consumption of poultry, especially so in North America, while beef and pork seem to have stabilized, with the former decreasing slightly. Poultry, who is generally regarded as a healthier alternative to red meat, does not really seem to substitute for it, since its rise is not counterbalanced by an equivalent decrease in the consumption of beef and pork. It is to be noticed that North America has very high levels of meat consumption per capita, currently at more than 300g per day and forecasted to grow, while Europe scores at about 200g[5]. With respect to the subgroups, consumption of beef in North America is double that of Europe. In

[5]It is to be remembered that these quantities are measured as 'food available', and are therefore likely to over-estimate slightly the actual consumption.

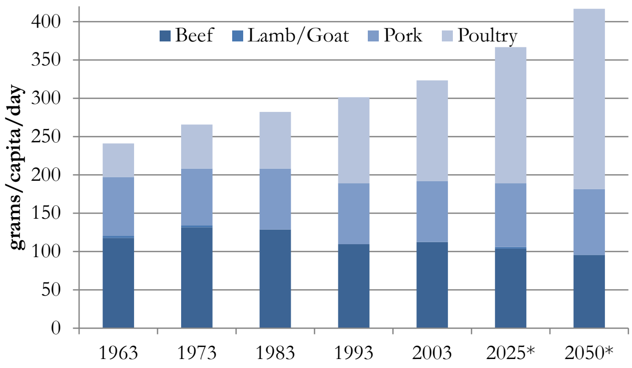
developing countries the quantities consumed are much lower, although the trends for all the meat subgroups are dramatically on the rise, an important piece of evidence that a 'nutrition transition' is in place, and it is occurring at a much faster pace than it did in the developed world, as highlighted by Popkin (1994, 2002). The two leading countries in this race to catch up with the Western world are in particular China and Brasil, which display similar growth trends for overall meat consumption, but very different within composition. China's staggering growth rates are forecasted to match European levels by 2050, while Brasil, together with Argentina, has already overtook the European per capita consumption level and is set to reach North America in the next decades. If on one side the Brasilian diet appears to be based predominantly on the consumption of beef and poultry, on the other side more than two-thirds of China's meat demand is made of pork. These different preferences may be attributed to the fact that Brasil and China are respectively among the top producers of cattle and pork. Moving to eggs and dairy products, the global trend appears to be completely flat, but the aggregation actually hides a much more interesting and complex picture, where the developed and the developing world are moving in opposite direction. North America in particular is dramatically cutting on these non-meat animal-foods, while Latin America and Asia are rapidly closing the gap. An important player in this field that did not appear as a significant meat consumer in the previous figure is India. Together with the highly positive trends that we can see in per capita consumption for China and India, we must further factor in the fast growth of their population in order to understand the scope of the nutrition revolution in place (Delgado et al., 2001).

3.1 Developing countries and the nutrition transition

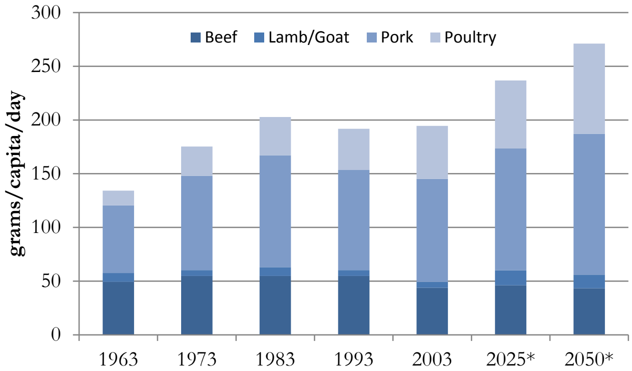
Income growth and development are not only linked to the widely known demographics and epidemiologic transitions, but also with a nutrition transition (Popkin, 1994; Drewnowski and Popkin, 1997), that is now happening at a much faster pace than it did for the advanced economies. Structural change in diets, both on a qualitative and quantitative point of view (Delgado et al., 1999; Pingali, 2007), have been particularly apparent in East Asia, due to their astonishing economic growth, but evidence can be found even in Africa and in much of Latin America. The key phenomenon of the nutrition transition is a 'westernization' of diets, where staples are being gradually replaced by animal-foods, fats and oils, but even vegetables and fruit, expanding the variety of choices (Pingali, 2007; Goodland, 2001). If rises in income, urbanization and lifestyle changes, linked for example to more sedentary occupations have led the convergence from the demand side, on the supply side, increasing availability due to increased interconnectedness and trade partnerships brought forth by reductions in transportations and conservations costs and customs fees, liberalization of FDI and multinational corporations, fast-foods and western supermarkets presence in lower income countries have facilitated these shifts (Pingali, 2007). And this dizzying trajectory of growth in animal-origin foods is likely to continue, unless intervention aiming at sustainable consumption, side by side the more talked about sustainable production, are put in place. Following industrialization and urbanization, in fact, the incidence of over-nutrition, obesity, and illnesses linked to excessive consumption of meat and animal fats is superseding that of de-nutrition and malnutrition in many regions of the less-developed world (Popkin et al., 2001). Maybe



(a) *Developed countries*



(b) *North America*



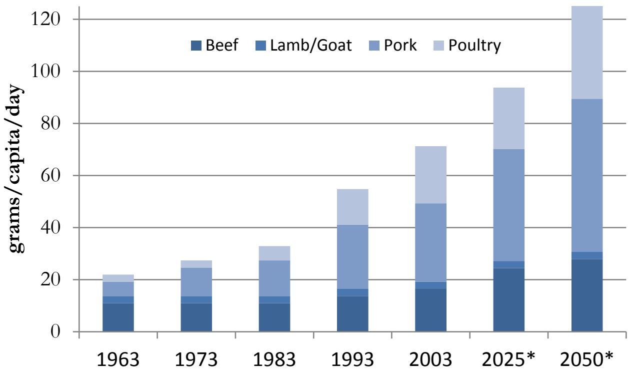
(c) *Europe*

counter-intuitively, these trends have been shown in some cases to hit harder on the poor, as revealed by Du et al. (2004) in China. Some staggering examples of the coexistence of de-nutrition and over-nutrition are presented by (Popkin, 2002) about India, where besides a 36% of under-nourished women (Body Mass Index$<$18.5), 11% can be classified as overweight and 2% as obese (BMI $>$ 25 and 30 respectively). Similarly, South Africa score rates of overweight women up to 44%, as well as widespread obesity. These two phenomena have even been seen even to coexist within same families, as found in Indonesia, proving that the drivers behind dietary choices are complex and go beyond a simple correlation with economic wealth. Importantly, despite the stunning economic and technological growth these newly industrializing countries are experiencing, their health systems are still struggling with the typical health problems of the less-developed countries, among which de-nutrition and malnutrition, and are not prepared to deal with the health consequence of over-nutrition and the chronic diseases they cause (Popkin, 2002).

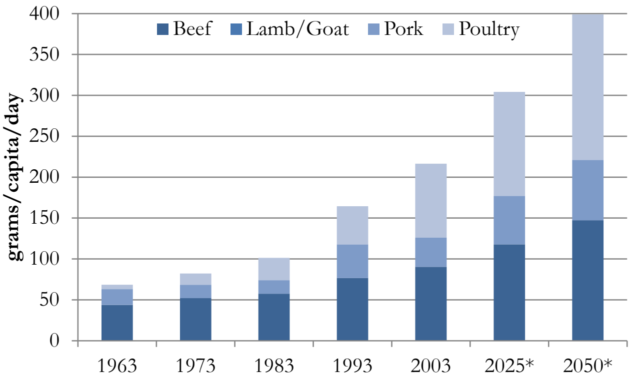
3.2 Developed countries and the USA

Contrarily to the nutrition revolution happening in vast part of the developing world, our hypothesis is that per capita consumption trends in the developed world ought be more stabilized and change little in the short-to-medium term. Even in countries where health campaigns are pushing towards changes in dietary habits, both in terms of quantities consumed and foods chosen, these are not likely to occur immediately nor dramatically. As behavioural studies have shown, habits are difficult to modify once they have been acquired (Gneezy et al., 2011; Becker and Murphy, 1988), and high rate of time discounting make health-problems that will arise in the future less of a compelling concern for present-time behaviour (see Frederick et al., 2002 for a review on the matter). In fact, although near-saturation levels of food consumption have been reached and the harmful effect of animal-foods overconsumption on health is now largely recognized, the data show no evidence of significant curtailment. *Au contraire*, even in countries where meat consumption has overstepped the 100kg per capita per year (corresponding to 250g per day), growth rates are still positive, and significantly so as Alexandratos and Bruinsma (2012) projections have alerted. At the regional level, AGRI (2012) has forecasted a 3% increase with respect to current per capita meat consumption in the EU by 2020. Similarly, although meat consumption in the United States has nearly doubled in the last century and Americans are now among the top per capita meat consumers in the world, with the average American eating more than three times the global average (FAO, 2009) and adult men having a protein intake double the level recommended, USDA (2010) expects a further 2% increase between 2010 and 2020. Another worrisome factor of Western diets is their composition. If on one hand the healthier poultry consumption has increased greatly in the last decades, this has not been followed by an equivalent reduction in red meat, that still makes up roughly half of the daily meat intake of Americans. The UK Department of Health in 2011 advised people who eat more than 90g of red meat[6] a day to reduce it down at least to 70g. As it can be seen from Figure 5, the US have been consuming almost double as much at least since the beginning of the last century.

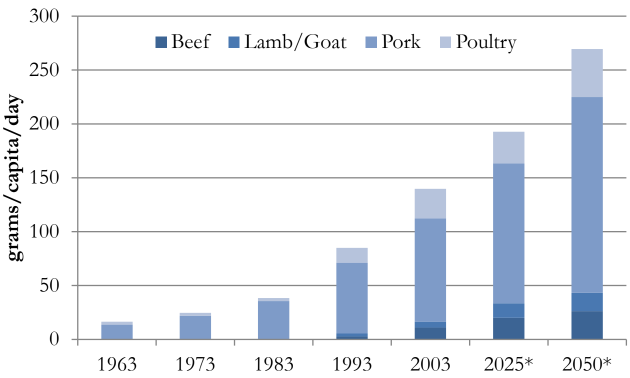
[6]The UK Department of Health's definition of red meat includes beef, lamb, pork, veal, venison, and mutton. We have adopted the same definition throughout the study.



(d) *Developing countries*

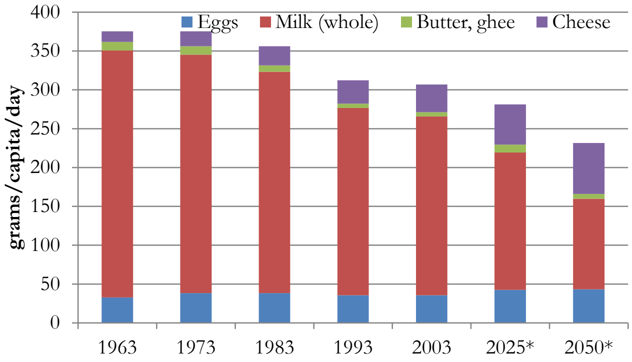


(e) *Brasil*

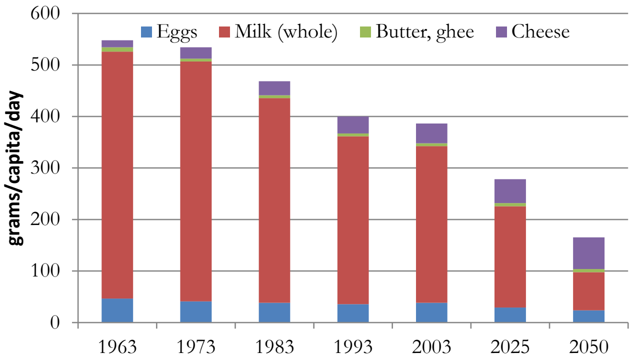


(f) *China*

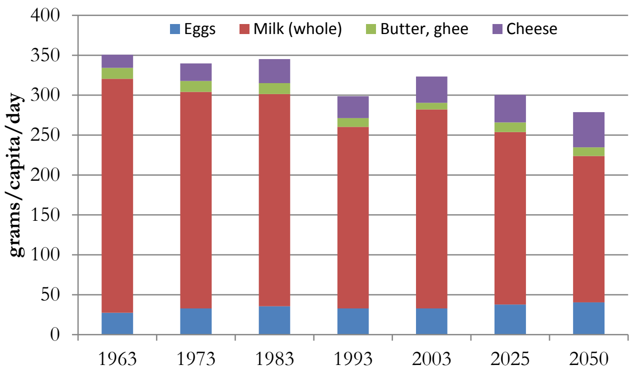
Figure 3: Trends in meat consumption in selected macro-areas and countries, disaggregated by meat types.
Source: Food and Agriculture Organization of the United Nations (2009) Food balance sheets: FAO, and Kearney (2010).



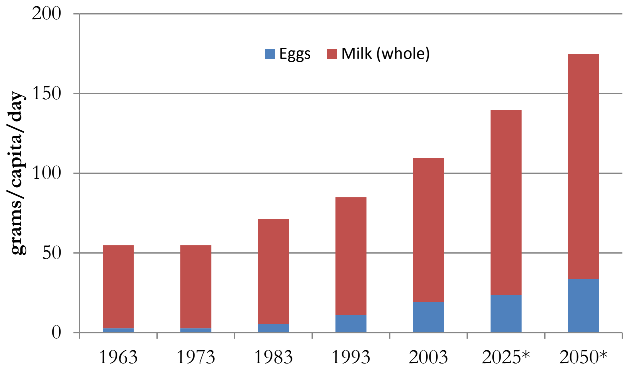
(a) *Developed countries*



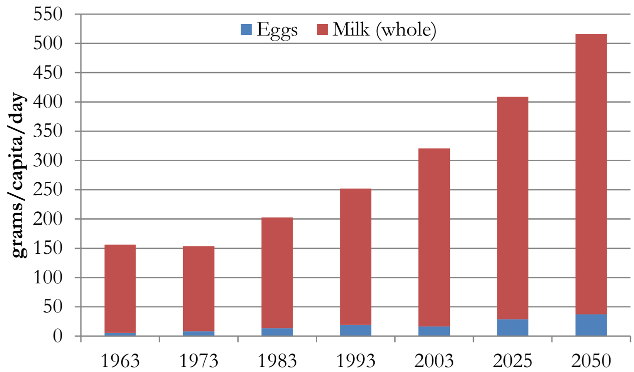
(b) *North America*



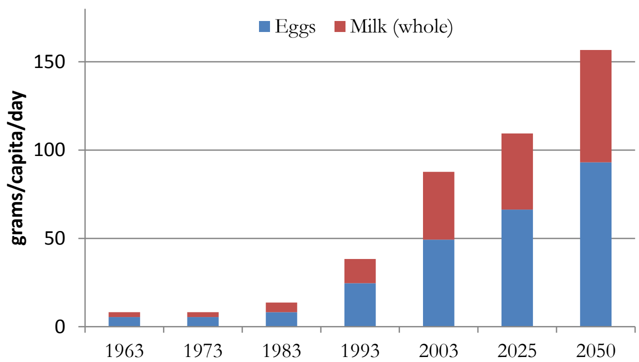
(c) *Europe*



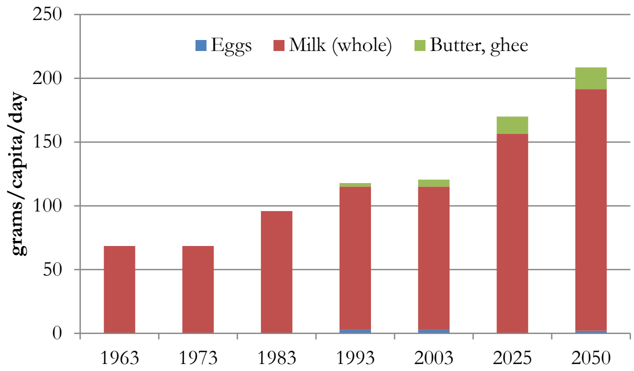
(d) *Developing countries*



(e) *Brasil*



(f) *China*



(g) *India*

Figure 4: Trends in eggs and dairy consumption in selected macro-areas and countries.
Source: Food and Agriculture Organization of the United Nations (2009) Food balance sheets: FAO, and Kearney (2010).

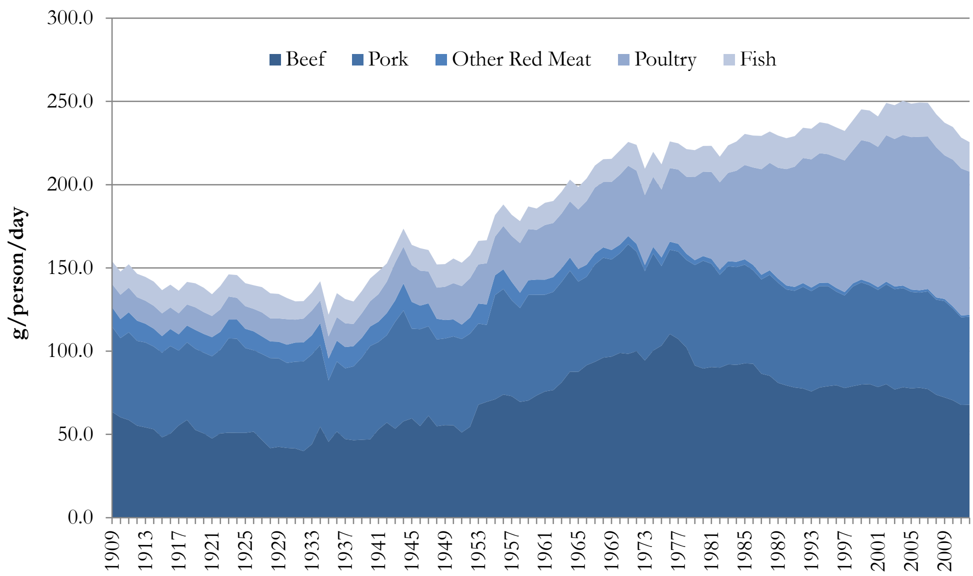


Figure 5: Per capita meat consumption in the US for the period 1909-2012, disaggregated by subgroups.
Source: USDA Economic Research Service (ESA) food availability documentation.

4 Data: out of the ivory tower, the ugly truth

4.1 What we want to do

As illustrated in the introduction, we are now interested in studying the drivers of the demand for animal-source foods among American residents, using survey micro-data, as advised by Hawkesworth et al. (2010). Once we manage to cast some light and estimate a model on the consumption of each food group, we will proceed to forecast the exogenous regressors to project the consumption towards 2020 and 2030. As we will see in what follows this will not be an easy task due to the less-than-ideal structure of the data. The food groups we are especially concerned with due to their impact on both human health and environmental sustainability are meat taken altogether, beef and dairy. Together with these ones, we are also analysing eggs, pork, poultry and fish, as they are the main substitutes for foods of the former types. In particular, poultry and fish are regarded as the healthy substitutes for red meat within the animal-origin food category (Hawkesworth et al., 2010).

4.2 What we ideally need

As Lin et al. (2003) concluded in their report, the trends for the consumption of animal-source food will continue to increase in the US in the near future, and while a main driver in this direction will be the growth in population – maybe slower than in the past, but still positive – the characteristics of the population and their distributions are very much likely to matter as well, especially in the choice of which foods to choose and how much of them to consume . The elements that influence dietary choices are many and varied, and probably not yet fully identified, as some of them might be unconscious, or instinctive, or hidden in deeply rooted habits. Arganini et al. (2012) reminded us that “a person does not necessarily have to be hungry to eat”, and even personal tastes have a low explicative power, as we hardly find ourselves eating our favourite dishes every day. In fact, nutrition is a complex component of human lives, and – especially once the threshold of subsistence is overstepped – many mechanisms beside the biological needs intervene in shaping it, ranging from social and cultural factors, to genetics, to economics. Ideally, we would therefore need data on all of the following characteristics to be able to have a comprehensive – although still probably incomplete – picture of the drivers of individual food choices:

- Economic: the main variables to be taken into consideration in this group are for sure per capita (real) disposable income – the main element of Engel curves – and prices of the good under consideration, and of its complements and substitutes, so to be able to estimate Marshallian demand functions and Slutsky equations and quantify the substitution effects. General equilibrium systems may also be built, combining both sides of supply and demand.

- Biological and physiological: characteristics such as age, gender, and health status are naturally associated with different biological needs. Children physiologically need less calories and proteins than adults. The same appears to hold for women, with Recommended Dietary Allowances (RDA) issued by the Institute of Medicine suggesting that they require about 20% less protein per day than men[7]. Moreover, age and gender have often be found to be linked to healthier diets (Verbeke, 2005; NU-AGE, 2012), with women in particular consuming systematically less red meat and animal-source foods in general, and to be more conscious of health consequences of nutrition (Lin et al., 2003). Nonetheless, many socio-cultural factors appear to influence the 'gender' gap in food consumption, not only physiological ones (Arganini et al., 2012; Prättälä et al., 2007).

- Socio-cultural: educated people are more likely to know more about nutrition and health (Lin et al., 2003; Daniel et al., 2011) although its impact on the quantity of – for instance – meat consumed is ambiguous, with some papers finding it not significant (Prättälä et al., 2007). Belonging to certain social-classes or ethnic groups may as well have an impact on nutrition, stimulating the diffusion of particular dietary patterns or fashions. The country and culture of origin is then necessarily linked with certain traditional and typical dishes, while religious beliefs may impose limitation on the consumption of certain food. Cultural diversity is also likely to play a part, allowing people to get into contact with new dishes and food traditions, and making exotic foods more easily accessible.

- Spatial and geographic: living in a rural rather than an urban area, or closeness to bodies of water, or to other particular geographic formations may favour self-production and availability of certain foods over others, as well as particular climate may be linked to preferences of certain foods over others.

- Seasonal: similarly, the season of the year may influence both availability and preferences (although thanks to the highly developed greenhouse agriculture and trade networks of the US all type of foods are likely to be available all over the year, without any significant seasonal constraint). Another 'seasonal' effect may be given by boom and bursts cycles in the economy, although it is a stylised fact in macroeconomics that consumption across macro categories hardly seem to change unless the crises drag on for an extended time span, preferences adjustment within each category may well happen. With particular regard to food, a traditional link seems to exist between times of crisis and restrictions on food consumption; think for instance to the every-day idioms still in use today 'to tighten one's belt', where curtailments in food is used to generally refer to contractions of consumptions in time of crisis.

- Ethical: personal beliefs and concerns over animal and environmental welfare, or even over labour conditions or political situations in production sites, may entice boycott and critical consumption.

- Food security and intra-household allocation: variables of this kind are more likely to matter in less developed countries, where transportation and refrigeration systems are still lacking in some areas, and poverty, subsistence agriculture, and periods of high inflation may constrain the choices of a significant portion of the population. Where food availability is constrained, allocation within the

[7] Source for Acceptable Macronutrient Distribution Range (AMDR) reference and RDAs: Institute of Medicine (IOM) Dietary Reference Intakes for Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein, and Amino Acids.

household does not always draw upon equal division. With respect to meat, the male members are usually given the lion share (Browning et al., 1994). Within this group, distribution may then be grounded on the consideration given to the members of the family, with the head of the household and the elderly being favoured, or on productivity concerns, meaning that the young and strongest members are preferred.

- Habits: finally, consumption naturally depends on what we are accustomed to eat and know how to prepare (Atkins et al., 2001), and it is not likely to change dramatically from one year to another. This means that the inclusion of lags for past consumption might have a strong explanatory power in modeling food consumption.

In order to disentangle the roles of these factors and control for fixed effects and individual specificities, the ideal structure we need for our data is a panel. In fact, panel data provide both cross-section and time-series components, allowing to get better insights than either of the two used solely. Unfortunately, as we are about to see, panel data for individual food consumption are incredibly rare and scarcely accessible, making our aim of getting projections harder to accomplish.

4.3 What we actually have

The main sources for individual micro-data on food consumption are national surveys carried out by countries. National surveys have the advantage of entailing a high level of details and distributional information which are collected with the aim of being representative of the overall population of the country. Nevertheless, implementation of micro-survey at such a scale is currently very costly and resource-demanding – and prohibitively so for many countries, especially less-developed ones. This means that the availability of data is dreadfully scarce, and even countries that do embark in such undertake, quite often only do it on a *una tantum* frequency, as it is the case for France's *Individuelle Nationale des Consomation Alimentaires* (INCA), Brasil's *Pesquisa de Orçamentos Familiares*, and many others. As if this was not bad enough, the resulting information are hardly comparable between countries – and sometimes even within a country, when responsibility for the survey move from an institution to another – due to lack of standardization in data collection. Anonymization techniques applied to prevent disclosure risks and identification of the participants – such as grouping the responses into classes and top-coding – add further noise to the analysis (cfr. Kearney, 2010). The best accessible databank that we could find among the countries we deemed of relevance for an analysis on food consumption was the US National Health and Nutrition Examination Survey (NHANES) conducted as a partnership between the U.S. Department of Health and Human Services (DHHS) and the U.S. Department of Agriculture (USDA). In particular, we are interested in the dietary intake interview section named 'What We Eat In America' (WWEIA), which has been released in two-year waves since 2001-2002, with the last available data being those for 2011-2012. A similar survey had also been carried out by USDA in the '90s under the name of Continuing Survey of Food Intakes by Individuals (CSFII), but the information collected and the methodology used differ slightly.

The NHANES is a multistage, representative survey with fixed sample-size targets for sampling domains defined by race, origin, sex, age, and low-income status, administered to a purposely selected sample of the US population, obtained through a four-stage probability sampling design (Curtin et al., 2013). In the first stratification, primary sampling units mostly equivalent to individual counties are drawn. Area segments comprising census blocks are then selected in the second stage, and households and dwelling units including also dormitories, etc. in the third stage. Within each household or dwelling unit, members were finally chosen based on gender, age, ethnic groups, and income in order to provide approximately self-weighting samples for each sub-domain while maximizing the number of interviewees per household. The final number of observations in each wave, for the variables we are interested in can be found among the summary statistics in the Appendix. Response attrition and over-sampling of groups of particular interest, such as minorities and elderly, make it necessary to devise a weighting scheme to insure the representativeness of the resident civilian non-institutionalized US population. Sample weights illustrating how many persons each respondent is 'representing' are therefore computed in a three-step procedure. The first-stage or base weight is given by the reciprocal of the sampled participant's probability of being drawn for the interview, i.e. the reciprocal of the sampling rate for the specific group they belong to. The objective of this first step is to counter-balance the differences in the probability of being selected due to the over-samplings. A second stage is then undertaken to account for non-responses and the specific characteristics that may have caused them. Variables most highly related to response and non-response propensity were identified by a Chi-squared Automatic Interaction Detector and adjustment factors computed as the reciprocals of the weighted response rate for the selected variables. Finally, weights were post-stratified using Census data with the aim of matching the US population totals (Mirel et al., 2013). Once the sample has been selected, data for the WWEIA section are collected in two separate 24-hour recall, the first conducted in person and the second by phone, on two different days of the week taken at random. A re-weighting for individuals that completed both interviews is provided. Importantly, the sample is re-drawn each year so that data do not have a panel structure, but are configured as repeated cross-section.

