Gluttony and Sloth:

Long-run changes in bodyweight, diet and activity*

Rachel Griffith † Rodrigo Lluberas ‡ Melanie Luhrmann §

March 30, 2015

Abstract

PRELIMINARY AND INCOMPLETE - PLEASE DO NOT CITE

There have been large increases in bodyweight in most developed countries. The economics literature has largely focused on increased calorie consumption as the cause. However, household surveys suggest reduced calorie consumption. We compile new comprehensive data on calories purchased and calories burnt in activity to show that in England there has been a substantial *decrease* in total calories purchased, despite an increase in some high calorie categories like fast food, snacks and drinks. Time use has changed in important ways, with work and other activities becoming less strenuous. We also show important age variation in these trends. We combine these data in a bio-physical model of bodyweight and show that together these data provide an accurate description of how the decline in *both* calories and the strenuousness of activity has resulted in weight gain.

*We gratefully acknowledge financial support from the European Research Council (ERC) under ERC-2009-AdG grant agreement number 249529 and the Economic and Social Research Council (ESRC) under, grant number ORA: ES/I012222/1.

†rgriffith@ifs.org.uk; Institute for Fiscal Studies and University of Manchester

‡rodrigo.lluberas.2009@rhul.ac.uk; Institute for Fiscal Studies and Royal Holloway

§melanie.luhrmann@rhul.ac.uk; Institute for Fiscal Studies and Royal Holloway

1 Introduction

The number of individuals that are overweight and obese has increased across the developed world. The economics literature has emphasised an increase in food consumption as the main factor (see, inter alia, Cutler et al. (2003), Bleich et al. (2008), Brunello et al. (2009)), while papers that have suggested that reduced physical activity and technological change also might play an important role have received less attention (Lakdawalla and Philipson (2009)). The medical literature has also highlighted excess food consumption (Swinburn et al. (2009), Duffey and Popkin (2011)), although the recent literature has highlighted a reduction in activity as an important factor (Ladabaum et al. (2014), Sturm and An (2008)). A confounding factor in understanding the factors leading to the rapid rise in obesity has been a lack of good time series data on calories and physical activity.

We make three contributions in this paper. First, we bring together household level expenditure data for England with nutritional information over thirty years to provide a more complete picture of how diets have changed than in the previous literature. We show that, while there has been an increase in some high calorie categories like fast food, snacks and drinks, overall there has been a *decrease* in total calories purchased, and there have been improvements in diet quality along a number of important dimensions. Second, we compile data on time use and physical activity over thirty years from a number of sources and show that these have changed in important ways, with work, travel and other activities becoming less strenuous, meaning that people now burn substantially less in their day to day life than they used to. Third, we show that together these data provide an accurate description of how the decline in *both* calories and the strenuousness of activity has resulted in weight gain.

Dobbs et al. (2014) identify 74 different policy interventions around the world designed to tackle obesity. It is important for policy formation to understanding what factors have driven the rapid rise in obesity. Our results suggest that most people are reducing their caloric intake, but just not by enough to compensate for the decline in the strenuousness of activities; these results do not say that policies focused at the food will not be effective, but they provide a more nuanced picture and suggest that policies targeting physical activity may also play an important and complementary role.

The fact that calories purchased fall over time in household expenditure survey data has been documented before (see Prentice and Jebb (1995)). This is in contrast to what food availability data show (e.g. Cutler

et al. (2003)). However, these household expenditure data do not include some types of food prior to 2001, importantly snacks, drinks and food out, and it has been thought that this the observed decline in calories in food at home can be attributed to a shift away from food at home to snacks and food out.

We use information from other population surveys and imputation methods to impute calories for these missing categories (Rubin (1987), Rubin (1996), Schafer (1997), Schafer (1999) and Royston and White (2011) amongst others; for applications of these methods see Attanasio and Pistaferri (2013)). We show that it is not plausible that these missing food categories explain the differences. The increase in calories purchased from fast food, snacks and drinks, is more than offset by a decrease in calories purchased for home consumption. Around 10% of the decline in calories for consumption at home is due to reduction in the calorie density within food categories, with the rest of the decline being due to shifts across food categories. Most notably, there has been a reduction in sugary products, such as jam and honey, and in fats, such as butter and meat. There has been a shift towards fruit and vegetables, fish and cereals.

This leads to a puzzle, if calories are declining why are people gaining weight? We provide evidence that changes in time use and the strenuousness of activities have led to an even greater reduction in calories expended. Changes in changes in labour force participation and the occupational mix of jobs has resulted in a reduction in the strenuousness of work. There have also been changes in time spent on housework and the strenuousness of these activities, and changes in leisure activities; together these have resulted in reductions in activity levels that plausibly explain the observed weight gain and reduced calories. The picture is somewhat different for males and females. The labour force participation and strenuousness of work for men has declined. While for women there has been an increase in labour market participation, but that has meant a shift towards less strenuous activity, away from more strenuous housework. In terms of leisure time there has been an increase in sedentary activities over time, such as watching TV and sleeping.

A challenge we face in marrying all of these facts up is that weight, calories and time use are not available in the same data. If we are looking at how means evolve this is misleading because there are important correlations between weight, calories and time/energy use that arise from endogenous choices made by individuals. We start by considering a single composite index of time/energy use per individual. We then consider multiple activities; this matters if there are differences in the correlation of the energy intensity of specific activities with calories and/or weight (e.g. if calories and energy intensity of work are positively correlated, but calories and energy intensity of leisure are negatively correlated).

Our work relates to several literatures. The literature from the US suggests that the declining cost of food plays an important role. In a highly cited paper, Cutler et al. (2003) stress changes in technology that have reduced the time and financial costs of calorie consumption, and argue that the amount of calories expended in work and exercise has remained roughly constant. Counter to this Lakdawalla and Philipson (2009) and Lakdawalla et al. (2005) suggest that 60% of weight growth is due to declining physical activity. Lakdawalla and Philipson (2007) suggest that job-related exercise affects weight.

Other work suggests that obesity has increased most among married women, that countries with more regulation supporting traditional agriculture and delivery systems have lower rates of obesity and that sociodemographic changes are important, including female labor force participation and urbanisation (see for instance: Ewing et al. (2003), Cutler et al. (2003), Bleich et al. (2008) and Baum II and Chou (2011)).

Brunello et al. (2009) study the relationship of obesity and family background and find that maternal education and weight are related to obesity among young Europeans. Finkelstein and Zuckerman (2008) argue that the rise in obesity in the US is due to the combination of declining food costs, particularly of processed high-calorie foods, and an increasing use of technology that makes the economy more productive but the population more sedentary. The authors state that this results in an imbalance between calories consumed and calories burned that explains the increase in bodyweight in the US over the last 30 years. Similarly, Philipson et al. (2003) state that obesity is the result of choices that the individual makes and emphasize the impact of the economic environment on these choices. They then argue that the increase in obesity could be explained by the lower cost of food intake and the increasing cost of physical activity due to technological change. Using data from the first three waves of the US National Health and Nutrition Examination Surveys, Rashad (2006) estimate a model of the determinants of adult obesity and finds that caloric intake, physical activity and smoking are determinants of obesity. They find that older people, blacks, Hispanic females and married males are more likely to have a higher BMI. On the other hand, BMI is negatively correlated with education and, only for women, income. The author also states that the increase in the availability of restaurants and the decline in food prices are also important contributors to the increase in obesity. By estimating a cross-country demand system, Seale et al. (2003) analyze how low, middle and high-income countries food consumption respond to changes in income and food prices. They find that both income and price elasticities for food items are larger in poorer countries. Finally, Chou et al. (2004) and Baum II and Chou (2011) suggest that increasing relative price of smoking accounts for 4% of weight

increase.