The average daily quantities consumed of meat, fish, dairy, eggs and meat subgroups' beef, pork and poultry are our target dependent variables. The NHANES datasets contain individual entry for each specific item eaten during the two-day survey[8]. After appending the first- and second-day datasets together, we therefore had to group the individual items in food groups, before summing up those in the same group and averaging them to obtain the daily average consumption we are interested in. Single items were codified according to the United States Department of Agriculture (USDA) "Food Code Numbers and the Food Coding Scheme" devised for the Food and Nutrient Database for Dietary Studies (FNDDS) (USDA). The first digit of this 8-digit code already identifies some major groups, namely: (1) milk and milk products; (2) meat, poultry, fish, and mixtures; (3) eggs; (4) legumes, nuts, and seeds; (5) grain products; (6) fruits; (7) vegetables; (8) fats, oils, and salad dressings; (9)

[8]In 2001-2002 the dietary survey focused on a single day, making the data for this wave not directly comparable with the rest, as the variance of the type of food and quantity consumed is much greater and less easy to link to our regressors of interest.

sugars, sweets, and beverages. We were interested only in the groups (1) to (3), i.e. the animal-source food groups. While the *dairy* and *eggs* groups were immediate to define, group (2) needed to be divided into *meat* and $fish$, and the former further into $beef$, *pork*, *poultry*, and other meat. Once obtained the average daily consumption within each food group for each individual, the dataset was appropriately merged with the demographic characteristics of the individual so to obtain our final dataset. Individuals that did not complete both days of the survey were discarded, and appropriately re-computed weights provided in the dataset were selected. The weights provided in the dietary dataset were preferred to those in the demographic dataset, so to be sure that the non-responses that abound in the dietary interviews are dealt with. The unit of analysis is the individual.

The main regressors we choose for our analysis are essentially two: income, which is the traditional component of Engel curves, and age. In fact, several studies have shown that food, as opposed to other macro-categories, seem to be a necessity good, and as such the elasticity of food consumption to income approximates zero, at least in developed economies. After a certain level of economic security has been reached, income is less likely to impact on the quantity of everyday food consumption, but rather on its quality (Manig and Moneta, 2014; Pingali, 2007). Yet, income may become a relevant variable in interaction with age, as it is not only a measure of purchasing power, but also an indicator for social classes and for the fashion trends and the concerns linked with them, which not only are likely to change with age, but also with time, becoming embroidered in habits. In turn, age is an interesting feature to take into account, as the amount of food consumed is in part physiologically linked to the development stage of the individual, and in part is again associated with dietary fashion and concerns in vogue among each age cohorts. For its own nature, the effect of age is likely to be non-linear, as we expect consumption to grow with age up to a certain point, and then decrease at older age (Verbeke, 2005; NU-AGE, 2012). This will be one of the reason leading to a semi-parametric specification of our reference models, rather than the more traditional parametric ones. Other socio-demographic, cultural and geographic variables are also added as controls, paralleling Lin et al. (2003) in as much as we could given the variables we could access to. The complete list of the variables used can be seen together with some summary statistics for the 2011-2012 wave in Table 1. Correlation coefficients for the same wave are presented in Table 2, while summary tables and correlation coefficients for other waves are given in the Appendix. Details for each variable used as independent regressors is given in what follows.

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
meat	7486	304938168	170.204	147.8465	0	1381.13
fish	7486	304938168	22.66933	62.45235	0	1043.925
dairy	7486	304938168	251.9536	259.697	0	3111
eggs	7486	304938168	24.27473	44.53427	0	700
beef	7486	304938168	58.00561	101.322	0	1381.13
pork	7486	304938168	15.23487	43.18309	0	747.76
poultry	7486	304938168	65.81226	95.39284	0	931.915
pcincome	7194	297082322	16964.13	19316.85	139.8148	187565.6
age	7486	304938168	37.43204	22.1845	0	80
gender	7486	304938168	0.5092513	0.4999478	0	1
hhsize	7486	304938168	3.357985	1.608195	1	7
black	7486	304938168	0.1246058	0.3302934	0	1
hispanic	7486	304938168	0.1666005	0.3726437	0	1
bornus	7486	304938168	0.8556331	0.3514849	0	1
hrgender	7486	304938168	0.451686	0.4976935	0	1
hrbornus	7486	304938168	0.7889985	0.4080467	0	1
winter	7486	304938168	0.4512302	0.4976491	0	1

Table 1: Summary statistics for 2011-2012 data.

	meat	fish	dairy	eggs	beef	pork	poultry	pcincome	age	gender	hhsize	black	hispanic	bornus	hrgender	hrbornus	winter
meat	1																
fish	-0.0491	1															
dairy	-0.0945	-0.0501	1														
eggs	0.0546	0.0118	-0.0234	1													
beef	0.6425	-0.0246	-0.0516	0.0256	1												
pork	0.2792	-0.0097	-0.0373	0.0587	-0.0032	1											
poultry	0.5731	-0.0281	-0.0536	0.0039	-0.0629	-0.0093	1										
pcincome	-0.0353	0.0371	-0.0196	-0.0054	-0.0566	-0.0195	0.0176	1									
age	0.0729	0.1201	-0.2004	0.0779	0.0544	0.0552	0.0132	0.1523	1								
gender	-0.1947	-0.0294	-0.0971	-0.0753	-0.1268	-0.0803	-0.0549	-0.0136	0.0463	1							
hhsize	-0.0175	-0.0334	0.0913	-0.0274	-0.0155	-0.0142	-0.009	-0.2379	-0.4825	-0.0514	1						
black	0.0477	0.0268	-0.1298	-0.0001	-0.0357	-0.0032	0.0748	-0.122	-0.065	0.0367	0.0461	1					
hispanic	-0.0098	-0.009	0.0176	0.0531	0.0168	-0.0573	0.0227	-0.1729	-0.1641	-0.0195	0.2135	-0.1645	1				
bornus	-0.0322	-0.0718	0.0516	-0.0335	-0.0311	-0.0131	-0.0549	0.0826	-0.0454	0.0011	-0.1047	0.0798	-0.374	1			
hrgender	-0.0696	-0.0097	-0.0339	-0.0385	-0.0517	-0.0289	-0.0492	-0.1073	-0.0344	0.2381	-0.0219	0.1214	0.0019	0.0481	1		
hrbornus	0.0116	-0.0402	0.004	-0.0303	0.0041	0.0348	-0.0463	0.1097	0.0973	0.0084	-0.2017	0.0744	-0.4502	0.6619	0.0751	1	
winter	0.0506	0.0103	-0.0197	0.0165	0.0464	-0.0099	0.0417	-0.1102	-0.0769	-0.019	0.0586	0.0188	0.1914	-0.0538	-0.0094	-0.0641	1

Table 2: Correlation table for 2011-2012 data.

The economic indicator we use (*pcincome*) is given by the overall annual income imputed to the household to which the individual belongs, adjusted for the household size. In fact, the survey asked interviewees for the income of the family[9]. From it, a derived variable was obtained estimating the total household income. Due to disclosure risks and confidentiality concerns, this value is not released as such, but is censored and re-coded into the following range values:

- **Class1:** 0 to 4,999 USD
- **Class2:** 5,000 USD to 9,999 USD
- **Class3:** 10,000 USD to 14,999 USD
- **Class4:** 15,000 USD to 19,999 USD
- **Class5:** 20,000 USD to 24,999 USD
- **Class6:** 25,000 USD to 34,999 USD
- **Class7:** 35,000 USD to 44,999 USD
- **Class8:** 45,000 USD to 54,999 USD
- **Class9:** 55,000 USD to 64,999 USD
- **Class10:** 65,000 USD to 74,999 USD
- **Class11:** 75,000 USD and Over
- **Class12:** 20,000 USD and Over
- **Class13:** Under 20,000 USD
- **Class14:** 75,000 USD to 99,999 USD
- **Class15:** 100,000 USD and Over

Classes 12 and 13 were included in order to provide for interviewees that were unable or unwilling to report greater detail. Up to the 2005-2006 wave, only the first thirteen classes were used, and income were truncated at 75,000 USD. Due to the growth in income and to the large portion of individuals consequently included in the uppermost category, class 11 has been replaced by classes 14 and 15 since 2007, and truncation moved forward to 100,000 USD. In order to make the variable usable for our purpose, we had to make some preliminary adjustments. Data including a refused, 'don't know', or missing answer for the income are assumed to be missing at random, due to the devices and techniques to account for different rates of responsiveness among sample groups, and therefore ignored. As they represent less than 5% of the total dataset, representativeness is assumed not to be significantly affected[10]. We then had

[9]The Census Bureau defines the term 'family' for use in the Current Population Survey (CPS)as "a group of two people or more (one of whom is the householder) related by birth, marriage, or adoption and residing together; all such people (including related subfamily members) are considered as members of one family" (http://www.census.gov/cps/about/ cpsdef.html). This definition is the one used by NHANES to distinguish members of different families within the same household.

[10]As a rule-of-thumb in survey analysis, weights for representativeness need not be re-evaluated if the missing values are less than 10%(Johnson et al., 2013).

	Mean value by class					
Class	2001	2003	2005	2007	2009	2011
0-$4,999	1538	1404	1306	1350	1173	1271
$5,000-$9,999	7672	7720	7702	7807	7911	7939
$10,000-$14,999	12371	12353	12405	12433	12384	12414
$15,000-$19,999	17270	17304	17321	17192	17340	17331
$20,000-$24,999	22280	22294	22217	22199	22242	22225
$25,000-$34,999	29662	29639	29576	29508	29578	29609
$35,000-$44,999	39531	39587	39511	39530	39544	39499
$45,000-$54,999	49682	49625	49564	49540	49592	49603
$55,000-$64,999	59633	59605	59598	59512	59616	59565
$65,000-$74,999	69638	69569	69574	69605	69594	69409
$75,000-$99,999	-	-	-	86053	85988	85995

Table 3: Mean values by income class, up to the last closed class.
Source: Current Population Survey (CPS), Annual Social and Economic Supplement, US Census Bureau.

to deal with the problem of the open classes 12 and 13. Ignoring the data was not an option, firstly because these classes included a non-negligible number of individuals, and secondly because removing such an amount of data would require adjustments in the weights to insure representativeness. Instead, we used an imputation technique commonly applied to missing data (Kalton, 1983): individuals in class 12 and 13 were allocated at random respectively in the classes 5 to 15 and 1 to 4, with the probability to be allocated in each class proportional to the number of people already in that class, so to preserve relative frequency of the incomes. Again, we chose to assume that missingness of more specific details on the income occurred at random and was not correlated to individuals characteristics, thanks to the devices and techniques put in place to account for different rates of responsiveness among sample groups. Next, in order to get to an approximately continuous distribution of the income, we need to transform the classes back into point values. The usual solution in these cases is to replace the classes with their midpoint value. The problem with this approach is that it assumes that the distribution of the variable is uniform within each class, therefore ignoring the actual intra-class distribution and providing a staircase overall density function. To fix this problem and make the overall distribution as representative as possible, we chose to determine the mean-value for each class using data about USA household income from the National Census[11], for each of the year under consideration, as shown in Table 3. Finally, we are left to deal with data in the last income class. The simpler way would be again to assign to the individuals in the last class the mean-value computed from the National Census statistics. Yet, this would not help with the problem of the right-censoring of the distribution. Instead, we borrow the uppermost subclasses for which the US Census Bureau provides relative frequencies and mean values. We then impute the individuals proportionally to the various class and then assign to each subclass its mean-value. In this way we have

[11]Current Population Survey (CPS), Annual Social and Economic Supplement, US Census Bureau, various years. In particular, the US Census Bureau provide for each year the frequency, mean value and standard error of the annual household income of the American population for a wide array of classes and demographic characteristics.

managed to smooth the right tail of the income distribution. The result of the imputations for 2011-2012 can be seen in Figure 6. Similar results were obtained for the other waves. An alternative approach to impute the censored class, as well

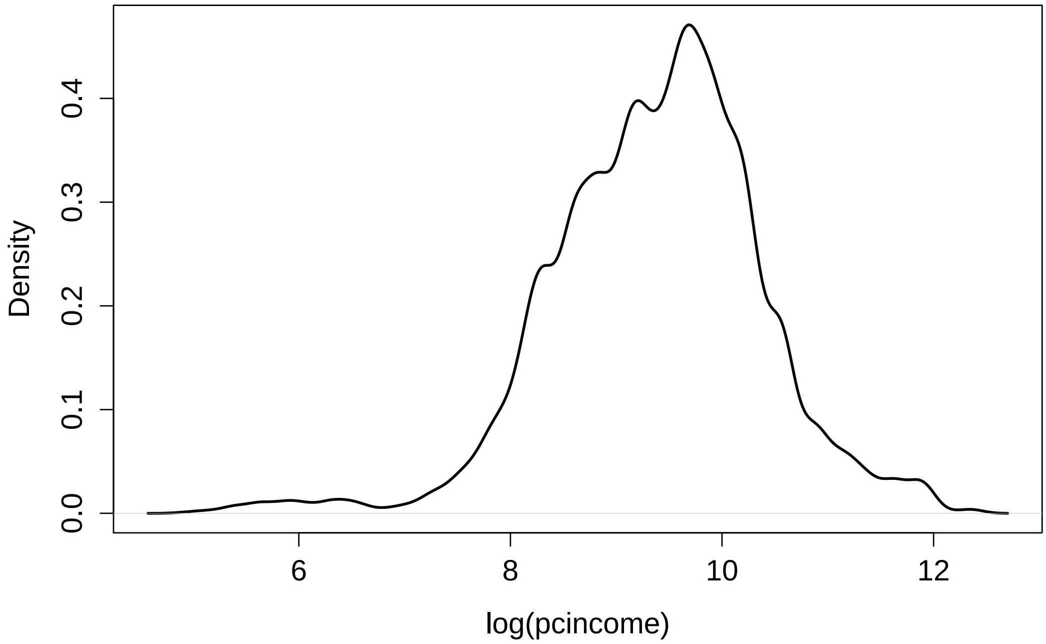


Figure 6: Density function of the derived per capita income variable for survey wave 2011-2012.

as classes 12 and 13, is to model the income for the remaining individuals using other characteristics provided for in the dataset. This strategy is nonetheless hardly applicable to our case, especially when taking into consideration that the income for the remaining individuals is imputed as the mean value of the class, and as such there is no intra-class variability on which to base the estimation, and even more importantly the income is a household-level variable, while most of the other characteristics at our disposal are at the individual-level, so that the correlation levels are too low to devise a fitting model. In order to adjust the overall income of the household for the household size and composition, various scale of adult-equivalency have been proposed in the literature (Atkinson et al., 1995). In particular, the OECD firstly devised an equivalence scale that assigned a value of 1 to the first household member, of 0.7 to each additional adult and of 0.5 to each child[12], then in the late 1990s changed the equivalence scale to 0.5 for additional adult members and 0.3 for children, as suggested by Hagenaars et al. (1996). This ‘OECD-modified scale’ was also adopted by the Statistical Office of the European Union (EUROSTAT). Although we do have data for the number of children and adults in each household in 2011-12, this is not true for previous years, when these pieces of information were not yet collected and only the household size is available. In order to make the variable homogenous across years, we therefore have to use a simplified version of the OECD adult equivalency scaled, where a value of 1 is assigned to the head of the household

[12]This scale is also called ‘Oxford scale’.

and 0.5 is instead given to all the other members.

$$adultequiv = 1 + (hhsize - 1) * 0.5$$

The individual disposable income is therefore finally computed as:

$$pcincome = hhincome/adultequiv$$

As we need to compare data for the different waves, we further proceed to deflate the income variable according to the US Consumer Price Index, as provided by the US Bureau of Labor Statistics[13], obtaining our final variable. Boxplots for the consumption of meat in different income classes (defined on a logarithmic scale) are shown in Figure 7. Data for other years give similar results.

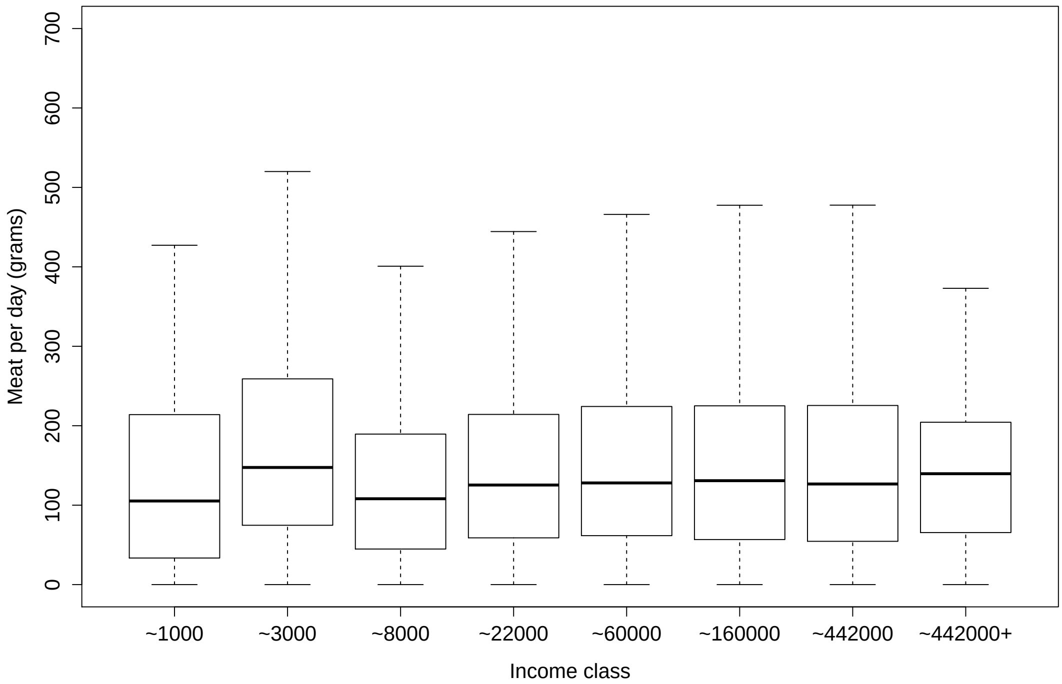


Figure 7: Boxplot of the consumption of meat by income class.

The variable *age* is given in years at the time of the initial screening interview for the survey. A top-code is adopted for disclosure risk at 80. For our analysis we restricted the dataset to the adult population, 16 years old and older, as younger individuals are unlikely to take their own decisions in what and how much food to buy or eat, and their responses to the survey were not self-reported but given by an older member of the family. Age classes used for the boxplot in Figure 8 follows those suggested by Johnson et al. (2013). Consistently with our hypothesis, age appears to have a non-linear effect on meat consumption, with a peak at 20-39 for men and at 40-59 for women

[13]Bureau of Labor Statistics. Series id: CUSR0000SA0. US city average, all items. Base period: 1982-84=100. http://www.bls.gov/cpi/data.htm.

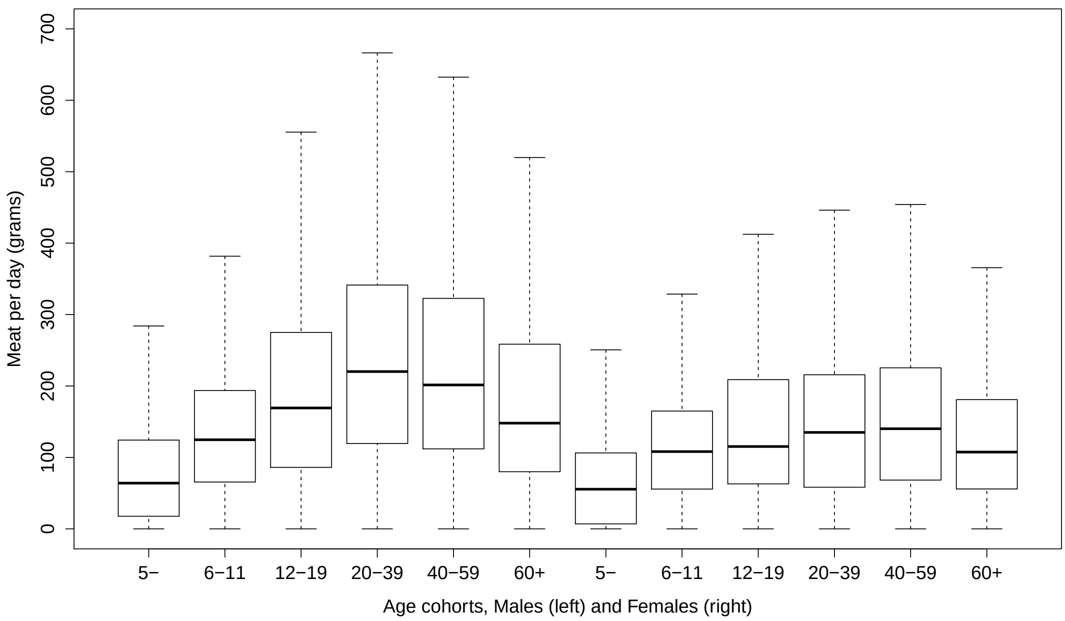


Figure 8: Boxplot of the consumption of meat by age class and gender.

Among the control variables included in the model, the *gender* dummy assumes a value of 1 for women and 0 for men. Consistently with the literature and with the physiological dimension, we expect women to consume less of most foods with respect to men. *hhsize* gives the total number of household members, from 1 up to 7, with top-code applied to preserve anonymity of the participants. We then used the variable on self-identified ethnic group to build the dummies *black* and *hispanic* to account for culturally-based dietary patterns. We could not build a similar variable for Asians as this characteristic was only collected starting from 2011. Nevertheless it would be a very interesting component to take into consideration, given the recent waves of first-generation migrants from China and South-East Asia, which are less likely to have already moulded their dietary habits to those prevailing in the USA. Another control linked to the habit dimension is the dummy *bornus*, assuming value 1 for individuals born in one of the 50 States or Washington DC and 0 for who is instead born abroad. As the traditional American diet is known to be one of the richest in meat and especially beef, we expect this variable to play a positive role in the consumption of these foods. Two other controls are added to take into account the role of the reference person in the household, who is likely to be the one doing or at least deciding the shopping and in some cases, cooking. *hrgender* controls for the gender (1 female, 0 male), while *hrbornus* for the origin (1 if born in one of the 50 States or Washington, DC, 0 if born abroad) of the reference person. Finally, *winter* is a seasonal dummy provided in the NHANES dataset, with value 1 for interviews taken in the colder months, November 1st through April 30th, and value 0 for those taking place between May 1st and October 31st.

Unfortunately, the dataset does not provide any geographical reference to account for the effect of living in different US regions, or in the countryside rather than in an un urban area. This is certainly a limitation on the analysis, as these variables appear to have a strong effect on dietary choice in the literature (see for instance Lin et al., 2003). Another notable no-show which geographic dummies are often used to control for, are relative prices. In fact, since expenditures and prices are not reported in the surveys, we are bound to assume that relative prices remain constant not only during the survey period but also over the whole projections period (cfr. Lin et al., 2003). Nevertheless, as the US are one of the biggest producer of meat and related animal-source foods in the world, the food groups we are analysing are highly available and easily affordable, and their prices do not fluctuate much[14]. The lack of sufficient relative price variation which is often recorded in advanced Western economies, would be particularly problematic with semi-parametric estimation (Pesaran and Schmidt, 1995, p.178). Finally, as we are considering quite broad categories of foods, including goods of different quality, and therefore not homogenous in price, we take prices as fixed. In fact, we hypothesize that price affect more the choice of the quality rather than the quantity of food consumed (Manig and Moneta, 2014), although verifying it is outside the scope of this work. Anyway, for interpretations purpose, similarly to the rest of the Western countries, fish is on average the most expensive among the foods we chose, followed in the order by beef, pork and poultry. As well as poultry, another relatively cheap food is eggs, while dairy are quite heterogenous in price, including very cheap products such as milk and butter, and more expensive cheese[15]. To conclude, a severe limitation is given by the absence of time lags of consumption, which prevent us from accounting for the habit component of dietary choice. Hopefully, with future release and changes in the sampling and data collection procedure, these drawbacks will be overcome. In particular, we hope that a panel-data structure will be adopted at least for part of the surveyed sample in the future, in order to capture the time trends and allow to disentangle whether differences in consumption linked to age are actually due to habits or to the physiological aspects of growing up and ageing. In fact, the lack of a panel-data structure, forces us to make the implicit assumption that when an individual moves from one age cohort to another their preferences immediately change accordingly. Data on prices and expenditure should also be collected together with quantity so to better complement the analysis and control for quality of the food. Finally, a standardization of the methodology involved in ascertaining food intake is advisable, in order to make it easier and possible to perform international comparisons (Kearney, 2010).

[14] See Consumer Price Index and Average Price Data from the U.S. Bureau of Labor Statistics http://www.bls.gov/data/.

[15] See Consumer Price Index and Average Price Data from the U.S. Bureau of Labor Statistics http://www.bls.gov/data/.

5 Model and estimates: let the data speak

5.1 Identifying the regular consumers

With the purpose of reducing the noise in the data and better isolate the drivers of animal-source food consumption, we only estimate the model in the next section for the individuals with a strictly positive consumed quantity of that specific food group. In fact, we observed that in so doing the portion of deviances explained improved substantially. Although individuals with zero and positive consumption can easily be identified in the starting datasets, in order to reproduce this truncation in the projected datasets, we need to model the choice of individuals to either enter or not enter in the category of the regular eaters[16] of each type of food. Following a study on a similar dataset for China, carried out by Du et al. (2004), we assume that this choice does not occur at random, but it is made more or less likely due to certain individuals' characteristics, which need not be the same that will then drive the choice of the amount to consume. The preliminary step of the model is therefore to estimate the following probit, for each food group (meat, fish, dairy, eggs, beef, pork, poultry), using data for each individual (l):

$$
\begin{aligned}
Pr\{dummy_l = 1|X_l\} = \Phi(\alpha X_l) = &\\
&= \Phi(\alpha_1 + \alpha_2 \log(pcincome_l) + \alpha_3 ridageyr_l + \\
&+ \alpha_4 gender_l + \alpha_5 hhsize_l + \alpha_6 black_l + \\
&+ \alpha_7 hispanic_l + \alpha_8 bornus_l + \alpha_9 hrgender_l + \\
&+ \alpha_{10} hrbornus_l + \alpha_{11} winter_l + \eta_l
\end{aligned}
$$

$\Phi(.)$ is the cumulative distribution function for the standard normal, i.e.:

$$\Phi(z) = \int_{-\inf}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Estimates are obtained by maximum likelihood estimation techniques, using an outlier-robust log-likelihood (Toomet and Henningsen, 2008).

5.2 The Generalized Additive Model

Our main model is a modification of the traditional cross-section Engel curves, which study the relationship between income and consumption, augmented to take into account the age structure of the population and the non-linearities that such structure and the distribution of income may entail in the demand for certain foods. For this and other reasons that will be delved into more depth in what follows, we decided to adopt a data-driven approach and added a non-parametric component in the specification of the model. It is to be noted that the first application of non-parametric techniques to the estimation of Engel curves, dates back to Ernst Engel himself, who in his 1857's seminal work used a data-fitting technique very similar

[16] Given that the survey concerned the dietary choice of two random days only, the absence of consumption cannot be directly interpreted as e.g. being 'vegetarians'. Similarly, the individuals that are found consuming a positive daily amount, may well not consume that much every day. The underlying idea of the survey is that these two biases cancel out on average. For this reason we prefer to distinguish the individuals into regular and non-regular consumers of each food.

to what is modernly known as a regressogram[17]. In the core mainstream economics that followed, Engel curves were mainly estimated parametrically, imposing models derived from micro-economic theories on the data. A non-negligible problem with this approach is that such parametric modeling is particularly prone to misspecification errors, and subsequent biases in the resulting estimates. On the contrary, non-parametric methods use an unspecified function to substitute for fixed coefficients and specified functional forms, asymptotically removing specification errors, similarly to what is done with Fourier expansions. While this works fine with univariate regressions, the so-called 'curse of dimensionality' arises when more than one regressor is considered – as in our case. In fact, with multiple explicative variables, the number of parameters that need to be estimated in non-parametric models according to the asymptotic theory associated with Fourier series increases dramatically fast (Pesaran and Schmidt, 1995). For this reason, semi-parametric models combining parametric and non-parametric components are often preferred in econometrics when many regressors are involved (Cameron and Trivedi, 2005). Semi-parametric models have the important advantage to be able to capture non-linearities and not force coefficients such as elasticity of demand to stay fixed, but allow them to change for different segment of the population under analysis. As a consequence, semi-parametric models have found a natural application in the estimation of consumer demand and Engel curves for those goods that present more variety of curvature than traditional parametric modeling would allow for, as noted by Blundell et al. (1998). In particular, age and other demographic variables are likely to introduce additional sources of non-linearity when included in demand models, and we therefore chose, following Gozalo (1997) and Blundell et al. (1998), to adopt this approach to obtain the estimates we need. Nevertheless, a caveat of non-parametric and semi-parametric estimation to be aware of is the over-fitting risk, which we have tried to control expanding the number of observations by pooling datasets for years that showed similar results, as explained below. Within the family of semi-parametric models, a convenient form for the purpose of this work is the flexible and widely used Generalized Additive Model (GAM), originally developed by Hastie and Tibshirani (1986, 1987) and for which an established methodology of estimate exists. The GAM is essentially a generalization of the generalized linear model (GLM). Contrarily to non-parametric models where a completely unspecified function is used with regressors as arguments, the GAM assumes additivity and separability of the predictors, and assign a different smooth function g_j to each covariate. Smooth functions are then estimated using what they call a *local scoring algorithm*, iteratively applying a scatterplot smoother. Theoretically, the procedure is nothing else than an empirical method to minimize the Kullback-Leibler distance to the true model, or – equivalently – maximizing the *expected* likelihood (Hastie and Tibshirani, 1986). In practice, with regards to the issue of over-fitting of the estimates that would arise if absolutely any smooth function was allowed, a penalization is applied in the maximum likelihood procedure, based on the penalized regression splines, as described in Wood (2006). Since many of our controls are dummies rather than continuous variables, we do not turn to smooth functions for them, but include them in a conventional parametric component, as

[17]See Chai and Moneta (2010) for a review on Engel's original works and its legacy up to the present day and Engel and Kneip (1996) for an overview on the application of non-parametric statistics to Engel curves.

shown in Equation 1:

$$E[y|X_1, X_2] = c + \beta X_1 + \sum_j g_j(x_{2j}) \tag{1}$$

where βX_1 represents the parametric component for control variables, and $g_j(.)$ the smooth functions, i.e. the non-parametric components, with the main regressors as input. This variant of the original Hastie and Tibshirani's GAM is generally referred to as 'partial generalized additive model', or simply as 'generalized additive model' (cfr. Cameron and Trivedi, 2005; Wood, 2006). As mentioned above, the GAMs have the crucial advantage to be able to deal with highly non-linear and non-monotonic effects of the predictors, and, as well as being an excellent exploratory tool, they also provides direct data driven interpretations and predictions, with the advantage *"of being completely automatic, i.e., no 'detective work' is needed on the part of the statistician"* as Hastie and Tibshirani (1986) themselves stated.

We try two specifications of the model, given by Equations 2 and 3 below, both of which give very similar results. Each model is estimated separately for each food type: meat, fish, dairy, eggs, beef, pork, poultry. As underlined in the introduction, our measure for consumption is quantity (grams per day) and not expenditure as in the traditional Engel curves. This is because our final aim is to obtain projections of the quantity demanded and consumed for health and environmental considerations. Since we are interested in the effects elicited by the age structure, the unit of analysis we choose is the individual, l:

$$\begin{aligned} c_l = {} & \beta_1 + g_1(\log(pcincome_l)) + g_2(ridageyr_l) + \\ & + \beta_2 gender_l + \beta_3 hhsize_l + \beta_4 black_l + \\ & + \beta_5 hispanic_l + \beta_6 bornus_l + \beta_7 hrgender_l + \\ & + \beta_8 hrbornus_l + \beta_9 winter_l + \epsilon_l \end{aligned} \tag{2}$$

As an exploratory step we run regression separately for each year. We noticed similarities in the significance of variables and in the structure of income and age elasticities across the waves 1994-1996, 2003-2004, 2005-2006, and 2011-2012 confirming our hypothesis that nutritional choices concerning animal-source foods in a developed country such as the US are quite stable in the medium term, as opposed to the nutrition transition in the developing world. In all these cases income does not seem to play any significant role – at least when no interaction with other variables is considered – although there are some evidence of a middle-class effect, while demographic and cultural variables such as age, gender, and ethnicity are highly significant. Remarkably, the effect of age does show non-linearities, as we foretold. The situation changes for the waves that include periods classified as recessions by the Business Cycle Dating Committee of the National Bureau of Economic Research (NBER), that is 2001, and 2007-09. In these times of crisis, income becomes a significant variables in the choice of how much meat to consume. As hypothesized in the previous section, this may be a form of reaction that dates back to the old time, rather than or as well as a symptom of actual economic struggle for food. In order to constrain the over-fitting problem of non-parametric modeling, we pool together the various wave for the period 2003-2012[18]. In fact, data collection only became systematic in

[18]Pooling is suggested by CDC to increase reliability (Curtin et al., 2013).

Food type	Meat	Fish	Dairy	Eggs	Beef	Pork	Poultry
Actual number	13467	3868	13013	5291	7650	4420	8460
Fitted number	13509	3924	12935	5411	7725	4485	8557
Discrepancy	0.003	0.014	-0.006	0.023	0.010	0.014	0.012

Table 4: Comparison between the true number of individuals with consumption dummy equal 1 and the number obtained by probit estimation.

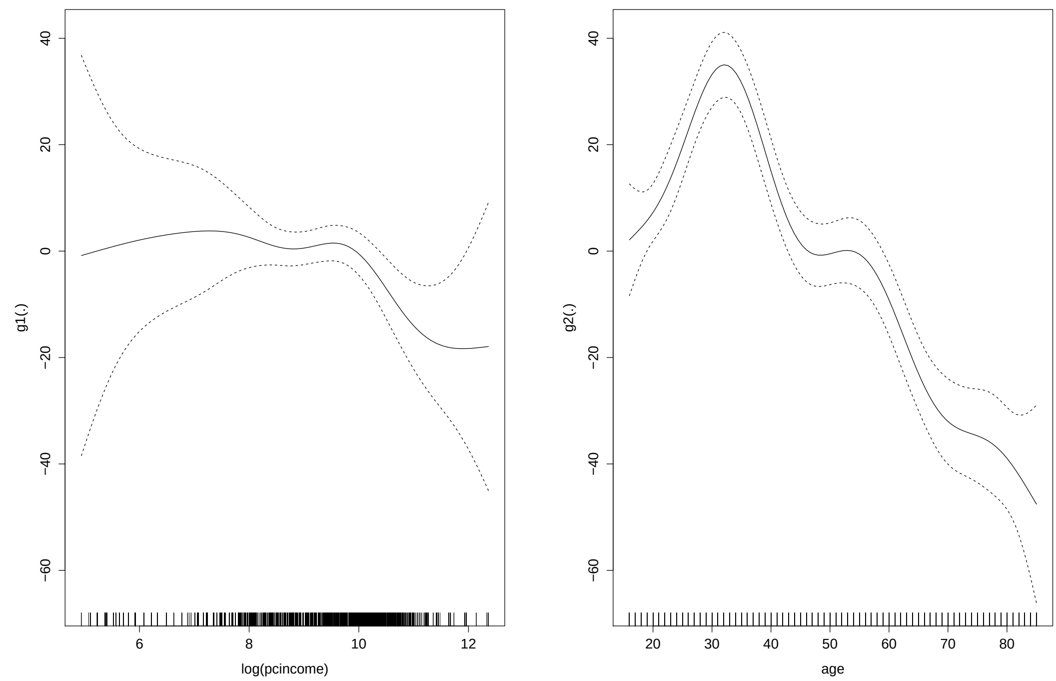
2003, while previous waves entailed slightly different methodologies, so that direct comparison between data collected in the 1994-96 and in the 2001-02 wave and data collected after 2003 is not advisable. A categorical variable is added to account for the different waves. Our observation on the difference between booms and bursts of the economy is confirmed, with the effects of the dummies for 2005-06 and 2011-12 being not significantly different from that of 2003-04, as opposed to the 2007-08 and 2009-10 in our main dataset, i.e. that for meat (Table 5). For other food types this difference is less marked, but since we suspect that recession periods may entail a structural change in the model and add further noise into the estimates due to fluctuations in income, we chose to remove the years of the crisis in all the pooled datasets. The estimates obtained using these datasets are those that will then be used for forecasting. Although income does not seem to be significant per se for some of the food groups, we suspect that it may have some interaction effects with the age variable. We therefore change the model into:

$$\begin{aligned} c_l = &\beta_1 + g(\log(pcincome_l), ridageyr_l) + \beta_2 gender_l + \\ &+\beta_3 hhsize_l + \beta_4 black_l + \beta_5 hispanic_l + \beta_6 bornus_l + \\ &+\beta_7 hrgender_l + \beta_8 hrbornus_l + \beta_9 winter_l \end{aligned} \tag{3}$$

The explained deviance improve slightly with respect to the previous specification, and the Generalised Cross Validation score and the Akaike Information Criterion are smaller for all the food types except for fish, for which we therefore stick to the previous specification.

5.3 Estimates for the various food groups

First of all, we want to know how good the selection procedure based on the probit model are in identifying who are the relevant consumers within the whole dataset. We compare the number of fitted response equal to 1 from the probit model specified in Equation1 with the number of actual observations whose dummy is 1 (Table 4). For all the food groups, the discrepancy between the two value is at or less than 1%, with the only exception of eggs, where the discrepancy is around 2%. We therefore deem the fit of the model good enough for our purpose. We can now move on to look at which variables are significant in the GAMs models and what effects they have, *ceteris paribus*. Table 5 shows the estimates of the two model specifications for *meat*, as well as of the preliminary model for the pooled dataset with recession years. As predicted, the gender dummy has a negative effect, with women eating on average 74 grams of meat less than men per day, consistent with findings in Arganini et al. (2012); Prättälä et al. (2007). A similar effect is also linked to the gender of the reference person in the household: if female, household members consume 6 grams of meat less per day each. Being Hispanic does not seem to have any role, while Afro-Americans eat on average 13 grams of meat more than non-black per day. Interviews taken from November to April were associated with a consumption of meat on average 8 grams greater than those taken in the warmer months. Age is also highly significant and its effect follows a non-linear and non-monotonic trend, with consumption of meat increasing with age for the youth, until it reaches its peak at about 30 years of age, and then decreasing, so that 50-year-olds consume on average as much meat as people in their 20s, and older people consuming even less (Figure 9, right). Income is only significant at the 90% significance level, and it appears to have a threshold effect: its role is rather flat and heterogenous up to a per capita income of 20,000 USD, after which consumption starts decreasing (Figure 9, left). Overall, it therefore seems that meat is a food for young adults and poor-to-middle-class people. This interpretation is confirmed by the second specification of the model, where the interaction term is again highly significant with a peak for individuals in their 30s and with a income of less than 500 USD, decreasing progressively for people between 25 and 40 with an income below 20,000USD. The very young, independently of their income, appear to eat an average level of meat. Among the elderly, income appears to matter more, with consumption decreasing faster with respect to age for the richer: a rich 45-year-old appears in fact to consume as much as poorer 55-year-olds, and the very wealthy over 70 years of age consuming less than anyone else (Figure 10). Size of the



(a) (b)

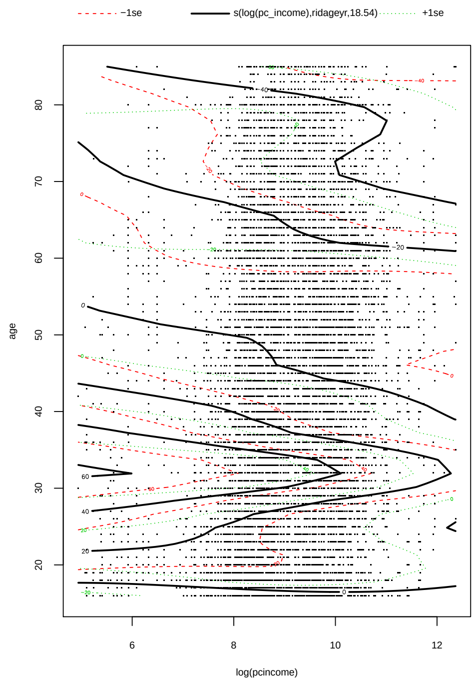
Figure 9: Smooth curves for the effect of income (left) and age (right) on meat consumption.

household, belonging to the Hispanic ethnic group, being born in the US or having the reference person in the household born in the US does not seem to affect the consumption of meat. The most significant variables are therefore the demographic ones (age and gender), together with a cultural effect specifically linked to being Afro-Americans and a seasonal effect linked to the time of the year. The income elasticity

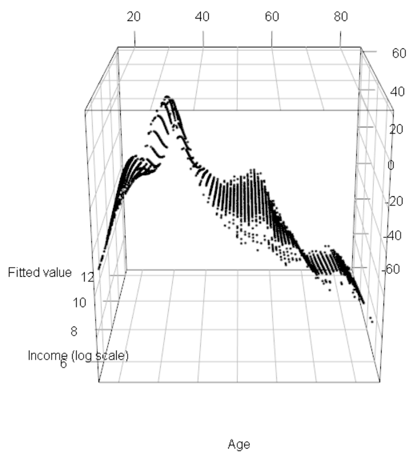
	Meat 0	Meat 1	Meat 2
(Intercept)	231.34***	229.32***	230.02***
	(4.82)	(5.69)	(5.68)
gender	−77.65***	−74.07***	−74.00***
	(2.03)	(2.59)	(2.59)
hhsize	−1.22	0.33	0.13
	(0.75)	(0.96)	(0.95)
black	7.89*	13.48***	12.98**
	(3.14)	(3.96)	(3.96)
hispanic	−0.30	−0.14	−0.46
	(3.57)	(4.59)	(4.57)
bornus	3.94	2.07	1.49
	(4.53)	(5.79)	(5.79)
hrgender	−4.29*	−5.26*	−5.50*
	(2.07)	(2.66)	(2.66)
hrbornus	−2.17	−2.23	−1.60
	(4.18)	(5.25)	(5.25)
winter	6.90***	8.35**	8.40**
	(2.03)	(2.56)	(2.56)
year_caty2005-2006	4.39		
	(3.13)		
year_caty2007-2008	10.30***		
	(3.13)		
year_caty2009-2010	11.97***		
	(3.14)		
year_caty2011-2012	4.79		
	(3.11)		
EDF: s(log(pcincome))	6.79**	4.56	
	(7.88)	(5.62)	
EDF: s(age)	6.91***	7.38***	
	(7.99)	(8.36)	
EDF: s(log(pcincome),age)			18.54***
			(23.11)
AIC	310682.42	179713.43	179699.30
Log Likelihood	-155313.51	-89834.77	-89821.12
Deviance explained	8.55%	8.91%	9.1%
GCV score	199207353.46	320144536.69	319809359.68
Num. obs.	23244	13467	13467

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 5: Comparison of model specifications for meat consumption.



(a)



(b)

Figure 10: Smooth curve for the effect of the interaction term given by income and age on meat consumption: contour plot (left) and 3-dimensional plot (right).

of the demand for meat was finally computed by differencing on the fitted values for the first specification of the model, using both meat and income in logarithmic scale. In fact, another of the main advantages of a semi-parametric approach is that it doesn't force the elasticity to stay fixed, but allows it to change for different segment of the population. As it can be noted from Figure 11, meat is regarded as a normal food or necessity, since its elasticity is nowhere significantly different from zero. Most of the other food types present very similar income elasticity, so we will not present them. The scarce relevance of income for many food groups is consistent with the coefficients found by Lin et al. (2003), using a Tobit model. Overall, the effects we found are in line with those presented by Daniel et al. (2011). Focusing

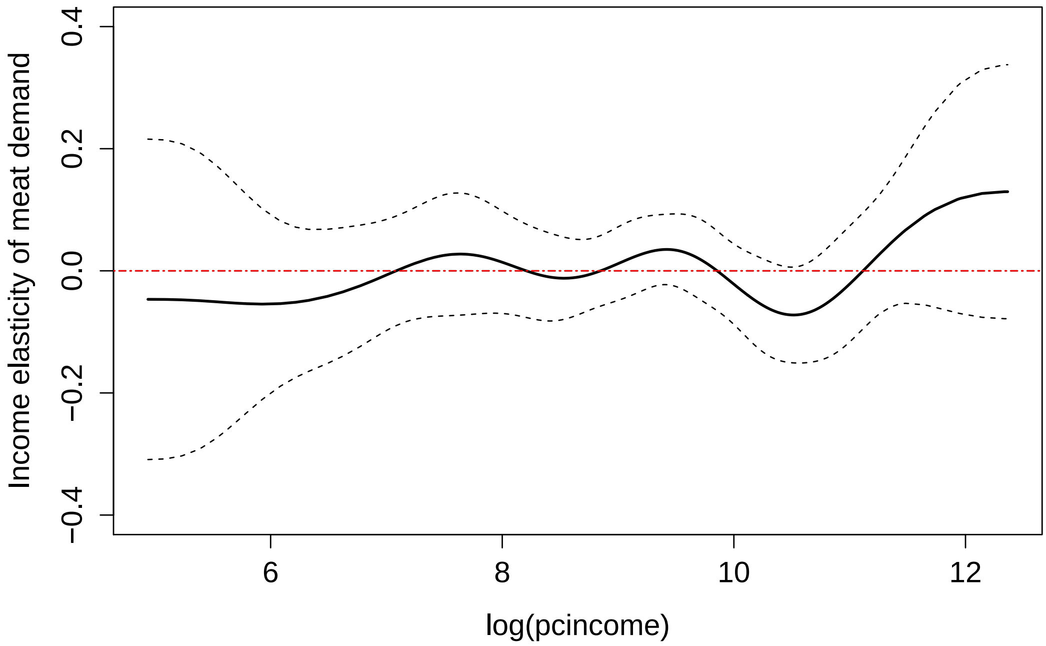
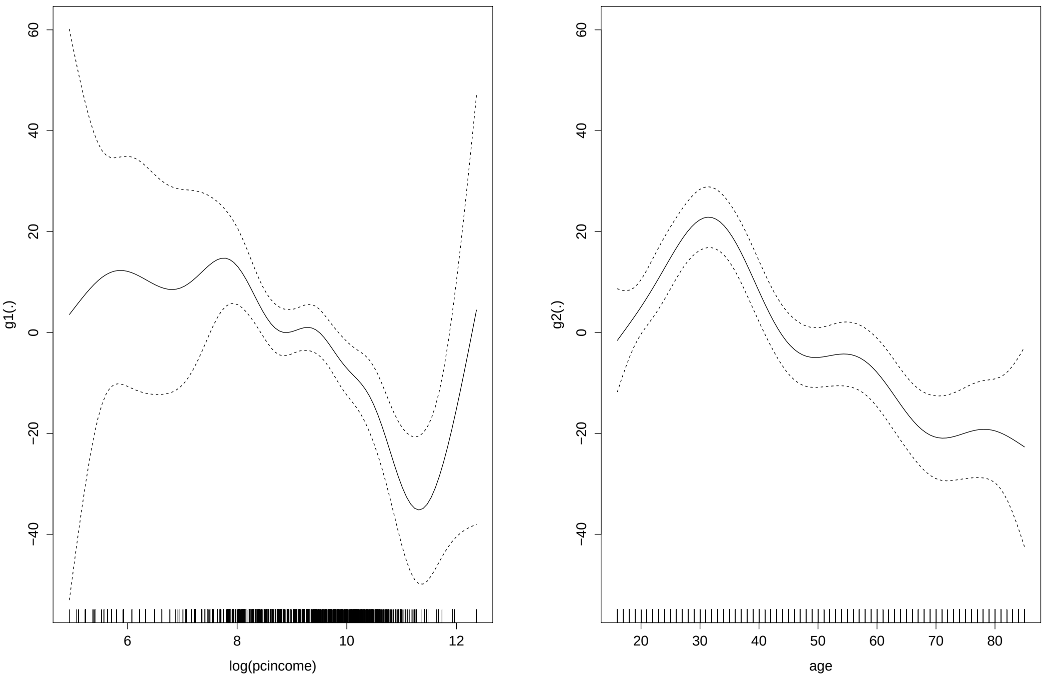


Figure 11: Structure of meat elasticity curves with respect to income.

on the meat subgroups (see Table 6), we find again a negative effect of gender on *beef* consumption, with women eating on average 41 grams less per day than men. Being black this time has a negative effect, with Afro-Americans eating about 17 grams less per day than non-black, while belonging to the Hispanic ethnic group is linked on average with a consumption of beef 12 grams greater than non-Hispanic. Smaller households also appear to consume more beef, with any additional member reducing by 2 grams the individual consumption. No seasonal effect appears to be linked with the consumption of beef, nor characteristics of the household reference person or the origins of the interviewee. Income and age are highly significant, both taken together and separately. Similarly to meat, the peak in beef consumption is reached by people between 30 and 35 years old, and then decreases progressively as the individual gets older (Figure 12, right). Income has an ambiguous effect up to 3,000 USD, then the relationship between income and consumption becomes a negative one, with individuals decreasing dramatically their consumption of beef as they get wealthier. The direction reverses for individuals with income over 60,000 USD, when consumption starts growing again (Figure 12, left). When the two are taken together, as in Figure 13, it can be seen that the peak in consumption is reached by the poorer share of the population between the age of 20 and 40. Up to about 35 years of age, consumption of beef seems to increase with age but decrease with income; after that age both income and age have a negative effect.



(a) (b)

Figure 12: Smooth curves for the effect of income (left) and age (right) on beef consumption.

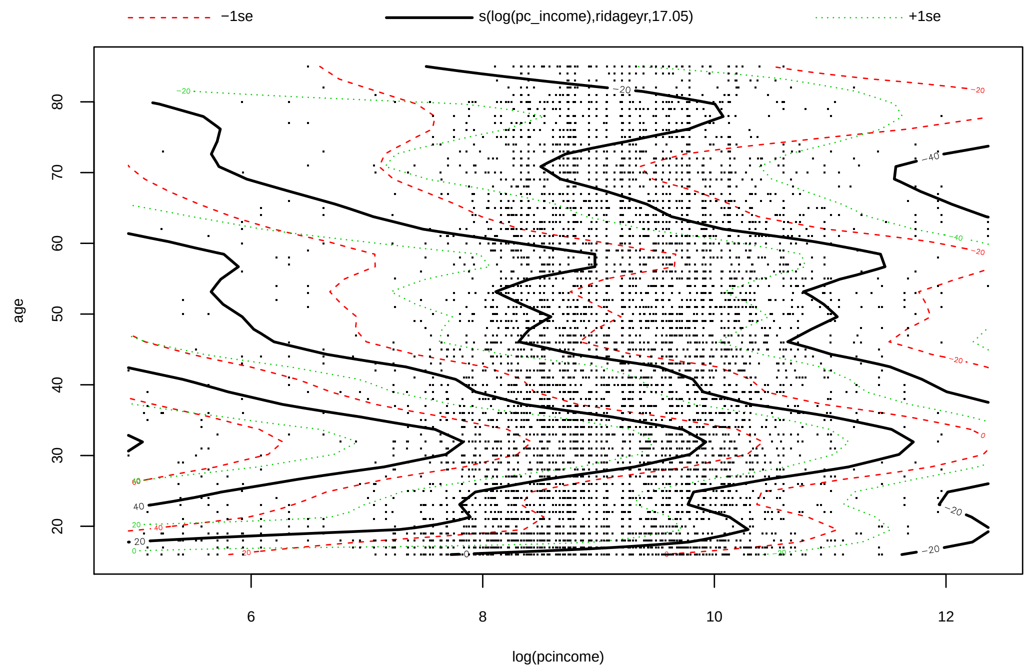


Figure 13: Smooth curve for the effect of the interaction term given by income and age on beef consumption: contour plot.