The structure of the rest of the paper is as follows: Section 2 presents evidence of the obesity pandemic in the UK, with a particular focus on weight gain. Section 3 shows the dynamics in calories, decomposed into food at home, eating out, snacks and takeaways, and alcohol. We show that the microdata employed in this study shows very different trends than previously found using aggregate calorie data. In Section 4, we lay out activity trends in time use, especially in work, and then combine these components in a physical model of weight in Section **??**. Given our partial information on time use of individuals, we derive the strenuousness of non-measured activities (mostly leisure) resulting from the dynamic changes in weight, diet and physical activity at work over time. We

2 Weight gain

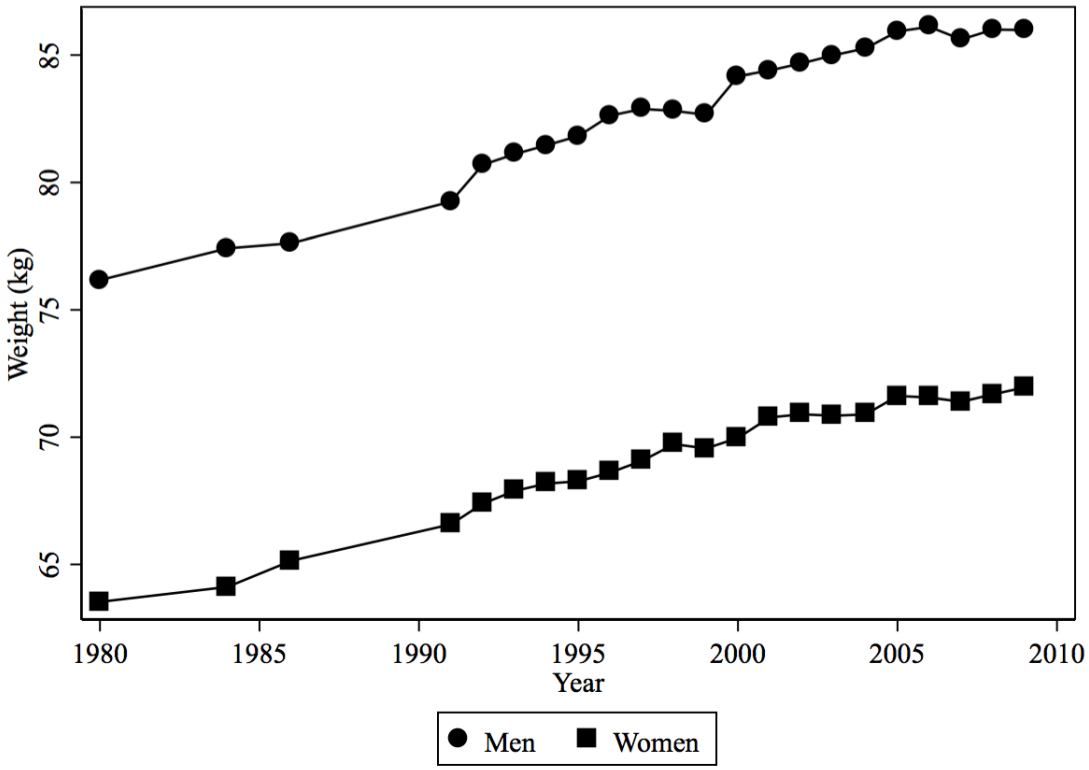
The dramatic increase in bodyweight has been well documented. Figure 2.1 shows the substantial rise in bodyweight from 1980 to 2008 for adult men and women in England; the patterns shown here are similar to those seen in the US (Ladabaum et al. (2014)). The mean weight of an adult male has increased from 75kg to to 86kg, and for adult women from 62kg to 72kg.

Figure 2.2 shows that we see this increase throughout the distribution, with the median also increasing by over 10kg for both men and women; the 10th percentile increasing from 51 to 56 for women and 61 to 69 for men and the 90th percentile shifting from 77 to 92 for women and from 90 to 105 for men. It also shows that we see if for people of different ages. We see similar patterns of weight gain across the socioeconomic distribution (Sturm and An (2008) shows this for the US).