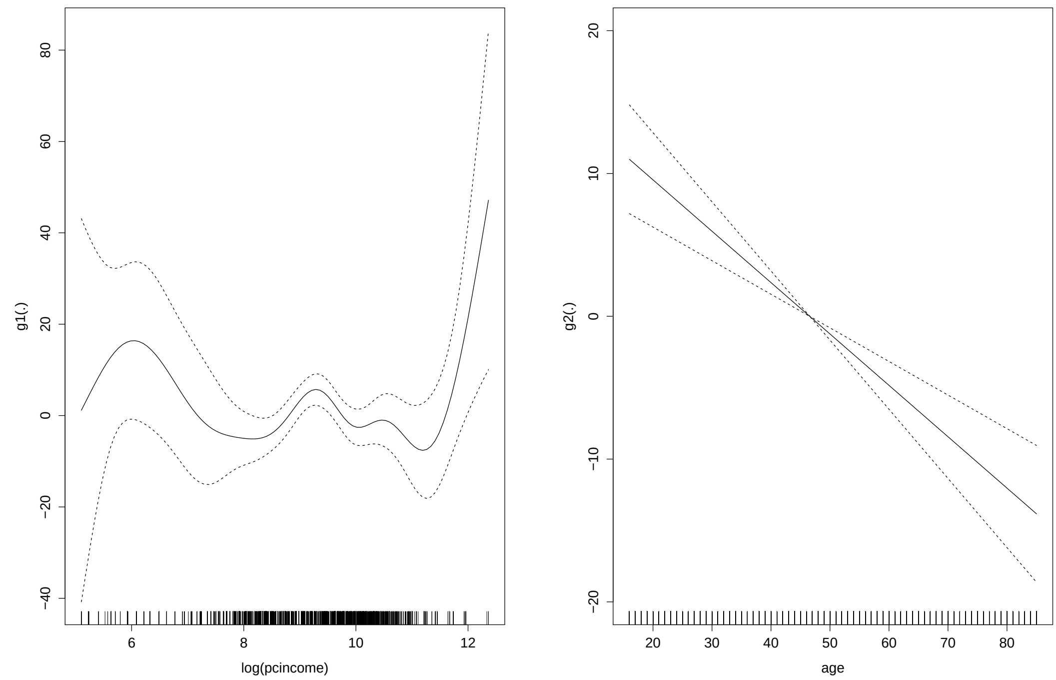
	Beef 1	Beef 2	Pork 1	Pork 2	Poultry1	Poultry2
(Intercept)	154.09^{***}	154.30^{***}	88.22^{***}	89.51^{***}	144.35^{***}	144.86^{***}
	(6.16)	(6.15)	(4.66)	(4.70)	(4.97)	(5.03)
gender	-40.69^{***}	-40.65^{***}	-21.46^{***}	-20.77^{***}	-27.40^{***}	-27.52^{***}
	(2.75)	(2.75)	(2.05)	(2.05)	(2.33)	(2.34)
hhsize	-2.41^{*}	-2.39^{*}	0.69	0.31	-1.65^{*}	-1.76^{*}
	(1.02)	(1.01)	(0.77)	(0.78)	(0.83)	(0.84)
black	-16.98^{***}	-17.12^{***}	-6.86^{*}	-6.72^{*}	5.92	5.73
	(4.34)	(4.33)	(2.98)	(2.98)	(3.31)	(3.31)
hispanic	11.55^{*}	11.53^{*}	-16.74^{***}	-17.22^{***}	-5.09	-5.24
	(4.88)	(4.88)	(3.76)	(3.77)	(4.00)	(4.00)
bornus	2.62	2.35	-19.04^{***}	-18.44^{***}	-10.45^{*}	-10.79^{*}
	(6.50)	(6.49)	(4.86)	(4.86)	(5.06)	(5.06)
hrgender	-2.92	-3.14	0.36	0.12	-4.13	-4.09
	(2.85)	(2.85)	(2.10)	(2.10)	(2.40)	(2.40)
hrbornus	-2.79	-2.98	-5.59	-5.89	0.53	0.64
	(5.80)	(5.79)	(4.35)	(4.35)	(4.59)	(4.59)
winter	2.25	2.33	-3.76	-3.30	6.50^{**}	6.55^{**}
	(2.73)	(2.73)	(1.99)	(2.00)	(2.27)	(2.28)
EDF: s(log(pcincome))	7.55^{***}		8.01^{**}		1.16	
	(8.46)		(8.73)		(1.31)	
EDF: s(age)	6.50^{***}		1.00^{***}		4.57^{***}	
	(7.63)		(1.00)		(5.61)	
EDF: s(log(pcincome),age)		17.05^{***}		16.16^{***}		16.75^{***}
		(21.58)		(20.67)		(21.28)
AIC	98806.00	98804.54	51766.17	51765.82	107220.12	107210.61
Log Likelihood	-49378.95	-49375.22	-25864.07	-25856.75	-53594.33	-53578.55
Deviance explained	5.36%	5.45%	5.48%	5.79%	3.65%	4.01%
GCV score	208392955.34	208353734.66	65507106.77	65503061.02	157709767.52	157533509.42
Num. obs.	7650	7650	4420	4420	8460	8460

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 6: Comparison of model specifications for beef, pork and poultry consumption.

Gender is again linked with a reduced intake of *pork* meat, although this time the effect is smaller, with 21 grams less per day. Black and Hispanic individuals eat 7 and 17 grams less pork per day respectively. Being born in the US rather than being an immigrant to the country also has a negative effect, with a gap of 19 grams less per day. Similarly to beef, no seasonal effect is apparent, as well as no effect for household size and characteristics of the household reference person. Income and age are both highly significant. Income appears to have a rather flat effect, causing a peak in consumption for people of the middle-class (Figure 14, left). The effect of age on the contrary has a much more marked and linear negative effect (Figure 14, right). The overall peak seems again to be for the very young and poor-to-middle classers, while the very young and very old with highest incomes are those who consume the least.

Given these results, *poultry* must be the meat subgroup that drive the significance



(a) .

(b) .

Figure 14: Smooth curves for the effect income (left) and age (right) on pork consumption.

of the seasonal effect while reducing the significance of income, maybe due to its high affordability as cheapest type of meat on average. While income appears to be irrelevant to the choice of how much poultry to eat, the consumption seem to increase by around 7 grams in the colder part of the year. The negative effect of gender is again confirmed, with women eating on average 27 grams of poultry less than men per day, and household with a female reference person consuming 4 grams less per member than other households, although the latter effect is only significant at the 90% confidence level. As for beef, smaller household consume more poultry per capita, with each additional member to the family reducing the per capita consumption by 2 grams. Ethnicity is scarcely relevant, with Afro-Americans consuming about 6 grams more (at a 90% confidence level) and Hispanics presenting no significant difference at all with the rest of the population. Similarly to pork, being born in the US appear to be linked to a smaller consumption of poultry of about 10 grams, somehow a proof

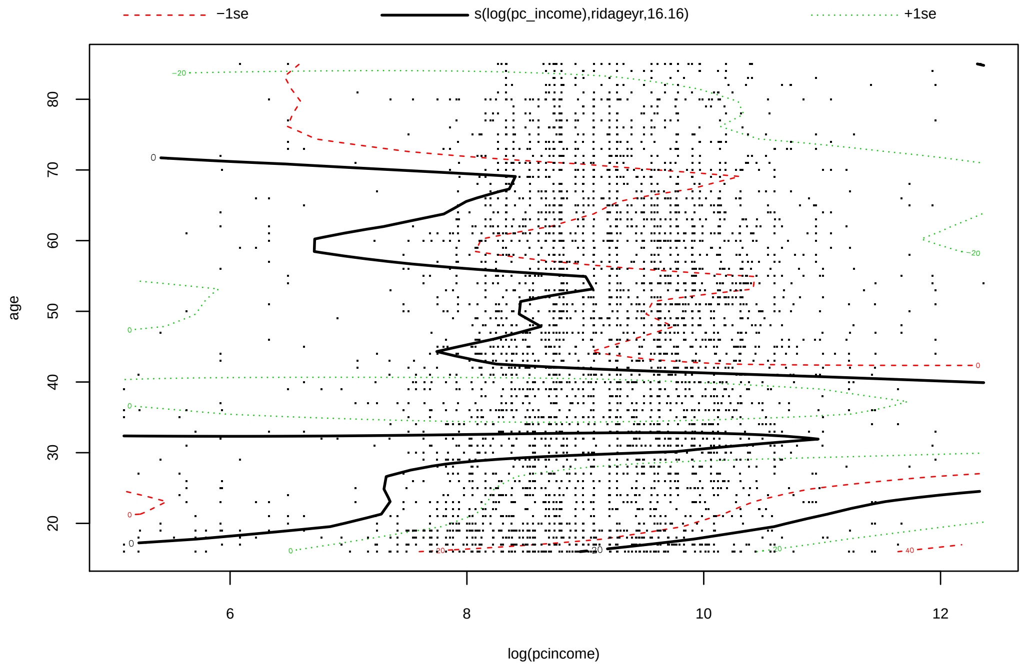
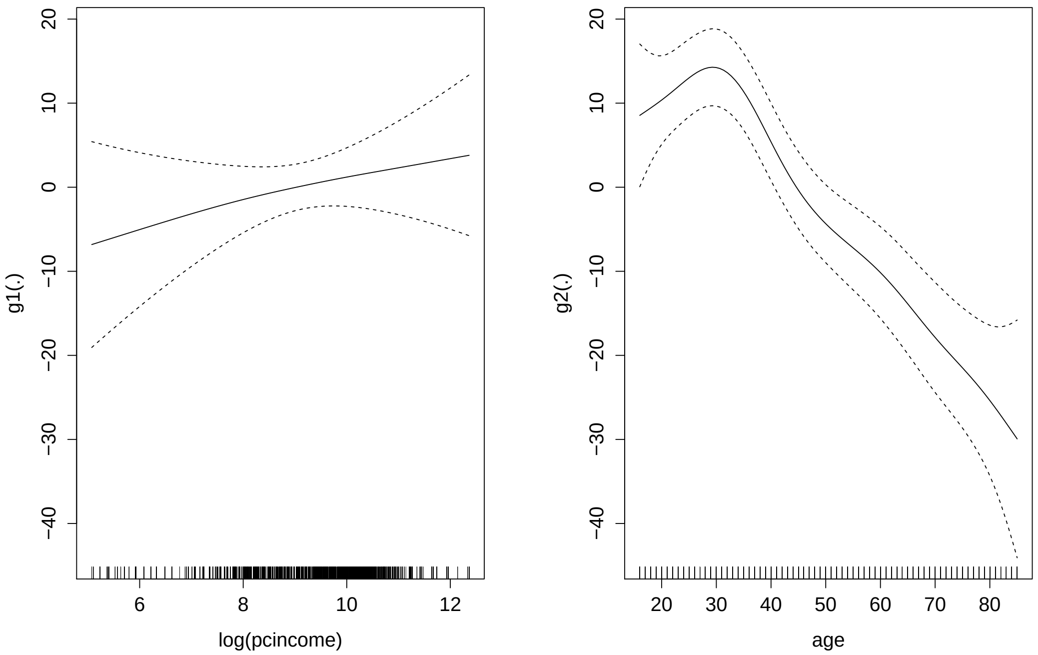


Figure 15: Smooth curve for the effect of the interaction term given by income and age on pork consumption: contour plot.

that the favourite meat type of *born-and-bred Americans* is beef, the most dangerous food at both individual health and environmental level. As reported above, income does not appear to be significant, although it may have a slight positive effect, while age is highly so. After peaking for individuals in their 20s, age appears to have a strong negative effect thereafter (Figure 16, right). While this may partly be due to the fact that elderly people consume less meat overall, for mid-aged people this is not true, signaling that new generations may be more sensitive to health concerns and prefer to substitute poultry for red meat. This is confirmed by the contour plot for the second model, where young adults appear to be the greatest consumer of poultry, notwithstanding their income, while among older generations income seems to have a positive effect. Again, this may signal a substitution effect due to health concerns among wealthier people, that are traditionally more educated and more concerned about healthy diets.



(a) (b)

Figure 16: Smooth curves for the effect of income (left) and age (right) on poultry consumption.

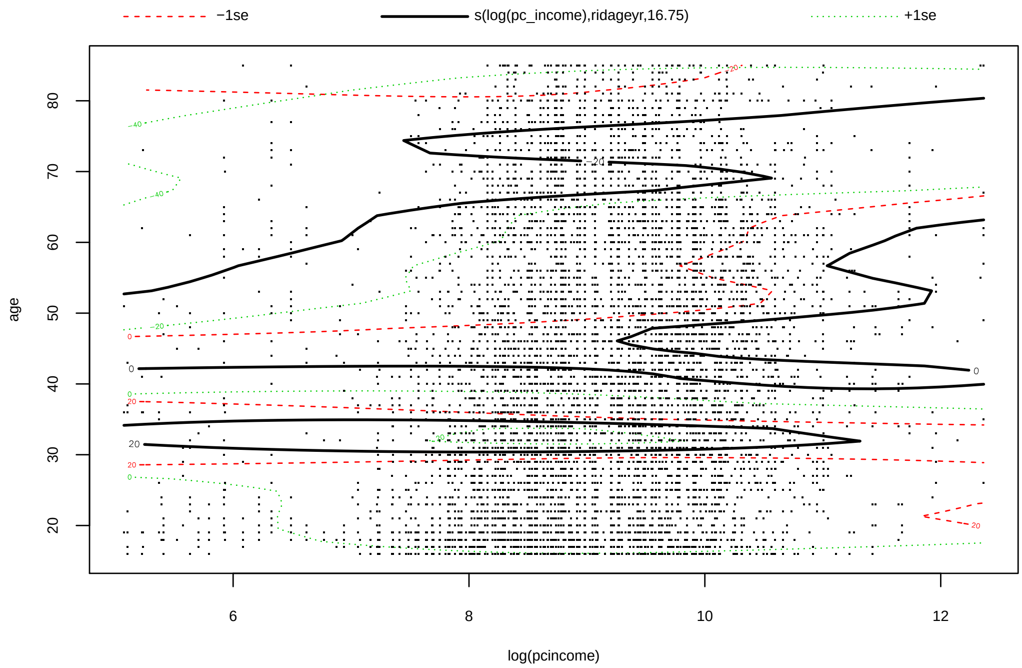


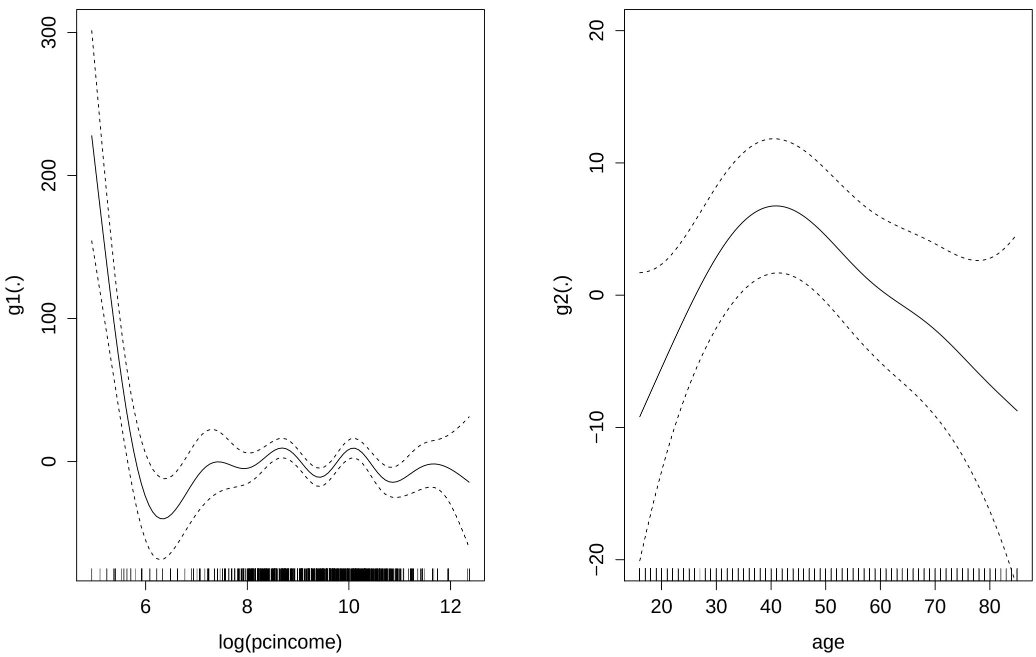
Figure 17: Smooth curve for the effect of the interaction term given by income and age on poultry consumption: contour plot.

	Fish 1	Fish 2	Dairy 1	Dairy 2	Eggs 3	Eggs 4
(Intercept)	102.77^{***}	102.96^{***}	311.25^{***}	312.28^{***}	81.08^{***}	80.70^{***}
	(6.13)	(6.14)	(10.67)	(10.65)	(3.14)	(3.13)
gender	-31.80^{***}	-31.76^{***}	-55.85^{***}	-55.65^{***}	-19.11^{***}	-19.15^{***}
	(3.03)	(3.04)	(4.85)	(4.84)	(1.42)	(1.42)
hhsize	1.40	1.61	-1.84	-2.27	-0.40	-0.39
	(1.14)	(1.13)	(1.81)	(1.80)	(0.53)	(0.52)
black	3.46	3.02	-109.95^{***}	-109.19^{***}	-3.21	-2.98
	(4.51)	(4.53)	(7.81)	(7.79)	(2.12)	(2.11)
hispanic	8.17	9.88	-32.44^{***}	-33.70^{***}	4.74	4.99^{*}
	(5.40)	(5.42)	(8.62)	(8.58)	(2.42)	(2.42)
bornus	-5.36	-5.67	1.05	0.43	-0.30	-0.64
	(6.16)	(6.19)	(10.74)	(10.73)	(3.12)	(3.11)
hrgender	2.78	2.62	17.26^{***}	16.82^{***}	-3.12^{*}	-3.06^{*}
	(3.12)	(3.14)	(4.98)	(4.98)	(1.45)	(1.45)
hrbornus	7.67	7.85	7.84	8.44	2.06	2.08
	(5.78)	(5.81)	(9.75)	(9.74)	(2.81)	(2.81)
winter	3.53	2.45	-29.41^{***}	-28.79^{***}	-0.60	-0.41
	(3.05)	(3.07)	(4.79)	(4.79)	(1.40)	(1.40)
EDF: s(log(pcincome))	8.79^{***}		6.50		5.52^{***}	
	(8.99)		(7.63)		(6.68)	
EDF: s(age)	3.03^{*}		8.52^{***}		7.74^{***}	
	(3.79)		(8.93)		(8.59)	
EDF: s(log(pcincome),age)		4.89		25.58^{***}		18.52^{***}
		(6.31)		(28.21)		(23.09)
AIC	47891.71	47938.89	189563.61	189527.29	59236.83	59225.62
Log Likelihood	-23924.03	-23954.55	-94756.79	-94728.07	-29595.16	-29584.29
Deviance explained	5.40%	3.90%	3.78%4	4.21%	6.32%	6.70%
GCV score	125101087.18	126634249.41	1098379011.93	1095321377.88	36893823.88	36816104.28
Num. obs.	3868	3868	13013	13013	5291	5291

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 7: Comparison of model specifications for fish, dairy and eggs consumption.

We finally analyse the results for the last set of models, presented in Table 7. Only two variables appear to substantially drive the choice of how much *fish* to consume, and contrarily to meat, one of this is income, sustaining the common idea that fish is rather a food for rich people than for the average population. It should also be noted that the number of observations for fish are much less than for the other food types, as not so many people eat fish regularly. The second relevant variable is gender, which consistently to the previous analyses shows a negative effect, with women consuming on average 32 grams of fish less per day than men. Age also appear to have some influence on fish consumption, although a smaller one, causing a peak in consumption for middle-age people. This time, the explained deviance of the interaction term between age and income is smaller than in the model specification where the two are separately considered, so that the latter will be our favourite model for fish. While being a woman is linked to a reduction in *dairy* consumption of 56



(a) (b)

Figure 18: Smooth curves for the effect of income (left) and age (right) on fish consumption.

grams, having a women as reference person for the household leads to an increase of 17 grams. Black and Hispanic people consume again much less than the rest of the population, with 110 and 32 grams less respectively. As for dairy, a seasonal effect is present, but is now of opposite sign, with interviews taken in the colder months being linked to 29 grams less of dairy consumed. Age is another highly significant variable, both when taken individually and when coupled with income. The different peaks in consumption that can be seen in Figure 19, right, seem to correspond to different income levels. In fact, in Figure 20 it can be seen that among the middle-to-high incomes, the very young are the greatest consumers, while among the less wealthy, consumption seems to peak for individuals in their 40s. Income per se, as well as household size and origin of the interviewee or of the household chief, is not significant. With respect to *eggs* consumption, the only relevant control variables

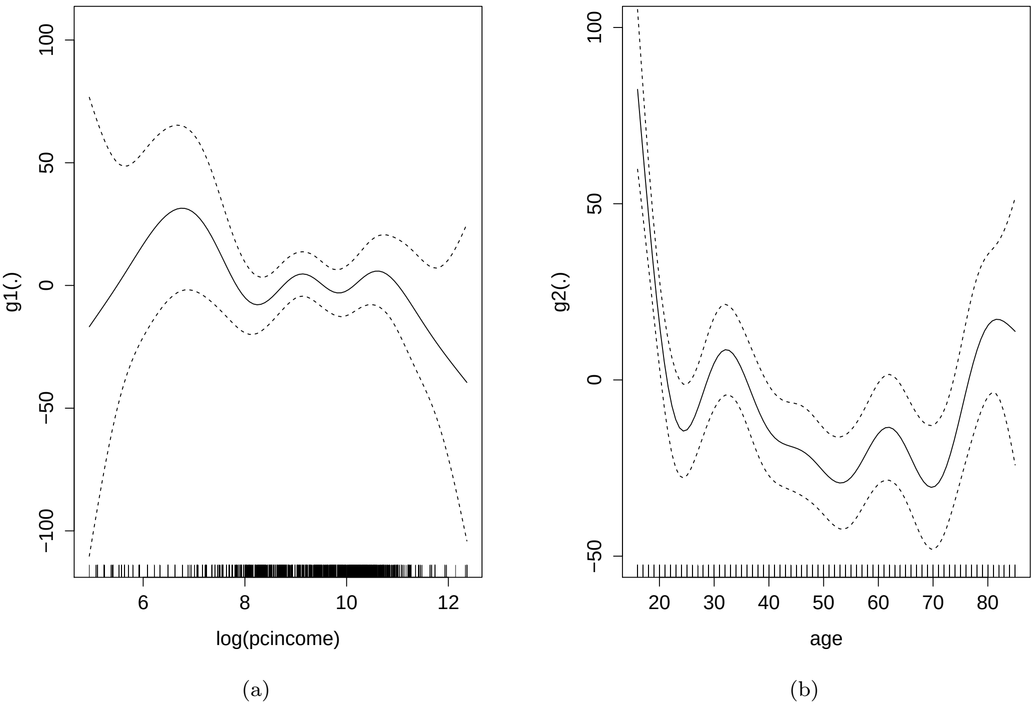


Figure 19: Smooth curves for the effect of income (left) and age (right) on dairy consumption.

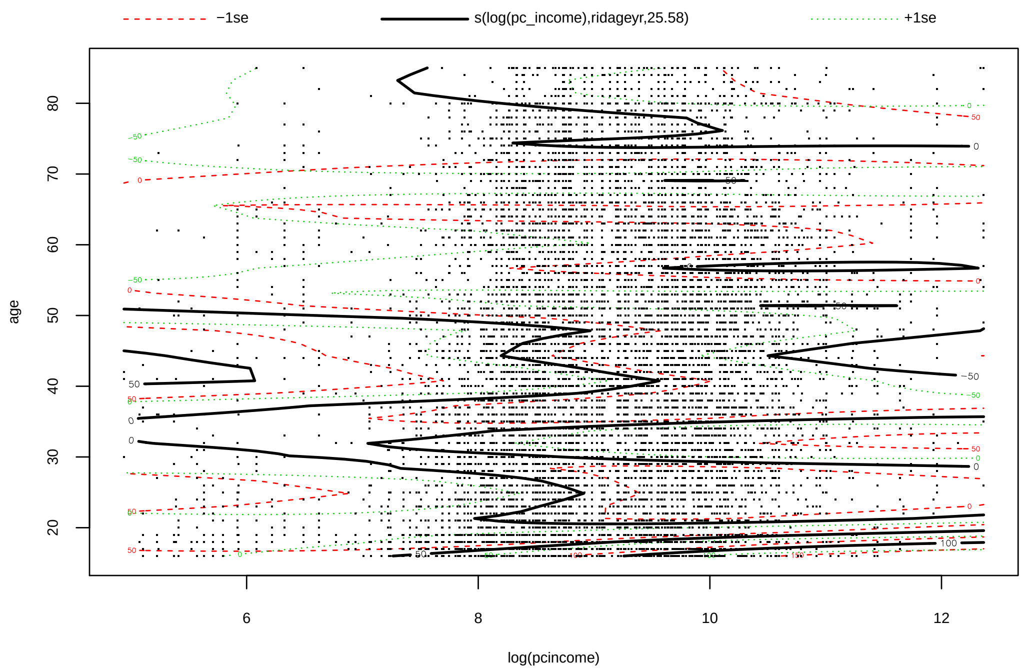
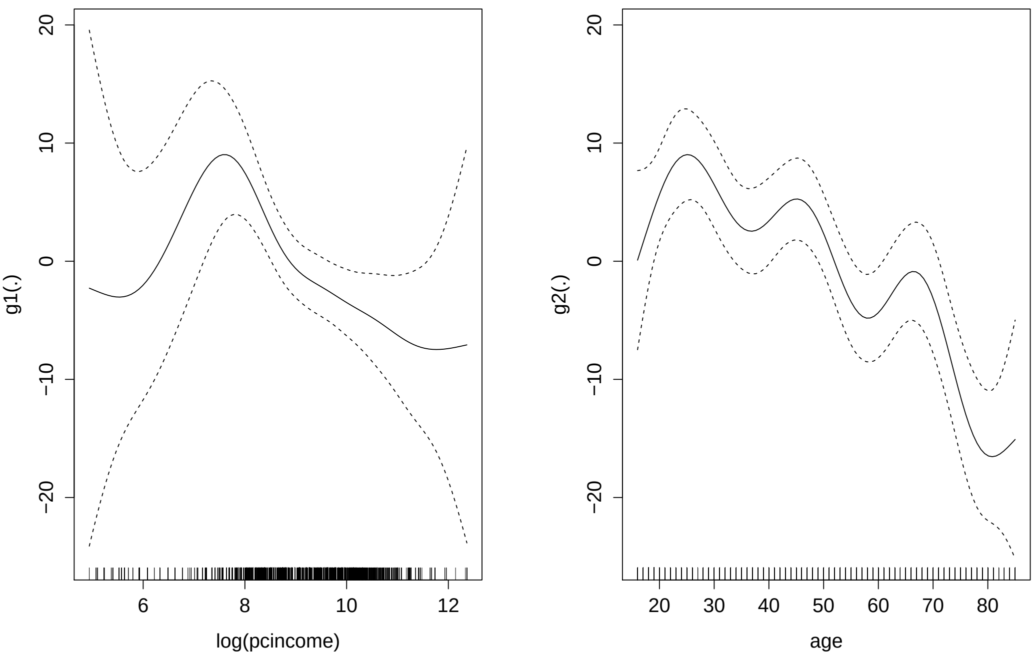


Figure 20: Smooth curve for the effect of the interaction term given by income and age on dairy consumption: contour plot.

are those linked to the gender, with women consuming 19 grams of eggs less than men and household with a woman as reference person consuming 3 grams of eggs per member less than other households. Income and age are both highly significant. The former appears to have a non-monotonic effect, peaking for the middle-classes and then decreasing substantially among the wealthy (Figure 21, left). The effect of age instead as a more negative trend, with substantial consumption by the youth and less among the elderly. In Figure 23 different situations can be identified. Among the wealthy, eggs consumption seems to peak both among the youth and then again for individuals around 70 years old. On the contrary, the poorest appear to have a peak between the age of 25 and 50, after which consumption decreases rapidly, while middle-class people have a rather constant consumption up to their 50s, that diminishes slowly thereafter. An interesting insight comes from the income elasticity structure of the demand for eggs in Figure 22. Eggs appear to be a superior food for lower-income classes, whose elasticity is positive and can get up to 0.4. It then becomes negative for individual belonging to the middle-class, meaning that they regard eggs as an inferior good, and then stabilizes at zero for people in the upper-classes, becoming a normal good.



(a) (b)

Figure 21: Smooth curves for the effect of income (left) and age (right) on eggs consumption.

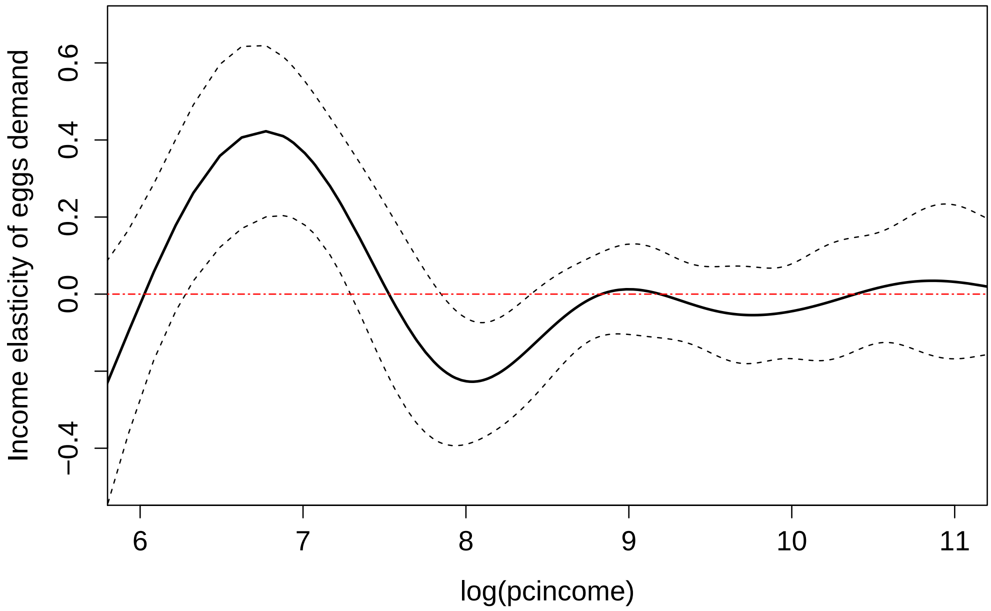


Figure 22: Structure of eggs elasticity curves with respect to income and age.

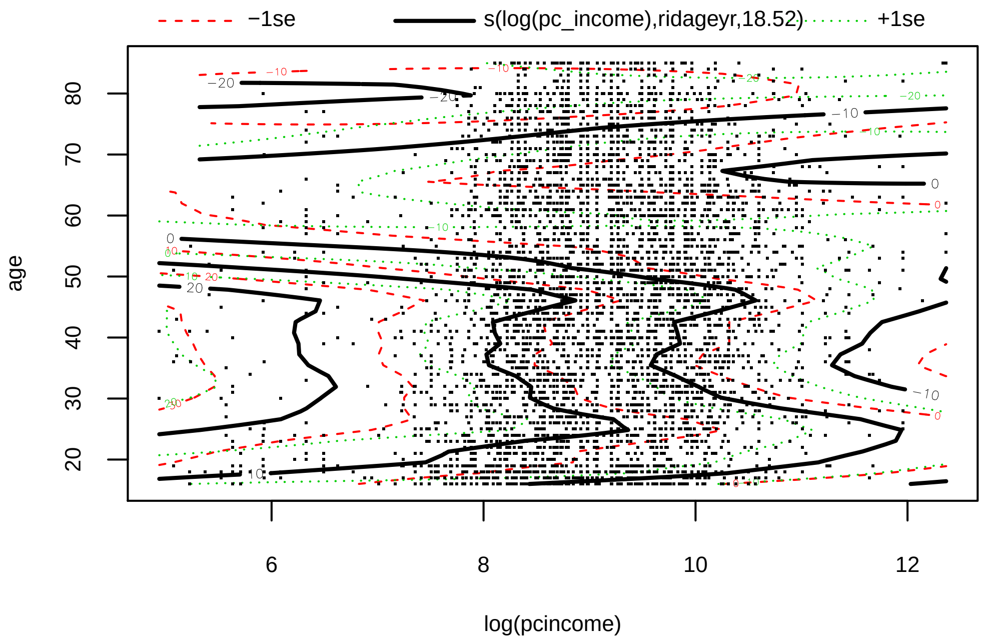


Figure 23: Smooth curve for the effect of the interaction term given by income and age on eggs consumption: contour plot.

6 Projections: a look at the future

Forecasting is never an easy task, and even less so when the structure of data is not explicitly dynamic. Since we only have cross-section data for different years, and not an actual panel data structure, we cannot track the 'movements' of individuals over time. As a consequence, this prevents the use of traditional forecasting procedures. In order to study the trend of meat and other animal-source foods consumption we therefore need to build a new dataset containing a time-series component, a sort of 'synthetic panel'. The most commonly used strategies to this purpose are the pseudo-panel approach, re-weighting, and microsimulations[19]. For simplicity's sake and due to the lack of high computational power and a ready-made economy-wide micro-model as those used by central banks and governments for simulations, the technique we employ is a less complex and more pragmatic one, drawing from the main grounding principles of the above-mentioned strategies: on re-weighting because we want to obtain new weights for individuals according to the evolution of the income and age distribution of the population; on micro-simulation, as we aim at obtaining a new dataset for the independent variables to fit in our model. And finally it draws on the pseudo-panel approach as we divide our individuals into fixed categories and then track the density within each of the category to obtain the time-trend component we lack. An important assumption we are making is that the structure of the coefficients estimated in the previous section stay the same over the short-to-medium term. This assumption is nonetheless not so extreme, as we have seen that per capita trends in consumption in the Western world have been quite stable in recent times, as opposite to the nutritional transition in developing countries. This is confirmed by our model, where we have checked that – with the exception of recession periods – the estimates does not appear to change much from one year to another. However, we have chosen not to force too much this assumption, and therefore selected a medium-term target year, the 2020, for plausible forecasting, together with the further away 2030.

6.1 Joint density of the predictors by year

We firstly take into consideration the main regressors we are interested in, i.e. the disposable income and the age of the individuals. Although the wage income earned by each person is likely to depend on the person's age, this is not the case for the 'income' variable we have been using, since it has been built not as individual earnings, but as the portion of the household income that each member of the household is supposed to have at their disposal. Similarly, even if fertility decisions – and therefore the age structure of the household or the whole population – may somehow be a consequence of the wealth at the household/family level, the latter cannot obviously influence the age of the single members, which evolve deterministically over time. We can therefore state the two variables to be independent. Given that in our model we are interested in the interaction between the two, we compute the joint density of age and income in our datasets for every year at our disposal[20]. To this purpose, we first build a matrix of joint frequency for a hundred income categories and for

[19]See Figari et al. (2014); Verbeek (2008); Li and O'Donoghue (2013) for reviews on these methodologies.

[20]Census data could also be used in alternative, for a more refined analysis, but because our data should in principle be representative of the whole US population, the two approaches should be indifferent.

every year of age, adjusted for the weight every individual carries. Starting with these joint frequency tables, we subsequently transform them into continuous two-dimensional functions by smoothing techniques. In particular, a kernel method for multivariate cases following Silverman (1986, pp.76-77) and Fukunaga and Hostetler (1975, p.175) is used . In the univariate case, a kernel estimator can be defined as ‘a sum of bumps, centred at the observation’ (Silverman, 1986). Generalizing this definition to the multivariate case, the kernel becomes a function of a d-dimensional argument x, satisfying the condition:

$$\int_{R^d} K(x)dx = 1$$

where R^d is the range over which the kernel is defined. Specifically, as K is usually taken to be a radially symmetric unimodal probability density function, we specify it as the standard multivariate Gaussian density (cfr. Silverman, 1986, p.76):

$$K(x) = (2\pi)^{-\frac{d}{2}} exp(-\frac{1}{2}x^T x).$$

To refine the estimation, the approach suggested by Fukunaga and Hostetler (1975) is adopted. Before smoothing using the kernel defined above, the data are linearly transformed so that their covariance matrix is unitary. At the end, the transformation is then reverted. Overall, this procedure is equivalent to the application of the following density estimate:

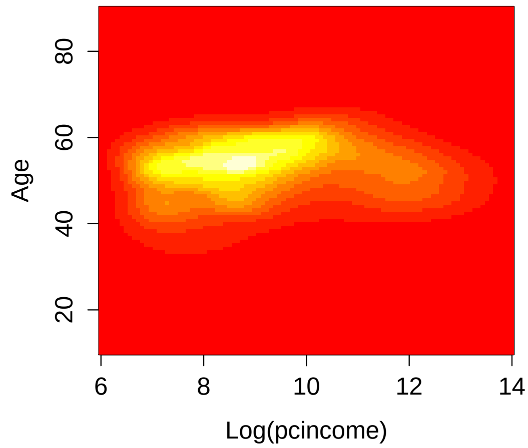
$$\hat{f}(x) = \frac{(\det S)^{\frac{-1}{2}}}{nh^d} \sum_{i=1}^{n} k\{h^{-2}(x - X_i)^T S^{-1}(x - X_i)\}$$

where S is the sample covariance matrix of the data and $k(x^T x) = K(x)$.

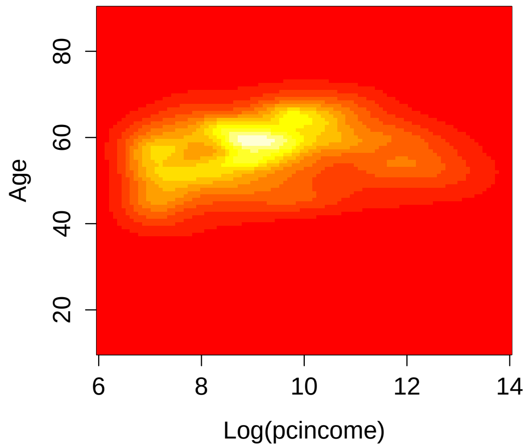
To further improve the smoothing, a correction is applied in estimating the tails of our joint distribution. In fact, when densities have long tails, i.e. regions of low density, it would be advisable to use a broader kernel in such regions. To this aim, Silverman (1986, pp.100-101) developed an *adaptive kernel*, combining together the typical elements of kernel estimates and nearest-neighbour approaches. The adaptive kernel uses a window width that is not fixed for every observation, but is allowed to change where needed. In order to identify whether a point belongs to a region of problematic low density, a preliminary step is implemented, getting a pilot estimate of the density $\tilde{f}(t)$. On this ground, bandwidths are computed for each observation, and then used in the adaptive kernel estimator, defined as follows:

$$\hat{f}(t) = \frac{1}{n} \sum_{i=1}^{n} h^{-d} \lambda_i^{-d} K\{\frac{1}{h\lambda_i}(t - X_i)\}.$$

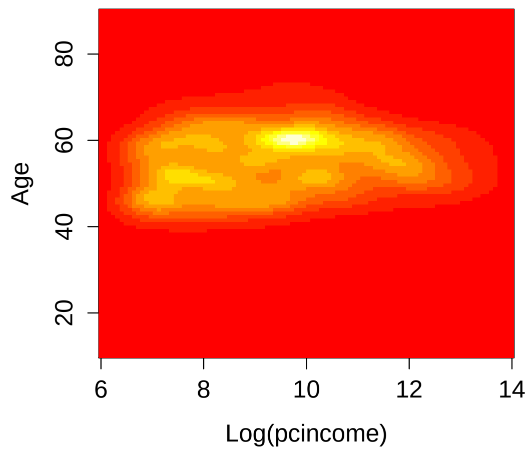
where λ_i is a local bandwidth factor estimated in the preliminary step. In principle, kernel estimation of the multivariate density could be applied for any number of dimensions d, and is not only bound to 2. This means that the joint density for all of the covariates could be estimated, adding precision to our analysis. Yet, lacking the computational power to make such estimate, we limit estimation to the joint density of our main predictors. The resulting density curves are shown in Figure 23. The aging pattern of the population is evident, while the evolution of the income is less regular, but tends to a progressive increase, too.



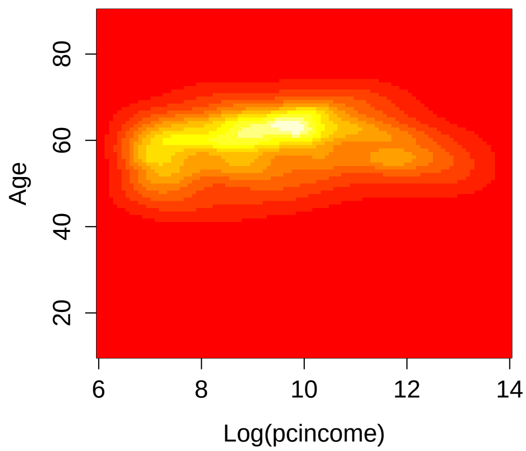
(a) *1994-1996*



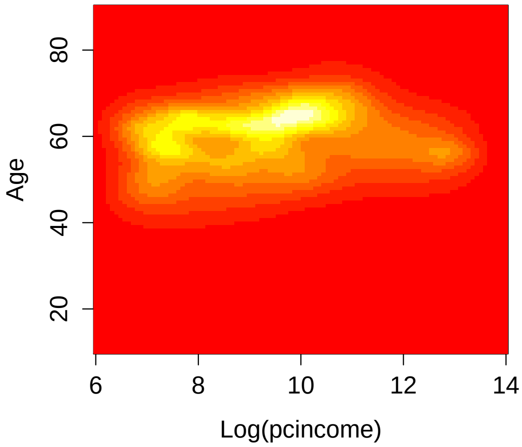
(b) *2001-2002*



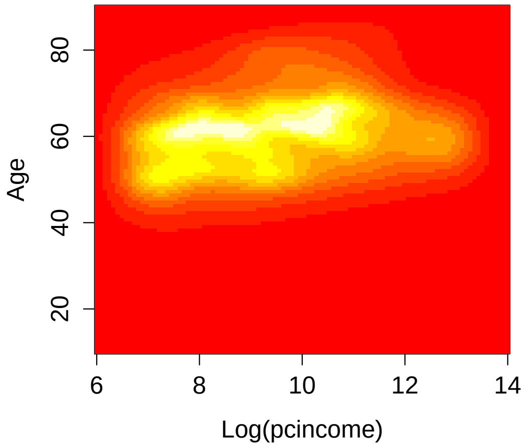
(c) *2003-2004*



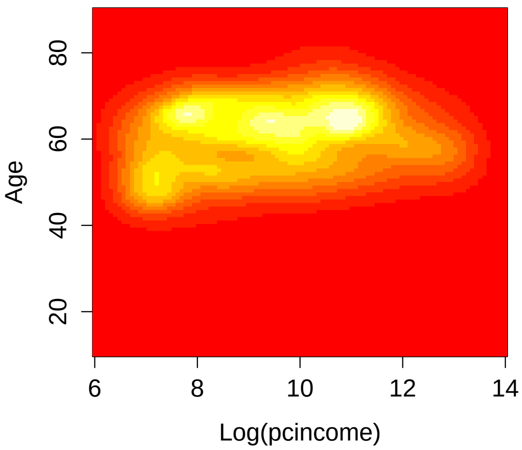
(d) *2005-2006*



(e) *2007-2008*



(f) *2009-2010*



(g) *2011-2012*

Figure 23: Age and income (log scale) joint density in different years.

6.2 Joint density dynamics and forecast

Even though we cannot track how each individual moves in the joint distribution, we can now follow how the density mass within each cross-category changes over the years. Before moving to that stage, we transform all the variables in their relative counterparts, by dividing them by their weighted average for that particular year. In doing so we isolate the distribution component from the overall trend due to continue economic growth and aging of the population. At this point we are only interested in the first one, while the second one will be included later on in the analysis, when we will multiply our projected relative distribution by the projected income and age average obtained from a demographic simulation. Then, taking into consideration one bin at a time, we build the time-series of the density mass contained in each of the bin representing the same characteristics over the years, and use it to estimate a model and extrapolate future values of the density mass in that bin[21]:

$$\pi_{ij,t} = f(\pi_{ij,t-1}, \pi_{ij,t-2}, \dots)$$

In order to increase the length of the time-series and counteract the subsequent high variability, values for the periods 1997-1998 and 1999-2000 are interpolated linearly from the available data, using an autoregressive specification of order 1, as the shortness of the series would not allow to take into consideration higher orders. A demeaning constant term is included and the Yule-Walker estimation method applied:

$$\text{AR(1)}: \ \pi_{ij,t} = \alpha_{ij} + \phi_{ij}\pi_{ij,t-1} \tag{4}$$

for $t = 0, ..., T$.The estimates obtained in Equation 4 are then used to forecast future values of the density mass in each bin up to 2030:

$$\begin{aligned} \pi^F_{ij,T+1} &= \alpha_{ij} + \hat{\phi}_{ij}\pi_{ij,T} \\ &\dots \\ \pi^F_{ij,T+\tau} &= \hat{\alpha}_{ij} + \hat{\phi}_{ij}\pi^F_{ij,T+\tau-1} \end{aligned}$$

The forecasted values are then normalized by dividing them by the result of the double integration over the two dimensions.

In order to preserve a certain degree of heterogeneity in the distribution of the control variables, we compute for each bin of the relative age and relative income joint density the weighted average of each controls, and interpolate linearly for bins in our region of interest where there are no individuals to compute from. For simplicity's sake we do this for the last years at our disposal, i.e. 2011-2012, and assume that this conditional distribution of the control variables stay more or less invariant with time. Since many of the controls are not highly significant, and the important issue in this step is to preserve the overall balance between men and women, black and non black, and so on, this assumption should not affect substantially our final projections. In fact, for a robustness check, we performed the same analysis using the overall weighted mean of each control variable for each individuals rather than a different value conditional on the income and age of the individual, and we obtained very similar results.

[21]Given the abysmal shortness of the time-series we obtain, the results of this exercise should be taken – as any forecast exercise in general – with a grain of salt.

6.3 Fitting the model

Before building our new dataset, we must remember that age and income have been used in relative terms, that is as their ratio to the average value in the specific year. In order to transform the data back to their absolute values, we therefore need the projected average value for each variable of interest (Table 8). In particular, we derive the forecasted data on the average (real) personal income by applying the growth rate predicted by the Economic Research Service (ERS) of the United States Department of Agriculture (USDA) for the real per capita GDP[22]. Projections on the annual growth rate of the population and on its specific age structure, were instead obtained through the DemProj module of the Spectrum Policy Modelling System, using data from the Population Division of the United Nations and their mid-scenario projections on fertility, mortality, and migration. In particular, the total population was estimated in order to be consistent with the 2010 Federal Census and the official population estimates for 2011. Total fertility rate, age mortality, and infant and child mortality were sourced from the country-specific official estimates on Life Tables ad Vital Statistics for the year 2010. Life expectancy at birth were based on official estimates of the age pattern of mortality from the Human Mortality Database. Finally, international migration estimates are based on official data on international migration and on the difference between overall population growth and natural increase through 2011. We can now build a new dataset, using the values of income and age of each cross-category as main regressors, the corresponding values of the other variables as controls, and the forecasted mass densities as new weights. The projected datasets are then fitted into the models we estimated in the previous section bin by bin, where each bin represents an individual observation and is then weighted by the corresponding mass in the joint distribution. First, the probit model is fit in order to know whether that 'individual' has a positive consumption of the food under consideration or not. If the resulting probability is above the threshold, the individual will be assigned a one, if it is lower, a zero.

$$
\begin{aligned}
Pr\{\hat{dummy}_{ij,T+\tau} = 1|X_{ij}\} = \Phi(&\hat{\alpha}_{1,T+\tau} + \hat{\alpha}_{2,T+\tau}\log(pcincome^F_{ij}) + \hat{\alpha}_{3,T+\tau}ridageyr^F_{ij} + \\
&+ \hat{\alpha}_{4,T+\tau}gender_{ij} + \hat{\alpha}_{5,T+\tau}hhsize_{ij} + \hat{\alpha}_{6,T+\tau}black + \\
&+ \hat{\alpha}_{7,T+\tau}hispanic_{ij} + \hat{\alpha}_{8,T+\tau}bornus_{ij} + \hat{\alpha}_{9,T+\tau}hrgender_{ij} + \\
&+ \hat{\alpha}_{10,T+\tau}hrbornus_{ij} + \hat{\alpha}_{11,T+\tau}winter_{ij})
\end{aligned}
$$

In case of a response equal to one, the GAM model corresponding to the food group is then fitted to obtain the quantity consumed.

$$
\begin{aligned}
\hat{c}_{ij,T+\tau} = &\hat{\beta}_{1,T+\tau} + \hat{g}_{T+\tau}(\log(pcincome^F_i), ridageyr^F_j) + \hat{\beta}_{2,T+\tau}gender_{ij} + \\
&+ \hat{\beta}_{3,T+\tau}hhsize_{ij} + \hat{\beta}_{4,T+\tau}black + \tau_{ij} + \hat{\beta}_{5,T+\tau}hispanic_{ij} + \\
&+ \hat{\beta}_{6,T+\tau}bornus_{ij} + \hat{\beta}_{7,T+\tau}hrgender_{ij} + \hat{\beta}_{8,T+\tau}hrbornus_{ij} + \\
&+ \hat{\beta}_{9,T+\tau}winter_{ij}
\end{aligned}
$$

This operation is repeated recursively for each bin, for each food group and for each two-year span. Overall consumption is then obtained by integrating over the two dimension of age and income the estimated curve of consumption obtained from the bin-by-bin estimation, weighted by the forecasted distribution of the population. The

[22] ERS International Macroeconomic Data Set, updated 18 December 2014.

latter is computed multiplying the normalized joint distribution of the population over age and income projected for each future date as explained above, by the projected total adult population (16-80+) obtained from DemProj, for each year under analysis as reported in Table 8:

$$\hat{C}_{T+\tau} = \int_i \int_j \hat{c}_{ij,T+\tau}(f^F_{ij,T+\tau} P^F_{T+\tau}) d_i d_j$$

where $\hat{c}_{ij,T+\tau}$ is the fitted 'individual' consumption, $f^F_{ij,T+\tau}$ is the normalized forecasted joint density function and $P^F_{T+\tau})$ is the forecasted population. It is to be noted that the US population is projected to increase by about 8% by 2020 and 16% by 2030 with respect to 2011 levels. Finally, we compute the growth rate of the estimated consumption for each projected year. As a further comparison, average per capita consumption are also computed. Red dots are added in the Figures to compare fitted values with true values for years 2003 to 2012. Although the average per capita consumption is forecasted to stay substantially the same for meat, with evidence of a slightly decreasing trend (Figure 26), business-as-usual is not a viable strategy, as growth in population will determine a 3 to 12% increase with respect to 2011-2012 levels by 2020 and a 9 to 26% increase by 2030 (Figure 24), consistently with what predicted in Lin et al. (2003). Annual rate of growth of the country consumption is projected to decrease, from the present average of 2% to about 1% in 2030, yet a positive rate, signaling no reversion in the meat consumption trend of the US. A very

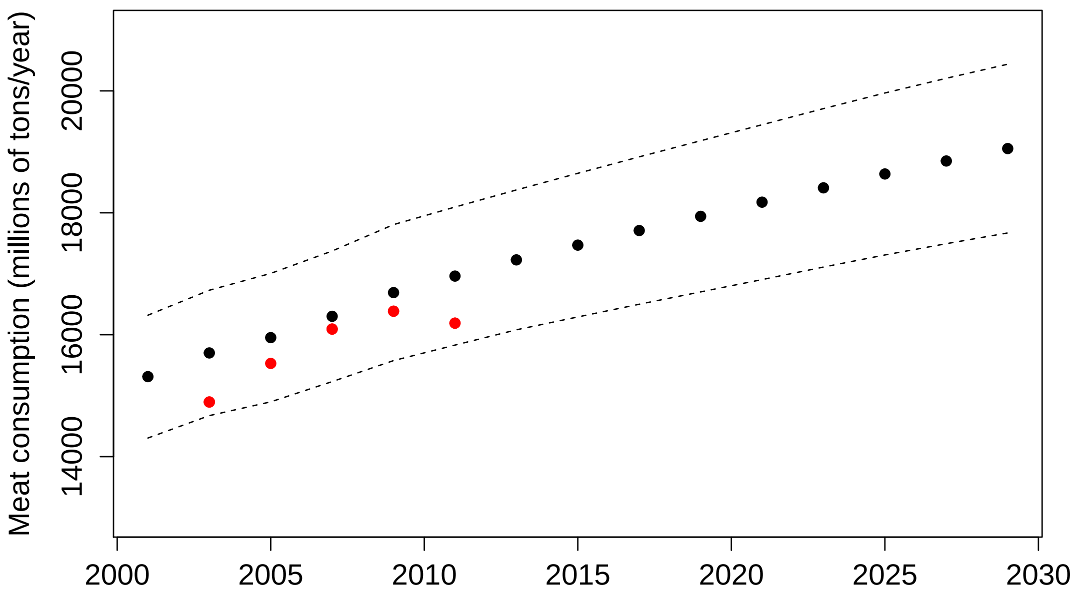


Figure 24: Projected overall meat consumption in the USA up to 2030, in tons per year. Red dots are true values, black dots are fitted values.

similar situation is also portrayed for beef consumption, whose per capita quantity may decrease slightly (Figure 29), and be even able to counteract the increase in total population for a while, with an overall growth by 2020 estimated to be between

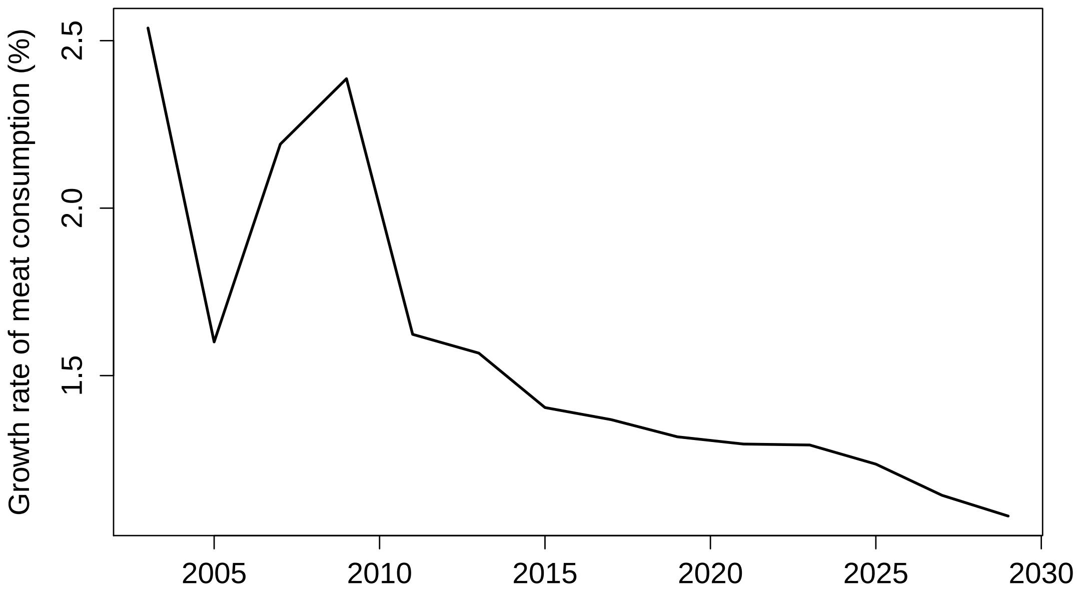


Figure 25: Projected annual rates of growth for overall meat consumption in the USA up to 2030, in percentage.

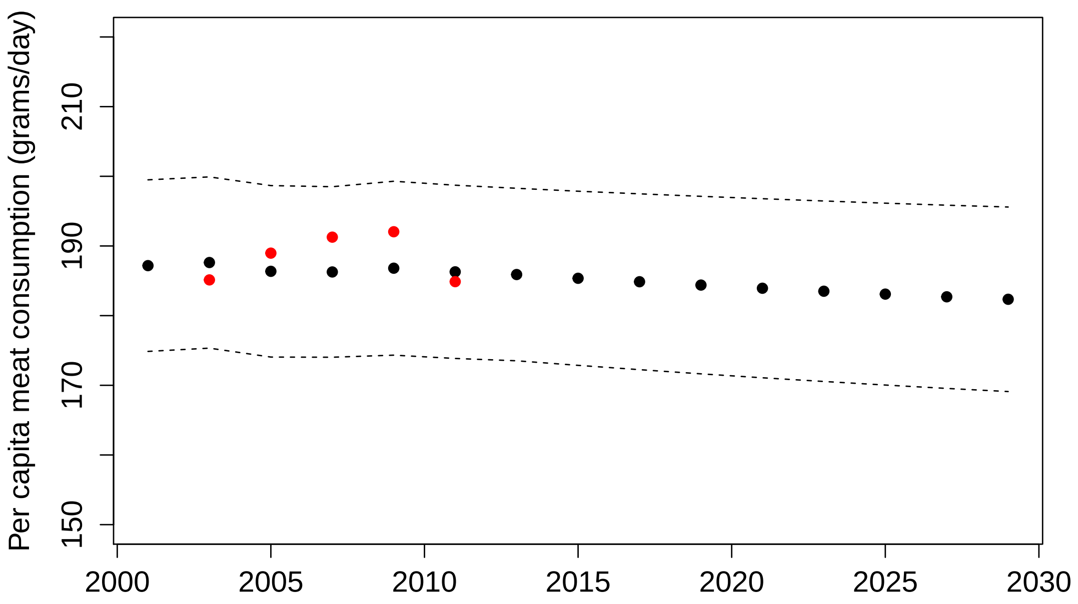


Figure 26: Projected average per capita consumption of meat in the USA up to 2030, in grams per day. Red dots are true values, black dots are fitted values.

-3 to 22% with respect to 2011-2012 levels. Nonetheless, even the reductions in per capita consumption presented by the lower-bound estimates would be too feeble to supersede the effect of population growth by 2030, when the quantity of beef consumed in the US is projected to be 2 to 30 The projections of pork consumption

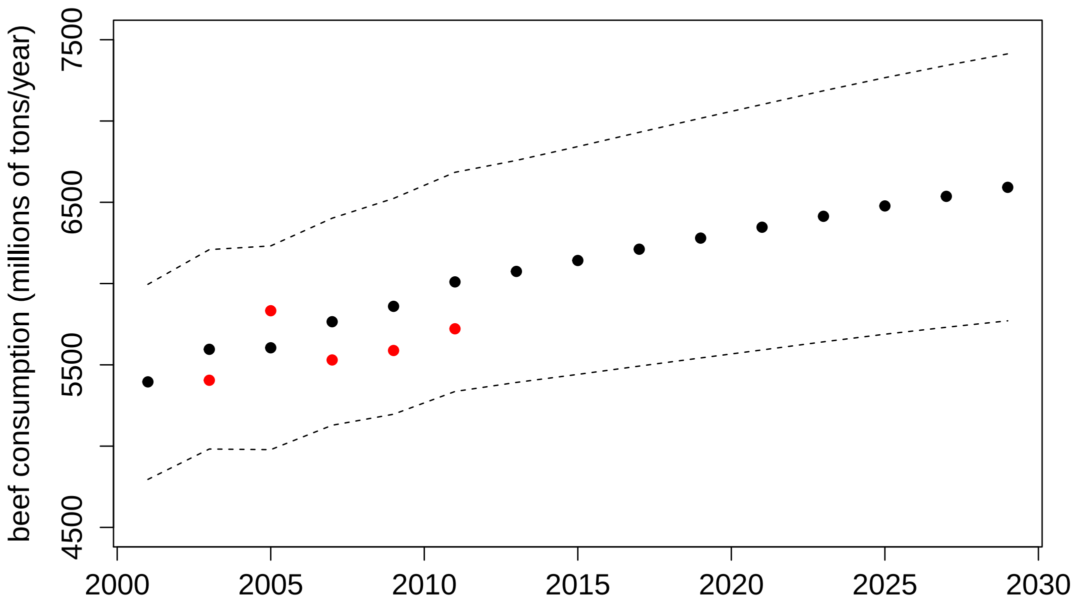


Figure 27: Projected overall beef consumption in the USA up to 2030, in tons per year. Red dots are true values, black dots are fitted values.

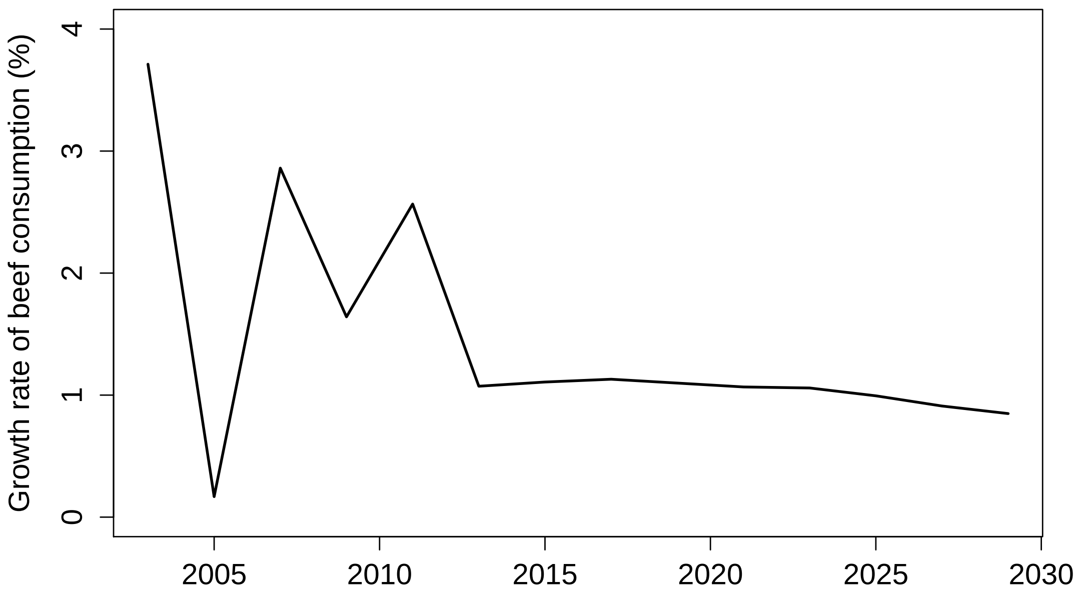


Figure 28: Projected rates of growth for overall beef consumption in the USA up to 2030, in percentage.

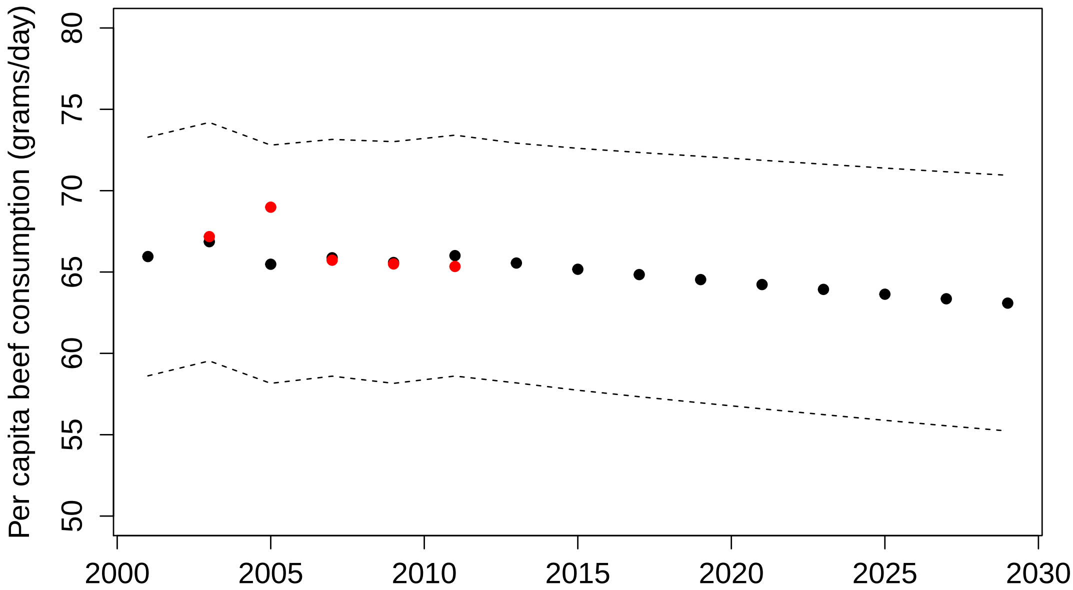


Figure 29: Projected average per capita consumption of beef in the USA up to 2030, in grams per day. Red dots are true values, black dots are fitted values.

are equally – if not more – worrisome, with a projected annual growth rate steadily over 1% as shown in Figure 30. Yet, the confidence interval of per capita (Figure 32) and – consequently – of overall consumption becomes wider with time, meaning that the direction of future trends is actually quite uncertain, with an overall quantity consumed in 2020 and 2030 that could be anywhere inbetween -12 and +30% with respect to the last data available, for the former, and between -8% and +39% for the latter. The estimated models appear to work quite well for both beef and pork consumption, as the true values are extremely close to the fitted ones. The demand for poultry and fish – the healthier and relatively less polluting alternatives for red meat – are both projected to grow steadily at about 1.5% and 2% per year respectively (Figures 33 and 36). If on one hand per capita poultry consumption does not show any appreciable variation in per capita consumption in Figure 35, meaning that its increase will be caused essentially by the growth in total population, the per capita demand for fish illustrated in Figure 38 present an upward trend. Overall, poultry consumption is forecasted to undergo a growth of 1 to 23% by 2020 and 8 to 32% by 2030 with respect to 2011-2012, while fish may score higher percentage of 7 to 38% by 2020 and 16 to 53% by 2030. While the model estimated for poultry appears to have a good fit with respect to the true values, the fitted values for fish seem to differ more significantly from the true values. This may be due to the fact that we had less observations and more variability for fish consumption, as not so many people actually ate this during the two days of the interviews. A dietary survey over longer span of time would surely improve the estimates. Per capita dairy consumption appears to remain flat in the next decades, although being an heterogeneous category this may hide variation in its within composition. Its yearly average growth rate is likely to slow down in the next decades, but stay positive throughout. Overall consumption

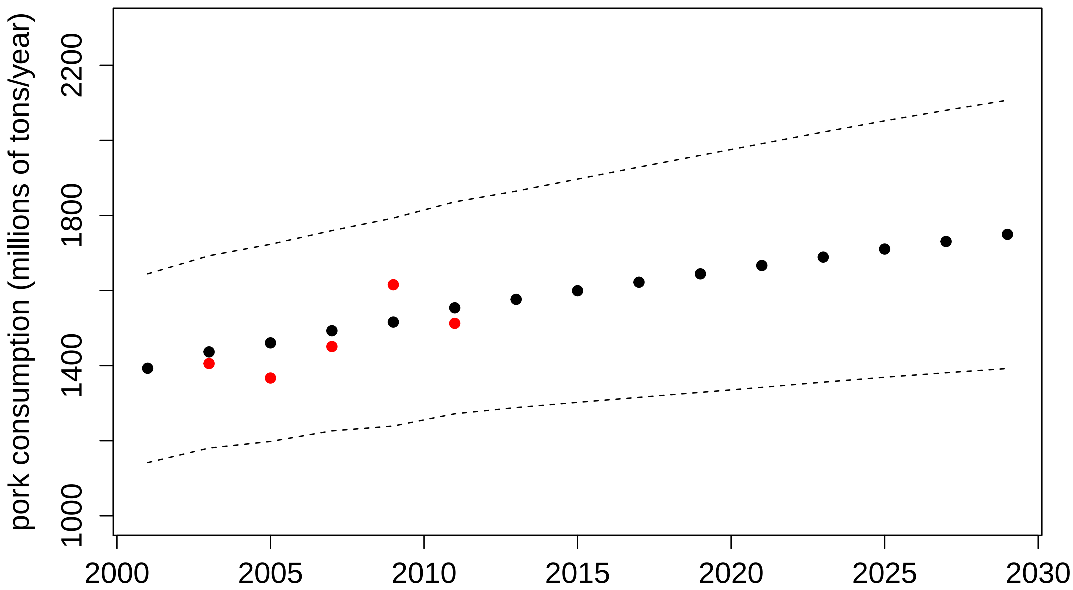


Figure 30: Projected overall pork consumption in the USA up to 2030, in tons per year. Red dots are true values, black dots are fitted values.

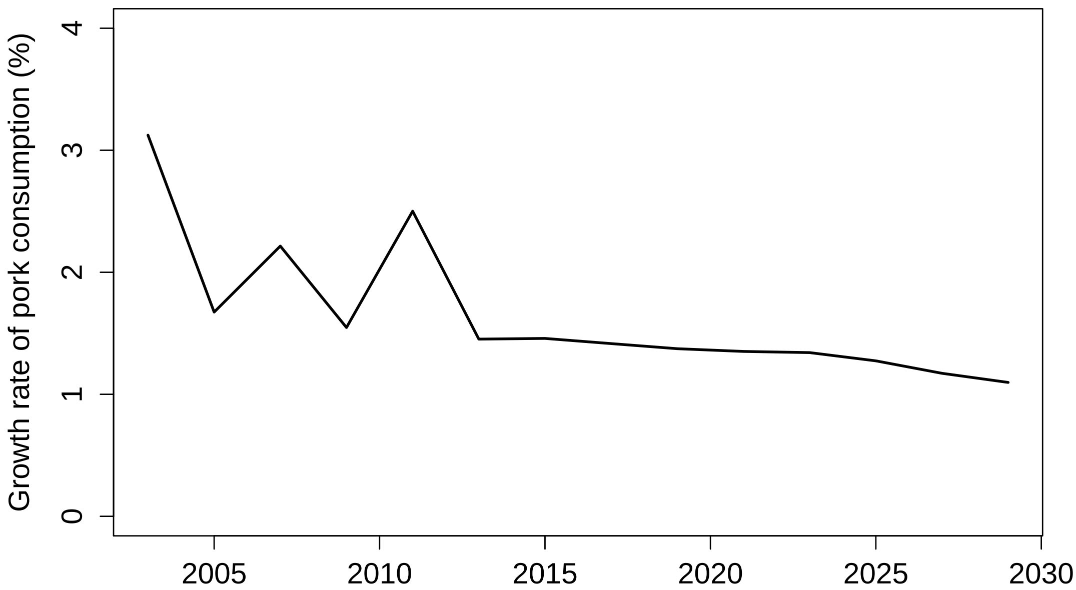


Figure 31: Projected rates of growth for overall pork consumption in the USA up to 2030, in percentage.

is in fact projected to grow 4 to 32% by 2020 and 12 to 43% by 2030 with respect to 2011-2012 (Figure 39). Eggs consumption on the contrary shows a flatter but lower average growth rate per year. Over the next few decades, consumption may even

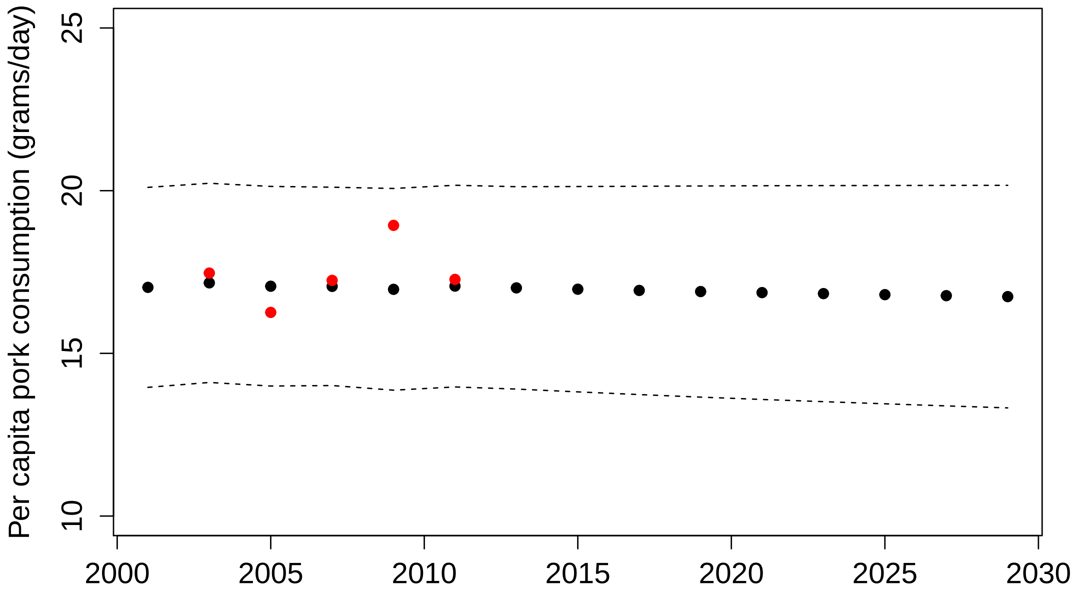


Figure 32: Projected average per capita consumption of pork in the USA up to 2030, in grams per day. Red dots are true values, black dots are fitted values.

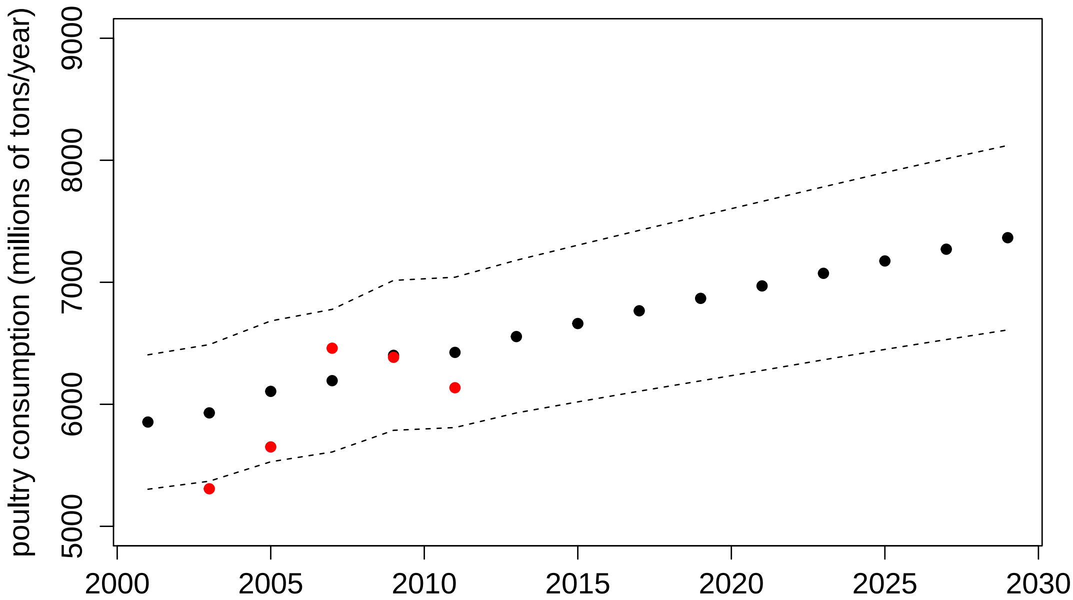


Figure 33: Projected overall poultry consumption in the USA up to 2030, in tons per year. Red dots are true values, black dots are fitted values.

decrease slightly and stabilize with respect to 2011-2012 levels, with an estimated variation between -4 and 19% by 2020 and 2 to 28% by 2030 (Figure 42). In both

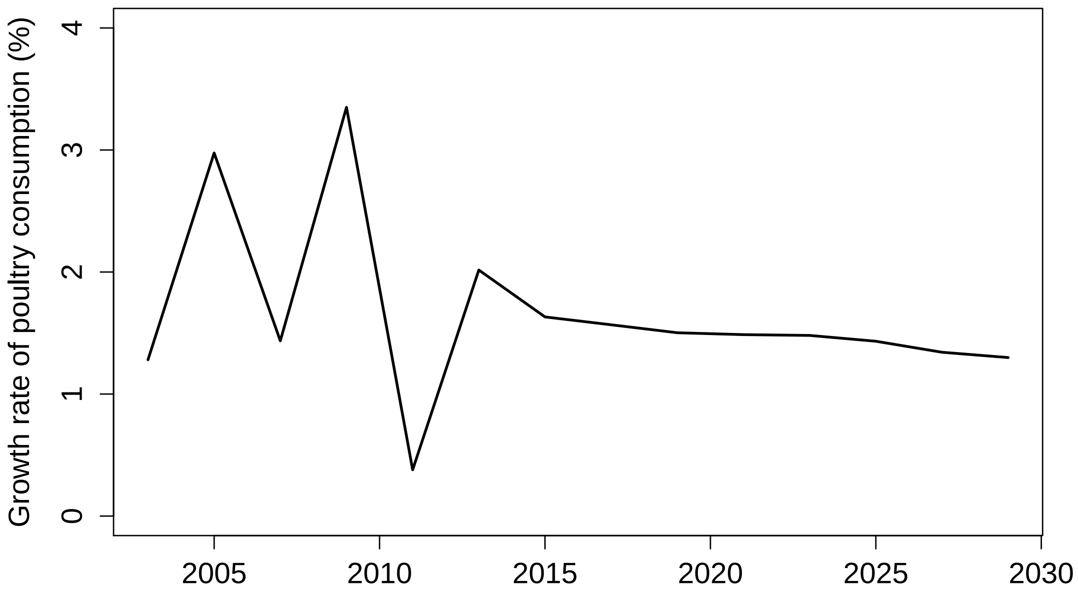


Figure 34: Projected rates of growth for overall poultry consumption in the USA up to 2030, in percentage.

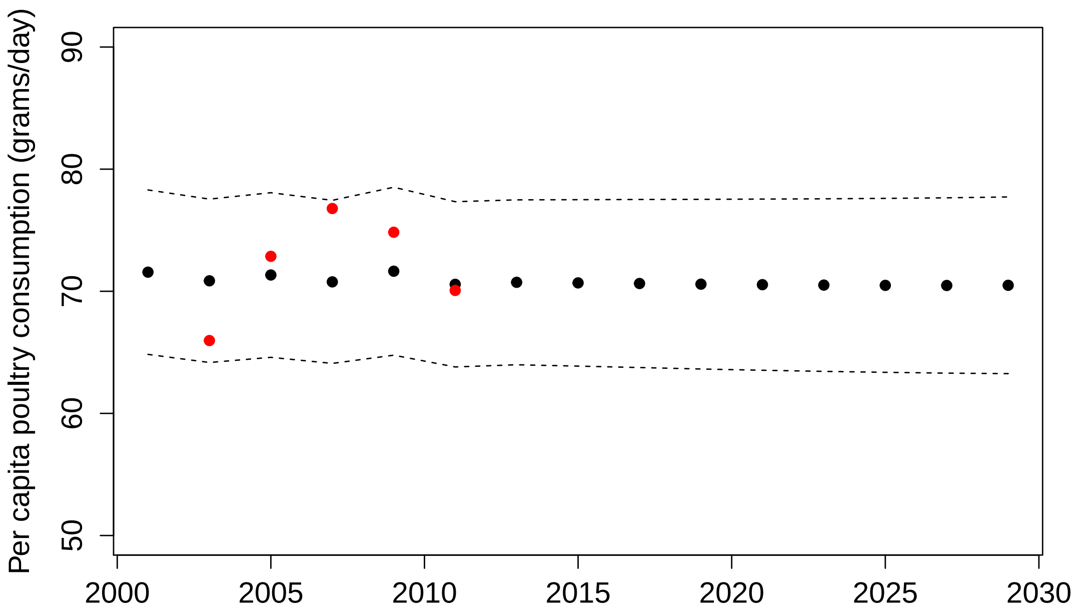


Figure 35: Projected average per capita consumption of poultry in the USA up to 2030, in grams per day. Red dots are true values, black dots are fitted values.

cases the true values (red dots) are well inside the confidence interval, a good sign for the fit of our model. Overall, our results are compatible with USDA (2010) baseline projections to 2021, which forecast a rise in per capita meat consumption of 2 percent

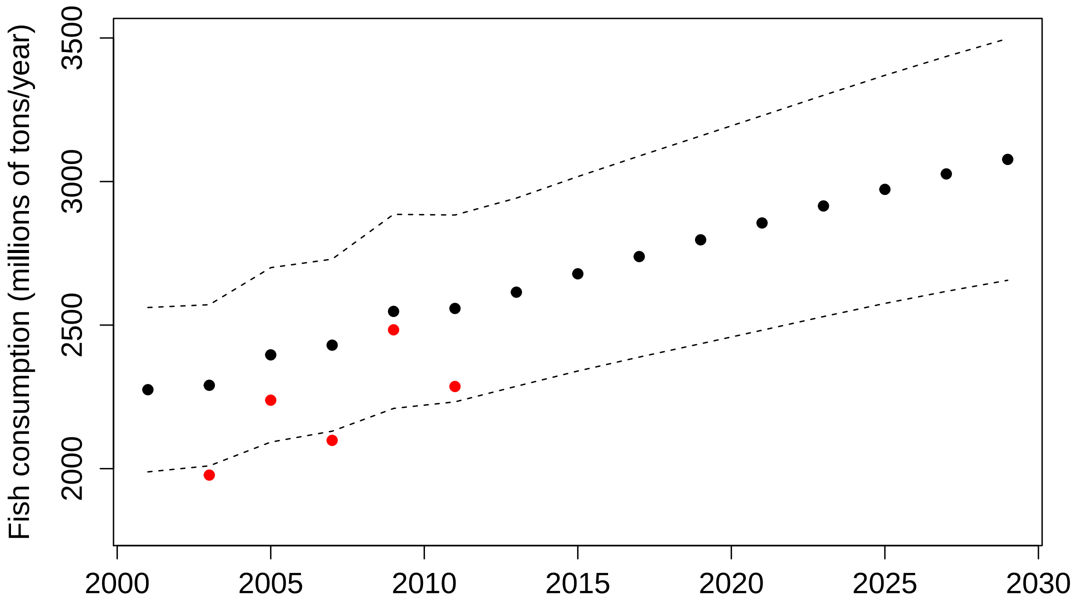


Figure 36: Projected overall fish consumption in the USA up to 2030, in tons per year. Red dots are true values, black dots are fitted values.

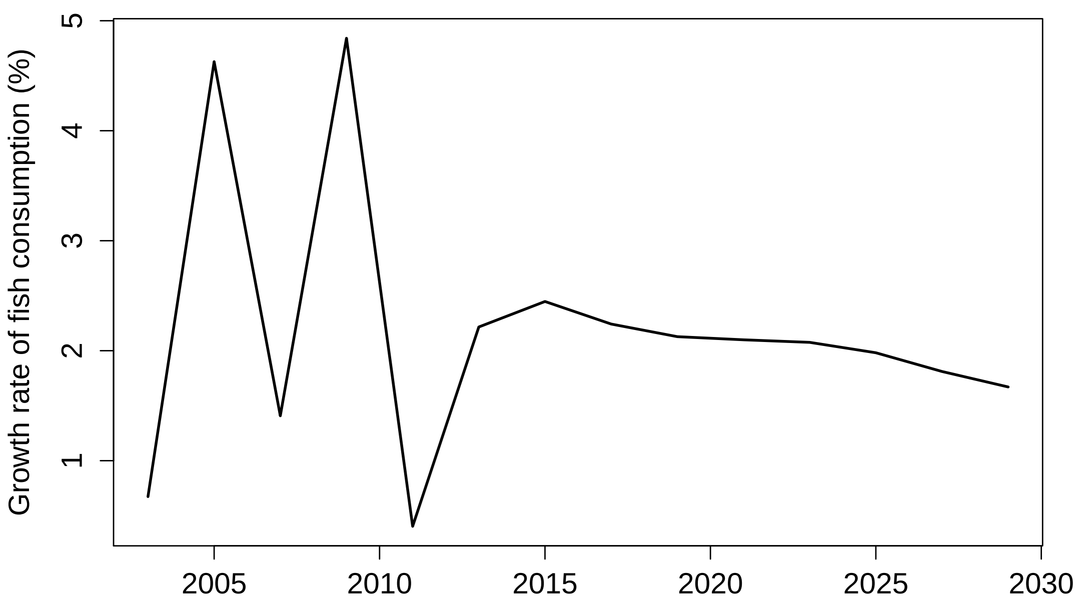


Figure 37: Projected rates of growth for overall fish consumption in the USA up to 2030, in percentage.

with respect to 2010, disaggregated into slight decrease in red meat and increase in poultry consumption, but on average more optimistic than the macro-based projections presented above, as most of the consumptions are projected to stay flat at the

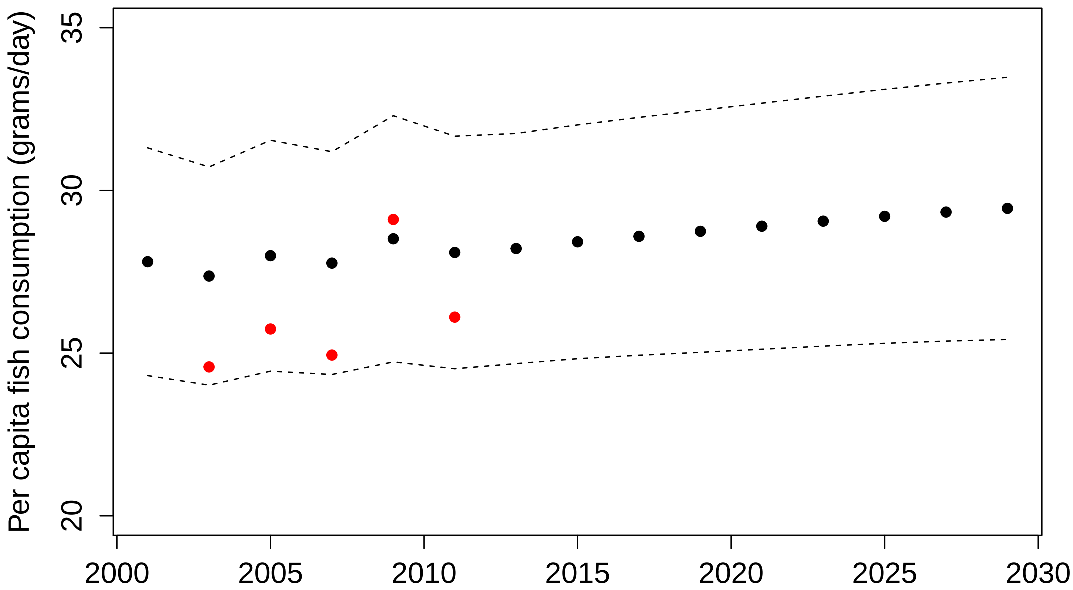


Figure 38: Projected average per capita consumption of fish in the USA up to 2030, in grams per day. Red dots are true values, black dots are fitted values.

per capita level. Similarly, the study conducted by Lin et al. (2003) on micro-data predicts a rise in fish, poultry and eggs and a fall in beef, pork and dairy. While the trends for fish, beef, pork and poultry are consistent with our projections, those for eggs and dairy seem less concordant. In line with our analysis, they also underline that although the main driver of increases in food consumption will be population growth, the within composition of this trend will depend on the distribution of the characteristics of the population, in particular the age structure.

Year	Population16-80+	Average age of pop. 16-80+	Real per capita GDP growth
2001	222,840,866	43.67	-0.03
2002	225,418,579	43.78	0.82
2003	227,984,085	43.89	1.86
2004	230,561,062	44.01	2.85
2005	233,197,674	44.14	2.39
2006	235,845,002	44.28	1.70
2007	238,482,313	44.43	0.81
2008	241,072,674	44.58	-1.22
2009	243,583,233	44.74	-3.63
2010	245,997,422	44.89	1.69
2011	248,327,013	45.05	0.87
2012	250,589,799	45.21	1.60
2013	252,804,311	45.38	1.50
2014	254,983,781	45.55	1.61
2015	257,142,321	45.72	2.78
2016	259,281,300	45.89	2.19
2017	261,399,650	46.06	1.89
2018	263,493,676	46.23	1.89
2019	265,564,535	46.40	1.90
2020	267,623,165	46.56	1.90
2021	269,681,803	46.72	1.90
2022	271,748,437	46.86	1.91
2023	273,824,012	47.01	1.91
2024	275,886,619	47.14	1.91
2025	277,903,668	47.27	1.91
2026	279,856,930	47.40	1.91
2027	281,744,226	47.53	1.91
2028	283,582,823	47.65	1.92
2029	285,388,379	47.76	1.93
2030	287,160,345	47.86	1.94

Table 8: Data and projections on the US adult population, its average age, and the real per capita GDP growth rate per year, for the period 2001-2030.
Source: DemProj module in Spectrum software for projections of population and age structure; USDA ERS International Macroeconomic Data Set for real per capita GDP projections.

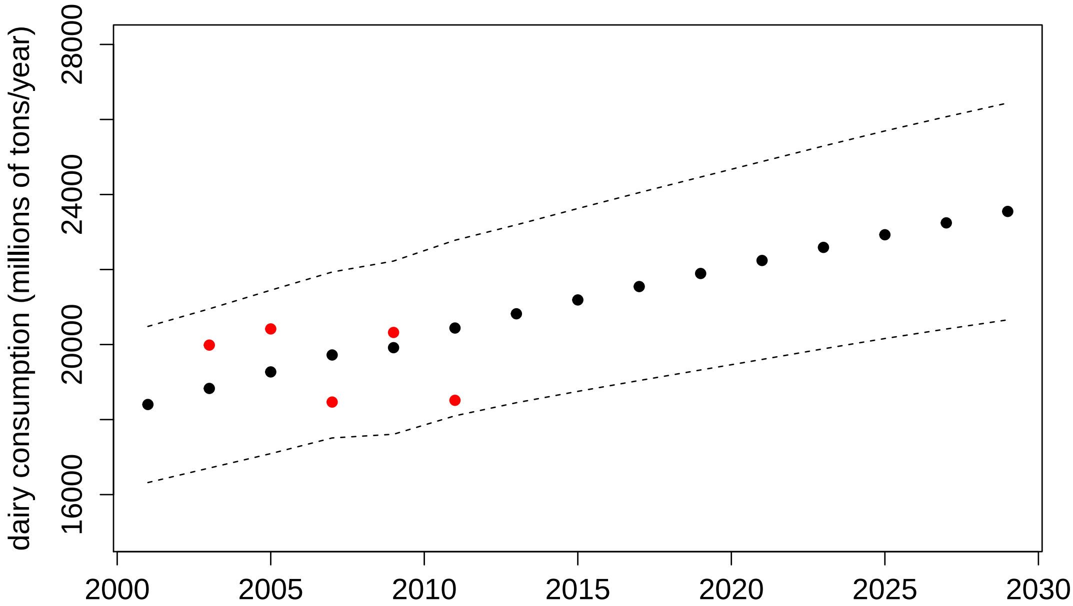


Figure 39: Projected overall poultry consumption in the USA up to 2030, in tons per year. Red dots are true values, black dots are fitted values.

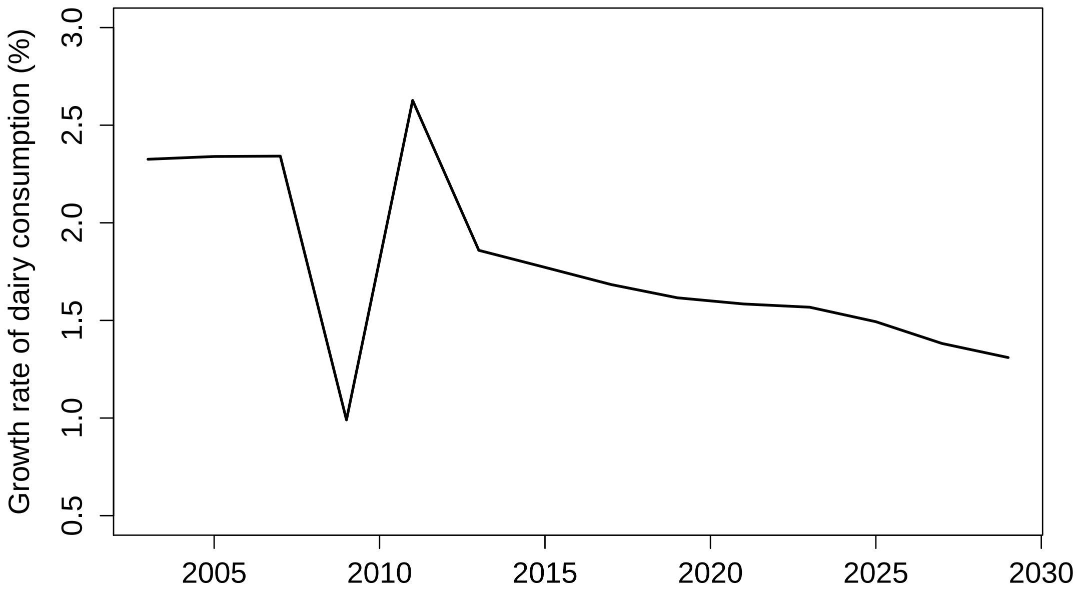


Figure 40: Projected rates of growth for overall dairy consumption in the USA up to 2030, in percentage.

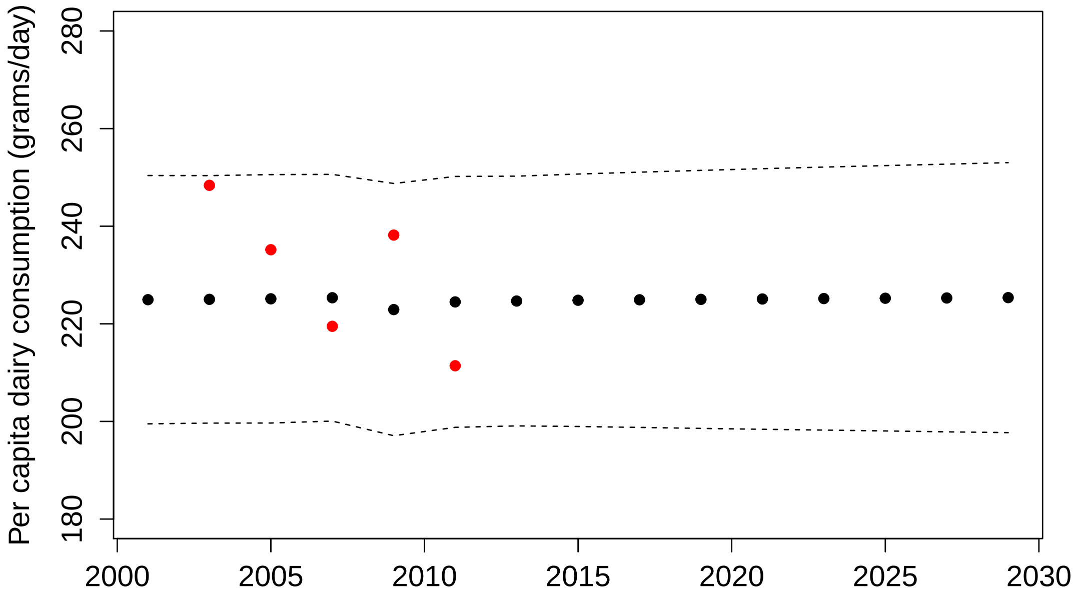


Figure 41: Projected average per capita consumption of dairyin the USA up to 2030, in grams per day. Red dots are true values, black dots are fitted values.

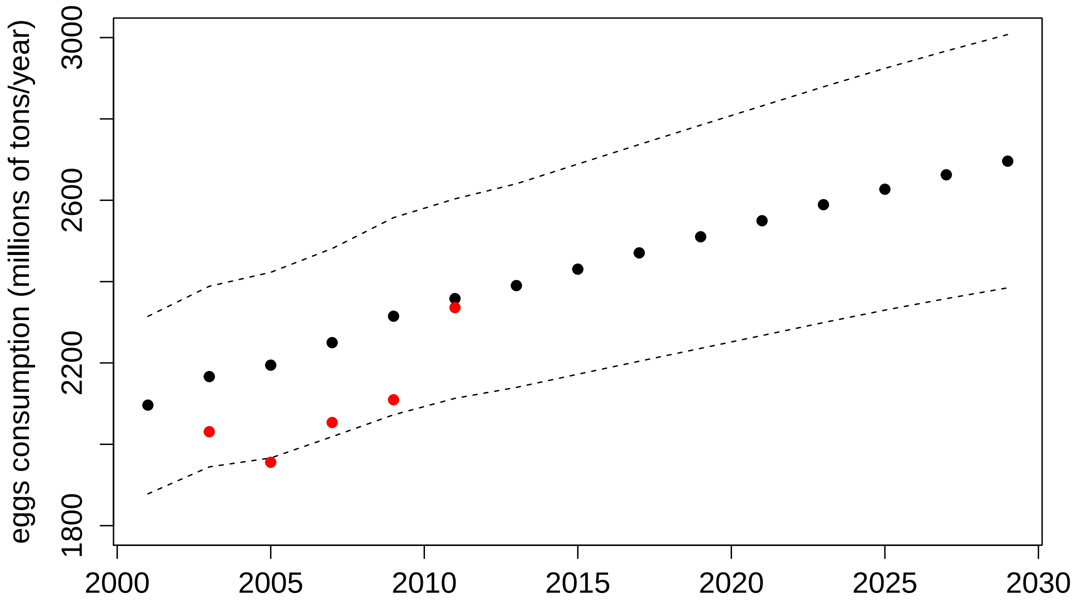


Figure 42: Projected overall eggs consumption in the USA up to 2030, in tons per year. Red dots are true values, black dots are fitted values.

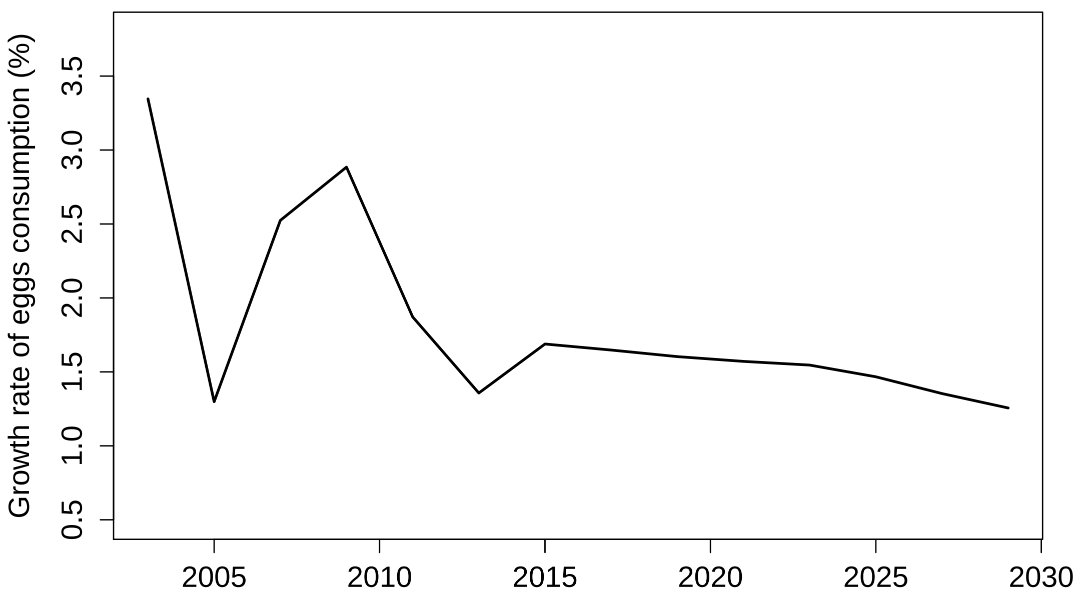


Figure 43: Projected rates of growth for overall eggs consumption in the USA up to 2030, in percentage.

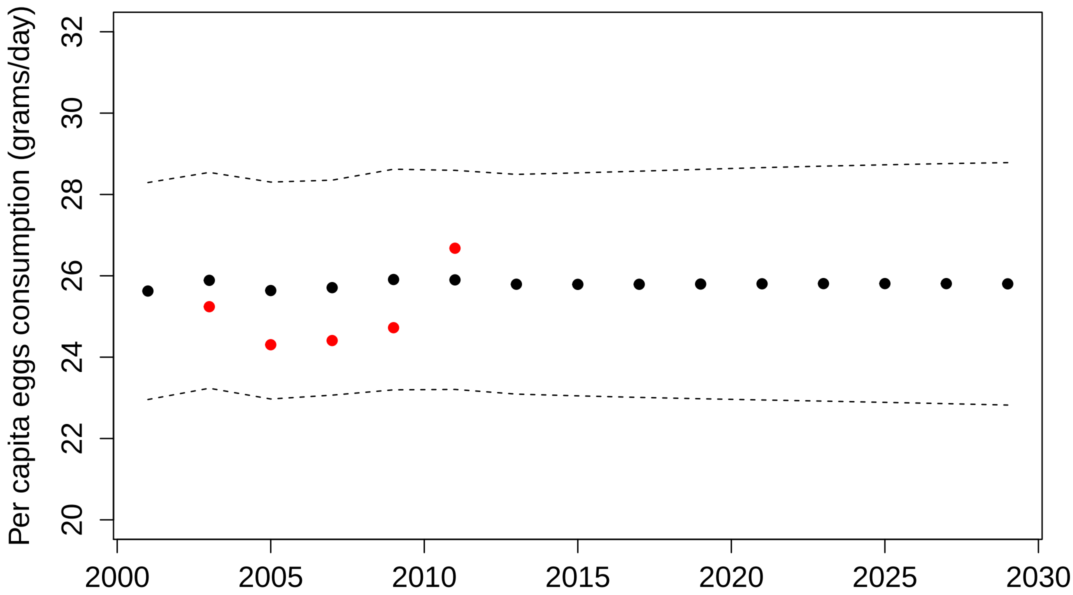


Figure 44: Projected average per capita consumption of eggs in the USA up to 2030, in grams per day. Red dots are true values, black dots are fitted values.

7 Conclusion

To sum up, our micro-based approach – contrarily to the macro-based projections seen in the first part of this work – allows us to take into consideration a neglected but important element of animal-food demand, namely age and the age structure of the population. In Section 5 we have seen that age has a strong non-linear and non-monotonous effect on the consumption of animal-origin foods, meaning that aging population such as that of the US and of other western countries might lead to a decrease in average per capita consumption, differently from what forecasted at the macro level. Yet, even in our best-case scenario, these reductions in per capita consumption seem unlikely to be able to counteract the increase in population of these same countries, resulting anyway in a decisively positive growth of the overall consumption levels, as we have seen in Section 6. Moreover, this approach allows us to analyse the distribution and the mix of consumptions on a disaggregated level, showing who are the more likely victims of the adverse effects linked to the 'abuse' of red meat and animal-source food in general. In fact, in Section 5 we have seen what are the characteristics that make an individual more likely to belong to the over-consumer category, with peaks being especially likely for male young-to-mid-age adults causing health problems that are irremediably going to last and worsen with age, and in some specific cases for low-income people that are also the portion of population with more limited access to health services in the US. Given the amount of externalities and public-good issues involved in environmental protection and public health, and the role of time-discounting and habits reversal linked to individual health, some policy intervention may actually be beneficial. Education on the health and environmental consequences of animal-source foods consumption, and food guidelines that take them into consideration may for instance be a way. Something seems to be moving in this direction, as this year (February 2015) the US Dietary Guidelines Advisory Committee (DGAC) has explicitly advised towards reduction in animal-origin foods for the first time in its history, pleading both environmental and health concerns (DGAC, 2015). Whether the US Department of Health and Human Services (HHS) and the US Department of Agriculture (USDA) will finally include the DGAC's recommendations in the forthcoming *2015 Dietary Guidelines for Americans* is nonetheless an open question, given the manifested opposition of livestock and food sector lobbies and their vested interest. The specific situation of the US is just a tile in the global mosaic, presented in Section 3. With nutrition transitions on the way in the newly industrialized countries coexisting side by side with still widespread protein deficiency and de-nutrition, the developed world needs to lead actions to insure sustainability of consumptions, before pressures on livestock production and cropland may cause detrimental increases in the price of food worsening the situation and threatening food security among the poorest, and before the damages on ecosystems and the global environment become irreversible. We hope to have been able through this work to cast some useful light on the mechanisms behind these problems and stimulate a fertile debate on possible solutions.

Appendix A Summary statistics and correlation tables for NHANES datasets

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
meat	8213	284948179	171.203	146.2511	0	2066.91
fish	8213	284948179	20.79407	54.42204	0	745.5
dairy	8213	284948179	293.1311	306.9425	0	3744.44
eggs	8213	284948179	23.16919	42.78874	0	426.98
beef	8213	284948179	61.48534	97.77188	0	1476.25
pork	8213	284948179	15.40764	41.99967	0	764.785
poultry	8213	284948179	61.7996	92.86437	0	2054.91
pc_income	7754	271118519	14533.82	17720.52	188.2038	227717.4
ridageyr	8213	284948179	36.21556	22.24063	0	85
gender	8213	284948179	0.5120786	0.4998845	0	1
hhsize	8213	284948179	3.342595	1.567032	1	7
black	8213	284948179	0.1226002	0.3279977	0	1
hispanic	8213	284948179	0.1261054	0.3319883	0	1
bornus	8213	284948179	0.8952653	0.3062299	0	1
hrgender	8213	284948179	0.433375	0.4955714	0	1
hrbornus	8213	284948179	0.8305793	0.3751459	0	1
winter	8213	284948179	0.4082316	0.4915363	0	1

Table 9: Summary statistics for 2003-2004 data.

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
meat	8257	289674102	173.3455	148.7652	0	1509.81
fish	8257	289674102	23.79763	59.94648	0	1514.565
dairy	8257	289674102	288.4995	292.7897	0	2941.705
eggs	8257	289674102	21.9348	42.72086	0	640.5
beef	8257	289674102	64.03956	103.7744	0	1091.825
pork	8257	289674102	14.56062	40.20445	0	640
poultry	8257	289674102	63.92954	92.58091	0	1323.565
pc_income	8026	284419203	18335.54	22492.32	164.5665	234350.8
ridageyr	8257	289674102	36.70658	22.23131	0	85
gender	8257	289674102	0.5179112	0.4997093	0	1
hhsize	8257	289674102	3.310892	1.594344	1	7
black	8257	289674102	0.1237879	0.3293594	0	1
hispanic	8257	289674102	0.1238664	0.329449	0	1
bornus	8257	289674102	0.8883917	0.3149029	0	1
hrgender	8257	289674102	0.4171357	0.4931156	0	1
hrbornus	8257	289674102	0.8333772	0.3726613	0	1
winter	8257	289674102	0.4184546	0.4933354	0	1

Table 10: Summary statistics for 2005-2006 data.

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
meat	7711	295675046	175.1567	154.427	0	1592.755
fish	7711	295675046	21.29755	58.15219	0	1334.375
dairy	7711	295675046	259.2714	261.3122	0	2548.815
eggs	7711	295675046	22.16625	41.87285	0	504
beef	7711	295675046	58.82249	96.37285	0	1592.755
pork	7711	295675046	15.26033	43.76927	0	581.875
poultry	7711	295675046	70.94806	104.5058	0	1036
pc_income	7477	287711506	16481.57	17553.58	159.7263	220044.6
ridageyr	7711	295675046	36.63574	22.02942	0	80
gender	7711	295675046	0.5291739	0.4991805	0	1
hhsize	7711	295675046	3.383309	1.596199	1	7
black	7711	295675046	0.1225546	0.3279465	0	1
hispanic	7711	295675046	0.1521667	0.3592057	0	1
bornus	7711	295675046	0.8794535	0.3256207	0	1
hrgender	7711	295675046	0.4699997	0.4991315	0	1
hrbornus	7711	295675046	0.8157386	0.3877223	0	1
winter	7711	295675046	0.3710499	0.4831171	0	1

Table 11: Summary statistics for 2007-2008 data.

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
meat	8279	299979897	177.8193	157.9367	0	1651.94
fish	8279	299979897	24.24362	63.06385	0	852.22
dairy	8279	299979897	281.4814	284.3136	0	4460.5
eggs	8279	299979897	22.6436	42.22367	0	503.5
beef	8279	299979897	58.78755	100.2532	0	1651.94
pork	8279	299979897	16.79015	47.34645	0	724.375
poultry	8279	299979897	70.69504	105.8755	0	1163
pcincome	7916	287829512	21262.45	26594.44	135.563	214937.1
age	8279	299979897	36.90317	22.05805	0	80
gender	8279	299979897	0.517274	0.4997317	0	1
hhsize	8279	299979897	3.439501	1.627723	1	7
black	8279	299979897	0.120763	0.3258713	0	1
hispanic	8279	299979897	0.1585297	0.365259	0	1
bornus	8279	299979897	0.8541447	0.3529824	0	1
hrgender	8279	299979897	0.4840872	0.4997769	0	1
hrbornus	8279	299979897	0.7962739	0.402792	0	1
winter	8279	299979897	0.4226746	0.4940144	0	1

Table 12: Summary statistics for 2009-2010 data.

Variable	Obs	Weight	Mean	Std. Dev.	Min	Max
meat	7486	304938168	170.204	147.8465	0	1381.13
fish	7486	304938168	22.66933	62.45235	0	1043.925
dairy	7486	304938168	251.9536	259.697	0	3111
eggs	7486	304938168	24.27473	44.53427	0	700
beef	7486	304938168	58.00561	101.322	0	1381.13
pork	7486	304938168	15.23487	43.18309	0	747.76
poultry	7486	304938168	65.81226	95.39284	0	931.915
pcincome	7194	297082322	16964.13	19316.85	139.8148	187565.6
age	7486	304938168	37.43204	22.1845	0	80
gender	7486	304938168	0.5092513	0.4999478	0	1
hhsize	7486	304938168	3.357985	1.608195	1	7
black	7486	304938168	0.1246058	0.3302934	0	1
hispanic	7486	304938168	0.1666005	0.3726437	0	1
bornus	7486	304938168	0.8556331	0.3514849	0	1
hrgender	7486	304938168	0.451686	0.4976935	0	1
hrbornus	7486	304938168	0.7889985	0.4080467	0	1
winter	7486	304938168	0.4512302	0.4976491	0	1

Table 13: Summary statistics for 2011-2012 data.

	meat	fish	dairy	eggs	beef	pork	poultry	pcincome	age	gender	hhsize	black	hispanic	bornus	hrgender	hrbornus	winter
meat	1																
fish	-0.0166	1															
dairy	-0.0195	-0.0363	1														
eggs	0.0367	0.046	-0.073	1													
beef	0.6296	-0.0177	0.0339	0.0336	1												
pork	0.2564	0.0027	-0.0354	0.0232	-0.0561	1											
poultry	0.5893	-0.0097	-0.0555	0.0135	-0.0466	0.0024	1										
pcincome	0.0259	0.0379	-0.0185	-0.0232	0.0013	0.0363	0.0214	1									
age	0.0515	0.1275	-0.2146	0.0877	0.0397	0.07	-0.0234	0.1181	1								
gender	-0.2255	-0.0715	-0.0866	-0.0809	-0.1422	-0.0873	-0.0794	-0.028	0.0456	1							
hhsize	-0.017	-0.057	0.0806	-0.0313	0.0229	-0.0581	-0.0003	-0.1807	-0.4997	-0.0303	1						
black	0.0335	0.0214	-0.1294	0.0082	-0.0138	-0.0107	0.0461	-0.0815	-0.0736	0.0164	0.0904	1					
hispanic	0.0119	0.0037	0.0006	0.0656	0.0693	-0.0158	-0.0109	-0.1083	-0.1312	-0.0012	0.2573	-0.1408	1				
bornus	-0.0408	-0.0882	0.0557	-0.0417	-0.0423	-0.0234	-0.0349	0.054	-0.0374	0.0068	-0.1612	0.0526	-0.4062	1			
hrgender	-0.0696	-0.04	0.0396	-0.0364	-0.055	-0.0579	-0.0193	-0.0992	-0.0525	0.2012	-0.0171	0.1434	-0.0152	0.0563	1		
hrbornus	-0.0048	-0.0392	-0.0005	-0.0234	-0.0253	-0.0013	-0.0228	0.0644	0.1013	-0.0215	-0.2589	0.0594	-0.4598	0.6736	0.0852	1	
winter	0.0248	-0.0127	-0.0217	0.0307	0.0013	0.0053	0.0477	-0.0081	-0.0141	-0.0119	0.0229	-0.0183	0.2616	-0.1288	0.0167	-0.161	1

Table 14: Correlation table for 2003-2004 data.

	meat	fish	dairy	eggs	beef	pork	poultry	pcincome	age	gender	hhsize	black	hispanic	bornus	hrgender	hrbornus	winter
meat	1																
fish	-0.0453	1															
dairy	-0.0384	-0.0282	1														
eggs	0.0382	0.0104	-0.0083	1													
beef	0.6661	-0.0317	-0.0266	0.0057	1												
pork	0.2804	0.0342	-0.0317	0.0456	0.0183	1											
poultry	0.5654	-0.0453	-0.0085	0.0071	-0.0652	-0.0348	1										
pcincome	0.0095	0.0416	-0.025	0.0034	-0.0191	0.0271	0.0115	1									
age	0.0564	0.1172	-0.2024	0.0731	0.0368	0.0584	-0.0049	0.1209	1								
gender	-0.2342	-0.05	-0.0932	-0.0878	-0.141	-0.1045	-0.0827	-0.0529	0.0522	1							
hhsize	-0.012	-0.0366	0.133	-0.0207	0.004	-0.0105	0.0012	-0.2188	-0.5029	-0.0407	1						
black	0.0268	0.0027	-0.1268	0.005	-0.0315	-0.003	0.0675	-0.1026	-0.0628	0.0285	0.0336	1					
hispanic	0.0038	-0.0082	0.0146	0.0451	0.0319	-0.032	0.0138	-0.0917	-0.1427	-0.0046	0.2085	-0.1377	1				
bornus	-0.0338	-0.0598	0.0542	-0.0173	-0.0093	-0.0282	-0.0606	-0.0047	-0.0222	0.0124	-0.1104	0.0522	-0.3754	1			
hrgender	-0.0732	-0.0357	-0.0052	-0.0077	-0.0553	-0.0189	-0.0248	-0.0614	-0.0034	0.2232	-0.0698	0.1124	-0.0111	0.0249	1		
hrbornus	-0.0113	-0.0263	0.0031	-0.0054	-0.0117	-0.0177	-0.0368	0.0234	0.1103	-0.0181	-0.1758	0.0496	-0.4343	0.6793	0.0664	1	
winter	0.0247	-0.0247	-0.1159	0.0275	-0.0091	0.0016	0.0483	-0.0149	-0.0256	0.0158	0.0056	0.0898	0.174	-0.0839	-0.0139	-0.0936	1

Table 15: Correlation table for 2005-2006 data.

	meat	fish	dairy	eggs	beef	pork	poultry	pcincome	age	gender	hhsize	black	hispanic	bornus	hrgender	hrbornus	winter
meat	1																
fish	-0.0262	1															
dairy	-0.0511	-0.0476	1														
eggs	0.0245	0.0708	-0.0391	1													
beef	0.617	-0.0281	-0.0245	0.0107	1												
pork	0.311	0.0324	-0.0115	0.0338	0.0557	1											
poultry	0.6281	-0.024	-0.0418	-0.0003	-0.0407	-0.0304	1										
pcincome	0.0461	0.0644	0.0074	-0.0226	0.0402	0.0232	0.0178	1									
age	0.0736	0.1038	-0.2065	0.0922	0.0578	0.0586	0.0261	0.1158	1								
gender	-0.2157	-0.0693	-0.0716	-0.119	-0.1278	-0.0845	-0.0885	-0.0246	0.0399	1							
hhsize	-0.055	-0.0631	0.0969	-0.0208	-0.0482	-0.0178	-0.029	-0.195	-0.5064	-0.0281	1						
black	0.0092	0.0037	-0.1225	-0.0029	-0.0281	-0.008	0.0262	-0.1095	-0.07	0.0203	0.0273	1					
hispanic	-0.0129	-0.0053	0.0347	0.0866	0.0001	-0.0189	0.0235	-0.1705	-0.1433	-0.021	0.2777	-0.1562	1				
bornus	-0.0049	-0.0924	0.0334	-0.071	0.0105	-0.0099	-0.0456	0.0587	-0.0555	0.0315	-0.1113	0.0906	-0.4122	1			
hrgender	-0.0764	-0.0442	-0.0244	-0.0586	-0.061	-0.0167	-0.0229	-0.1119	-0.0274	0.2126	0.0067	0.1118	-0.0088	0.0336	1		
hrbornus	0.0433	-0.0491	-0.0148	-0.0618	0.033	0.0014	-0.0102	0.0795	0.1159	0.0195	-0.2312	0.0844	-0.4798	0.655	0.0492	1	
winter	0.0178	0.0499	-0.0603	0.0821	-0.0111	-0.0464	0.0777	0.0521	-0.065	-0.0293	0.072	0.1651	0.2349	-0.1746	-0.0653	-0.2198	1

Table 16: Correlation table for 2007-2008 data.

	meat	fish	dairy	eggs	beef	pork	poultry	pcincome	age	gender	hhsize	black	hispanic	bornus	hrgender	hrbornus	winter
meat	1																
fish	-0.0262	1															
dairy	-0.0477	-0.0715	1														
eggs	0.0555	0.0261	-0.0581	1													
beef	0.598	-0.0079	-0.0312	0.0293	1												
pork	0.2936	0.0204	-0.0004	0.0468	0.0064	1											
poultry	0.6182	-0.0359	-0.0311	0.0274	-0.0544	-0.0317	1										
pcincome	0.0487	0.0718	0.008	-0.0084	0.0038	-0.0155	0.0731	1									
age	0.0826	0.1325	-0.2166	0.0501	0.0726	0.0617	0.0095	0.1413	1								
gender	-0.2393	-0.0497	-0.0678	-0.0899	-0.138	-0.1132	-0.1	-0.0535	0.0356	1							
hhsize	-0.0415	-0.0852	0.1079	-0.0098	-0.0319	-0.0256	-0.0055	-0.1653	-0.5024	-0.0129	1						
black	-0.0017	0.008	-0.1321	0.0156	-0.024	-0.0341	0.0328	-0.1383	-0.0554	0.0315	0.0214	1					
hispanic	0.0048	-0.0207	0.008	0.0886	0.0181	-0.0222	0.0458	-0.173	-0.1587	-0.0148	0.2935	-0.1566	1				
bornus	-0.0587	-0.0908	0.0782	-0.0424	-0.0031	-0.0604	-0.0874	0.0537	-0.0582	0.0262	-0.1337	0.05	-0.3504	1			
hrgender	-0.0528	0.0192	0.0014	-0.0275	-0.0059	-0.0165	-0.0434	-0.0707	-0.0508	0.2049	0.0253	0.1011	-0.0019	0.0201	1		
hrbornus	-0.0394	-0.0337	0.0094	-0.019	0.0073	-0.0433	-0.0842	0.0756	0.1194	0.0138	-0.23	0.0486	-0.451	0.6791	0.0456	1	
winter	0.0151	0.0125	-0.0675	0.0379	0.0014	0.0054	0.0409	-0.0822	-0.02	0.0269	0.1175	0.0681	0.2275	-0.122	0.0536	-0.1379	1

Table 17: Correlation table for 2009-2010 data.

	meat	fish	dairy	eggs	beef	pork	poultry	pcincome	age	gender	hhsize	black	hispanic	bornus	hrgender	hrbornus	winter
meat	1																
fish	-0.0491	1															
dairy	-0.0945	-0.0501	1														
eggs	0.0546	0.0118	-0.0234	1													
beef	0.6425	-0.0246	-0.0516	0.0256	1												
pork	0.2792	-0.0097	-0.0373	0.0587	-0.0032	1											
poultry	0.5731	-0.0281	-0.0536	0.0039	-0.0629	-0.0093	1										
pcincome	-0.0353	0.0371	-0.0196	-0.0054	-0.0566	-0.0195	0.0176	1									
age	0.0729	0.1201	-0.2004	0.0779	0.0544	0.0552	0.0132	0.1523	1								
gender	-0.1947	-0.0294	-0.0971	-0.0753	-0.1268	-0.0803	-0.0549	-0.0136	0.0463	1							
hhsize	-0.0175	-0.0334	0.0913	-0.0274	-0.0155	-0.0142	-0.009	-0.2379	-0.4825	-0.0514	1						
black	0.0477	0.0268	-0.1298	-0.0001	-0.0357	-0.0032	0.0748	-0.122	-0.065	0.0367	0.0461	1					
hispanic	-0.0098	-0.009	0.0176	0.0531	0.0168	-0.0573	0.0227	-0.1729	-0.1641	-0.0195	0.2135	-0.1645	1				
bornus	-0.0322	-0.0718	0.0516	-0.0335	-0.0311	-0.0131	-0.0549	0.0826	-0.0454	0.0011	-0.1047	0.0798	-0.374	1			
hrgender	-0.0696	-0.0097	-0.0339	-0.0385	-0.0517	-0.0289	-0.0492	-0.1073	-0.0344	0.2381	-0.0219	0.1214	0.0019	0.0481	1		
hrbornus	0.0116	-0.0402	0.004	-0.0303	0.0041	0.0348	-0.0463	0.1097	0.0973	0.0084	-0.2017	0.0744	-0.4502	0.6619	0.0751	1	
winter	0.0506	0.0103	-0.0197	0.0165	0.0464	-0.0099	0.0417	-0.1102	-0.0769	-0.019	0.0586	0.0188	0.1914	-0.0538	-0.0094	-0.0641	1

Table 18: Correlation table for 2011-2012 data.

References

AGRI
2012. Prospects for agricultural markets and income in the EU 2012-2022. *Directorate-General for Agriculture and Rural Development, European Commission (Brussels).*

Alexandratos, N. and J. Bruinsma
2012. World agriculture towards 2030/2050: the 2012 revision. ESA Working Paper 12-03, Food and Agriculture Organization of the United Nations.

Arganini, C., A. Turrini, A. Saba, F. Virgili, and R. Comitato
2012. *Gender differences in food choice and dietary intake in modern western societies*. INTECH Open Access Publisher.

Atkins, P., I. Bowler, et al.
2001. *Food in society: economy, culture, geography.* Arnold, Hodder Headline Group.

Atkinson, A. B., L. Rainwater, and T. M. Smeeding
1995. *Income distribution in OECD countries: OECD Social Policy Studies*. Organisation for Economic Co-operation and Development.

Bailey, R., A. Froggatt, and L. Wellesley
2014. Livestock: Climate change's forgotten sector. Global public opinion on meat and dairy consumption. Research paper, Energy, Environment and Resources, The Royal Institute of International Affairs, Chatham House.

Barnard, N. D., A. Nicholson, and J. L. Howard
1995. The medical costs attributable to meat consumption. *Preventive Medicine*, 24(6):646–655.

Becker, G. S. and K. M. Murphy
1988. A theory of rational addiction. *The journal of political economy*, Pp. 675–700.

Blisard, N., J. N. Variyam, and J. Cromartie
2003. Food expenditures by US households: looking ahead to 2020. Technical report, United States Department of Agriculture, Economic Research Service.

Bloomberg School of Public Health
2013. Health and environmental implications of US meat consumption and production. Accessed: 10-04-2015.

Blundell, R., A. Duncan, and K. Pendakur
1998. Semiparametric estimation and consumer demand. *Journal of Applied Econometrics*, 13(5):435–461.

Browning, M., F. Bourguignon, P.-A. Chiappori, and V. Lechene
1994. Income and outcomes: A structural model of intrahousehold allocation. *Journal of political Economy*, Pp. 1067–1096.

Cameron, A. C. and P. K. Trivedi
2005. *Microeconometrics: methods and applications.* Cambridge university press.

Chai, A. and A. Moneta
2010. Retrospectives engel curves. *The Journal of Economic Perspectives*, 24(1):225–240.

Cirera, X. and E. Masset
2010. Income distribution trends and future food demand. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554):2821–2834.

Curtin, L., L. Mohadjer, S. Dohrmann, D. Kruszon-Moran, L. Mirel, M. Carroll, R. Hirsch, V. Burt, and C. Johnson
2013. National health and nutrition examination survey: sample design, 2007-2010. *Vital and health statistics. Series 2, Data evaluation and methods research*, (160):1–32.

Daniel, C. R., A. J. Cross, C. Koebnick, and R. Sinha
2011. Trends in meat consumption in the USA. *Public health nutrition*, 14(04):575–583.

Deaton, A.
1985. Panel data from time series of cross-sections. *Journal of econometrics*, 30(1):109–126.

Delgado, C. L., M. W. Rosegrant, and S. Meijer
2001. Livestock to 2020: The revolution continues. Paper provided by International Agricultural Trade Research Consortium in its series International Trade in Livestock Products Symposium, January 18-19, 2001, Auckland, New Zealand.

Delgado, C. L., M. W. Rosegrant, H. Steinfeld, S. K. Ehui, and C. Courbois
1999. Livestock to 2020: The next food revolution. Food, Agriculture and the Environment Discussion Paper 28, International Food Policy Research Institute.

DGAC
2015. *Scientific Report of the 2015 Dietary Guidelines Advisory Committee.*

Drewnowski, A. and B. M. Popkin
1997. The nutrition transition: new trends in the global diet. *Nutrition reviews*, 55(2):31–43.

Du, S., T. A. Mroz, F. Zhai, and B. M. Popkin
2004. Rapid income growth adversely affects diet quality in china—particularly for the poor! *Social science & medicine*, 59(7):1505–1515.

Edenhofer, O., R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, A. Baum, A. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel, and J. C. Minx
2014. *Climate change 2014: Mitigation of climate change.*, Contribution of working group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK, and New York: Cambridge University Press.

Engel, E.
1857. Die productions-und consumtions-verhältnisse des königreichs sachsen. *Bulletin de l'Institut International de la Statistique*, 9.

Engel, J. and A. Kneip
1996. Recent approaches to estimating engel curves. *Journal of Economics*, 63(2):187–212.

FAO
2009. The state of food and agriculture 2009: Livestock in the balance. Technical report, Food and Agriculture Organization of the United Nations.

Figari, F., A. Paulus, and H. Sutherland
2014. Microsimulation and policy analysis. *Handbook of Income Distribution*, 2.

Finkelstein, E. A., J. G. Trogdon, J. W. Cohen, and W. Dietz
2009. Annual medical spending attributable to obesity: payer and service specific estimates. *Health affairs*, 28(5):w822–w831.

Frederick, S., G. Loewenstein, and T. O'donoghue
2002. Time discounting and time preference: A critical review. *Journal of economic literature*, Pp. 351–401.

Fukunaga, K. and L. Hostetler
1975. The estimation of the gradient of a density function, with applications in pattern recognition. *Information Theory, IEEE Transactions on*, 21(1):32–40.

Garnett, T.
2013. Food sustainability: problems, perspectives and solutions. *Proceedings of the Nutrition Society*, 72(01):29–39.

Gerber, P. J., H. Steinfeld, B. Henderson, A. Mottet, C. Opio, J. Dijkman, A. Falcucci, G. Tempio, et al.
2013. *Tackling climate change through livestock: a global assessment of emissions and mitigation opportunities.* Food and Agriculture Organization of the United Nations (FAO).

Gneezy, U., S. Meier, and P. Rey-Biel
2011. When and why incentives (don't) work to modify behavior. *The Journal of Economic Perspectives*, Pp. 191–209.

González, A. D., B. Frostell, and A. Carlsson-Kanyama
2011. Protein efficiency per unit energy and per unit greenhouse gas emissions: potential contribution of diet choices to climate change mitigation. *Food Policy*, 36(5):562–570.

Goodland, R.
2001. The westernization of diets: The assessment of impacts in developing countries. *Global Hunger Alliance.*

Gozalo, P. L.
1997. Nonparametric bootstrap analysis with applications to demographic effects in demand functions. *Journal of Econometrics*, 81(2):357–393.

Hagenaars, A. J., K. De Vos, and M. A. Zaidi
1996. *Poverty statistics in the late 1980s: Research based on micro-data.* Office for Official Publications of the European Communities.

Hastie, T. and R. Tibshirani
1986. Generalized additive models. *Statistical science*, Pp. 297–310.

Hastie, T. and R. Tibshirani
1987. Generalized additive models: some applications. *Journal of the American Statistical Association*, 82(398):371–386.

Hawkesworth, S., A. D. Dangour, D. Johnston, K. Lock, N. Poole, J. Rushton, R. Uauy, and J. Waage
2010. Feeding the world healthily: the challenge of measuring the effects of agriculture on health. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1554):3083–3097.

Hedenus, F., S. Wirsenius, and D. J. Johansson
2014. The importance of reduced meat and dairy consumption for meeting stringent climate change targets. *Climatic change*, 124(1-2):79–91.

Heller, M. C. and G. A. Keoleian
2014. Greenhouse gas emission estimates of US dietary choices and food loss. *Journal of Industrial Ecology*.

Heller, M. C., G. A. Keoleian, and W. C. Willett
2013. Toward a life cycle-based, diet-level framework for food environmental impact and nutritional quality assessment: A critical review. *Environmental science & technology*, 47(22):12632–12647.

Hertwich, E.
2010. *Assessing the environmental impacts of consumption and production: priority products and materials.* UNEP/Earthprint.

Hoekstra, A. Y. and A. K. Chapagain
2007. Water footprints of nations: water use by people as a function of their consumption pattern. *Water resources management*, 21(1):35–48.

Hoekstra, A. Y. and M. M. Mekonnen
2012. The water footprint of humanity. *Proceedings of the National Academy of Sciences*, 109(9):3232–3237.

Johnson, C. L., R. Paulose-Ram, C. L. Ogden, M. D. Carroll, D. Kruszon-Moran, S. M. Dohrmann, and L. Curtin
2013. National health and nutrition examination survey: analytic guidelines, 1999-2010. *Vital and health statistics. Series 2, Data evaluation and methods research*, (161):1–24.

Kalton, G.
1983. Compensating for missing survey data.

Kaluza, J., A. Wolk, and S. C. Larsson
2012. Red meat consumption and risk of stroke a meta-analysis of prospective studies. *Stroke*, 43(10):2556–2560.

Kearney, J.
2010. Food consumption trends and drivers. *Philosophical transactions of the royal society B: biological sciences*, 365(1554):2793–2807.

Kreith, M. and C. Davis
1991. *Water inputs in California food production.* Water Education Foundation.

Li, J. and C. O'Donoghue
2013. A survey of dynamic microsimulation models: uses, model structure and methodology. *International Journal of Microsimulation*, 6(2):3–55.

Lin, B.-H., J. N. Variyam, J. E. Allshouse, and J. Cromartie
2003. Food and agricultural commodity consumption in the United States: Looking ahead to 2020. Technical report, United States Department of Agriculture, Economic Research Service.

Mallin, M. A. and L. B. Cahoon
2003. Industrialized animal production—a major source of nutrient and microbial pollution to aquatic ecosystems. *Population and Environment*, 24(5):369–385.

Manig, C. and A. Moneta
2014. More or better? measuring quality versus quantity in food consumption. *Journal of Bioeconomics*, 16(2):155–178.

Marmot, M., T. Atinmo, T. Byers, J. Chen, T. Hirohata, A. Jackson, W. James, L. Kolonel, S. Kumanyika, C. Leitzmann, et al.
2007. Food, nutrition, physical activity, and the prevention of cancer: a global perspective.

McDonald, J. B. and M. R. Ransom
1979. Functional forms, estimation techniques and the distribution of income. *Econometrica: Journal of the Econometric Society*, Pp. 1513–1525.

McGuire, S.
2011. US department of agriculture and US department of health and human services, dietary guidelines for americans, 2010. *Advances in Nutrition: An International Review Journal*, 2(3):293–294.

Metz, B., O. R. Davidson, R. Bosch, P R andDave, and L. A. Meyer
2007. *Climate change 2007: Mitigation of climate change.*, Contribution of working group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK, and New York: Cambridge University Press.

Micha, R., S. K. Wallace, and D. Mozaffarian
2010. Red and processed meat consumption and risk of incident coronary heart disease, stroke, and diabetes mellitus a systematic review and meta-analysis. *Circulation*, 121(21):2271–2283.

Mirel, L. B., L. K. Mohadjer, S. M. Dohrmann, J. Clark, V. L. Burt, C. L. Johnson, and L. R. Curtin
2013. National health and nutrition examination survey: estimation procedures, 2007-2010. *Vital and health statistics. Series 2, Data evaluation and methods research*, (159):1–25.

Nijdam, D., T. Rood, and H. Westhoek
2012. The price of protein: Review of land use and carbon footprints from life cycle assessments of animal food products and their substitutes. *Food Policy*, 37(6):760–770.

NU-AGE
2012. Report on the socio-economic determinants of food choices and preferences of the elderly. Nu-age report, European Commission.

Pan, A., Q. Sun, A. M. Bernstein, M. B. Schulze, J. E. Manson, M. J. Stampfer, W. C. Willett, and F. B. Hu
2012. Red meat consumption and mortality: results from 2 prospective cohort studies. *Archives of internal medicine*, 172(7):555–563.

Pesaran, M. H. and P. Schmidt
1995. *Handbook of applied econometrics: Microeconomics*, volume II. Blackwell.

Pingali, P.
2007. Westernization of asian diets and the transformation of food systems: implications for research and policy. *Food policy*, 32(3):281–298.

Popkin, B. M.
1994. The nutrition transition in low-income countries: an emerging crisis. *Nutrition reviews*, 52(9):285–298.

Popkin, B. M.
2002. The shift in stages of the nutrition transition in the developing world differs from past experiences! *Public health nutrition*, 5(1A):205–214.

Popkin, B. M., S. Horton, S. Kim, A. Mahal, and J. Shuigao
2001. Trends in diet, nutritional status, and diet-related noncommunicable diseases in china and india: the economic costs of the nutrition transition. *Nutrition reviews*, 59(12):379–390.

Prättälä, R., L. Paalanen, D. Grinberga, V. Helasoja, A. Kasmel, and J. Petkeviciene
2007. Gender differences in the consumption of meat, fruit and vegetables are similar in finland and the baltic countries. *The European Journal of Public Health*, 17(5):520–525.

Salem, A. and T. Mount
1974. A convenient descriptive model of income distribution: the gamma density. *Econometrica: Journal of the Econometric Society*, Pp. 1115–1127.

Silbergeld, E. K., J. Graham, and L. B. Price
2008. Industrial food animal production, antimicrobial resistance, and human health. *Annu. Rev. Public Health*, 29:151–169.

Silverman, B. W.
1986. *Density estimation for statistics and data analysis*, volume 26. CRC press.

Sinha, R., A. J. Cross, B. I. Graubard, M. F. Leitzmann, and A. Schatzkin
2009. Meat intake and mortality: a prospective study of over half a million people. *Archives of internal medicine*, 169(6):562–571.

Smil, V.
2002. Worldwide transformation of diets, burdens of meat production and opportunities for novel food proteins. *Enzyme and Microbial Technology*, 30(3):305–311.

Stehfest, E., L. Bouwman, D. P. van Vuuren, M. G. den Elzen, B. Eickhout, and P. Kabat
2009. Climate benefits of changing diet. *Climatic change*, 95(1-2):83–102.

Steinfeld, H., P. Gerber, T. Wassenaar, V. Castel, M. Rosales, C. d. Haan, et al.
2006. *Livestock's long shadow: environmental issues and options.* Food and Agriculture Organization of the United Nations (FAO).

Toomet, O. and A. Henningsen
2008. Sample selection models in r: Packagesampleselection. *Journal of statistical software*, 27(7):1–23.

Tukker, A., R. A. Goldbohm, A. De Koning, M. Verheijden, R. Kleijn, O. Wolf, I. Pérez-Domínguez, and J. M. Rueda-Cantuche
2011. Environmental impacts of changes to healthier diets in europe. *Ecological Economics*, 70(10):1776–1788.

USDA
. *Food Code Numbers and the Food Coding Scheme.* United States Department of Agriculture.

USDA
2010. Usda agricultural projections to 2021.

Verbeek, M.
2008. Pseudo-panels and repeated cross-sections. In *The Econometrics of Panel Data*, Pp. 369–383. Springer.

Verbeke, W.
2005. Consumer acceptance of functional foods: socio-demographic, cognitive and attitudinal determinants. *Food quality and preference*, 16(1):45–57.

Walker, P., P. Rhubart-Berg, S. McKenzie, K. Kelling, and R. S. Lawrence
2005. Public health implications of meat production and consumption. *Public health nutrition*, 8(04):348–356.

Wang, Y. and M. A. Beydoun
2009. Meat consumption is associated with obesity and central obesity among US adults. *International Journal of Obesity*, 33(6):621–628.

Wood, S.
2006. *Generalized additive models: an introduction with R.* CRC press.

Zapata, H. O. and A. K. Sulgham
2006. A semiparametric approach to estimate engel curves using the US micro data. In *2006 Annual meeting, July 23-26, Long Beach, CA*, number 21092. American Agricultural Economics Association (New Name 2008: Agricultural and Applied Economics Association).