3 Gluttony?

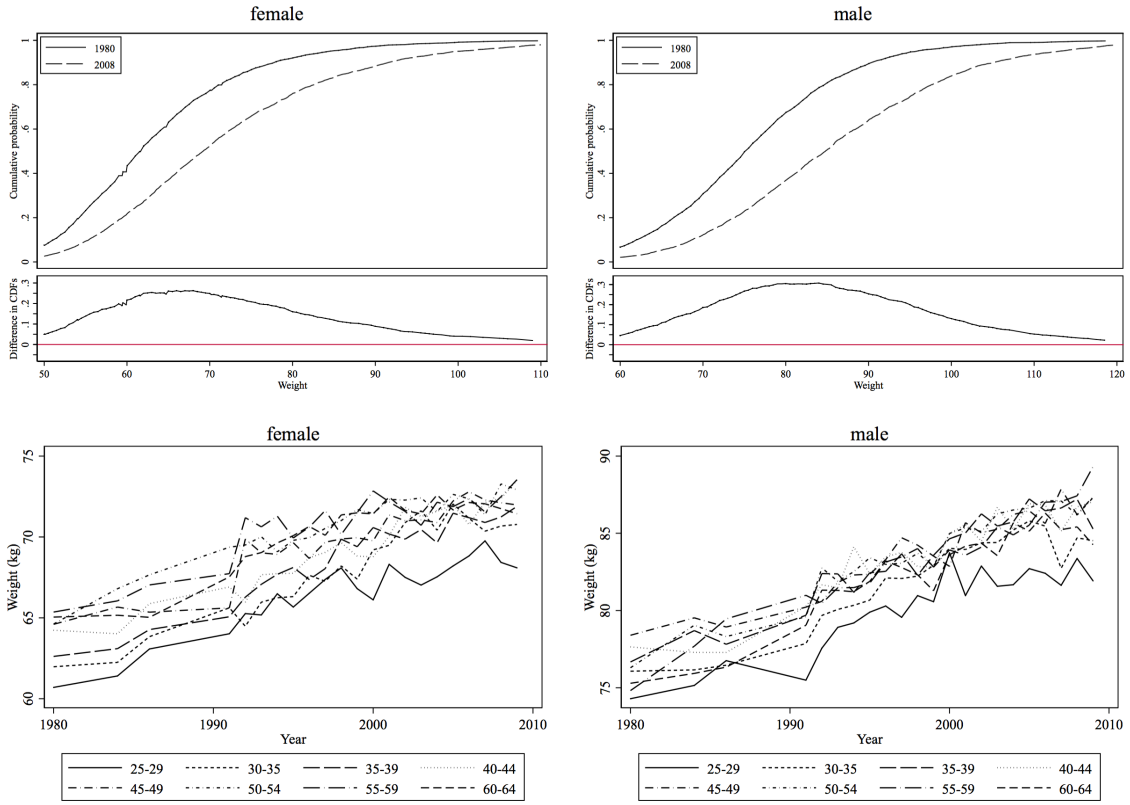
What has been less well documented is how calorie consumption has changed over the same time period. A number of studies have used food availability data; these data suggest an increase in calorie consumption over time (Cutler et al. (2003), Bleich et al. (2008), Brunello et al. (2009)). Food availability data are collected by the United Nations Food and Agriculture Organization (FAO) and the US Department of Agriculture. Food availability is measured as food supply (of each commodity) minus measurable uses, such as farm inputs

Figure 2.1: *Body weight*



Sources: Individuals aged 25-64. Data from Health Survey for England (HSE) for 1992-2009; Health and Lifestyle Survey (HALS) for 1984-85 and 1991-1992; Dietary and Nutritional Survey of British Adults (DNSBA) for 1986-1987; National Heights and Weights survey for 1980.

Figure 2.2: *Body weight*



Sources: Individuals aged 25-64. Data from Health Survey for England (HSE) for 1992-2009; Health and Lifestyle Survey (HALS) for 1984-85 and 1991-1992; Dietary and Nutritional Survey of British Adults (DNSBA) for 1986-1987; National Heights and Weights survey for 1980.

(feed and seed), exports, ending stocks, and industrial uses. See Appendix C.2 for further discussion.

Studies using food availability data have been influential, and policy makers and the media have largely attributed weight gain to increased calorie consumptions (REFS). An alternative source of information on calorie consumption comes from household expenditure surveys. In these data we see the opposite, calories have declined over time, see Figure **??**. There are a number of data issues that we have to tackle in constructing this long time series in calories at the individual level; these are discussed below and in Appendices C and D.

For males the average calories per day implied in UK household budget surveys fall from around 3600 to around 3000, and for females from around 2900 to around 2400. The median fall by a comparable amount, while the 10th percentile falls by 350 calories for females and 400 calories for men, and the 90th percentile falls from 700 calories for females and 900 for males.

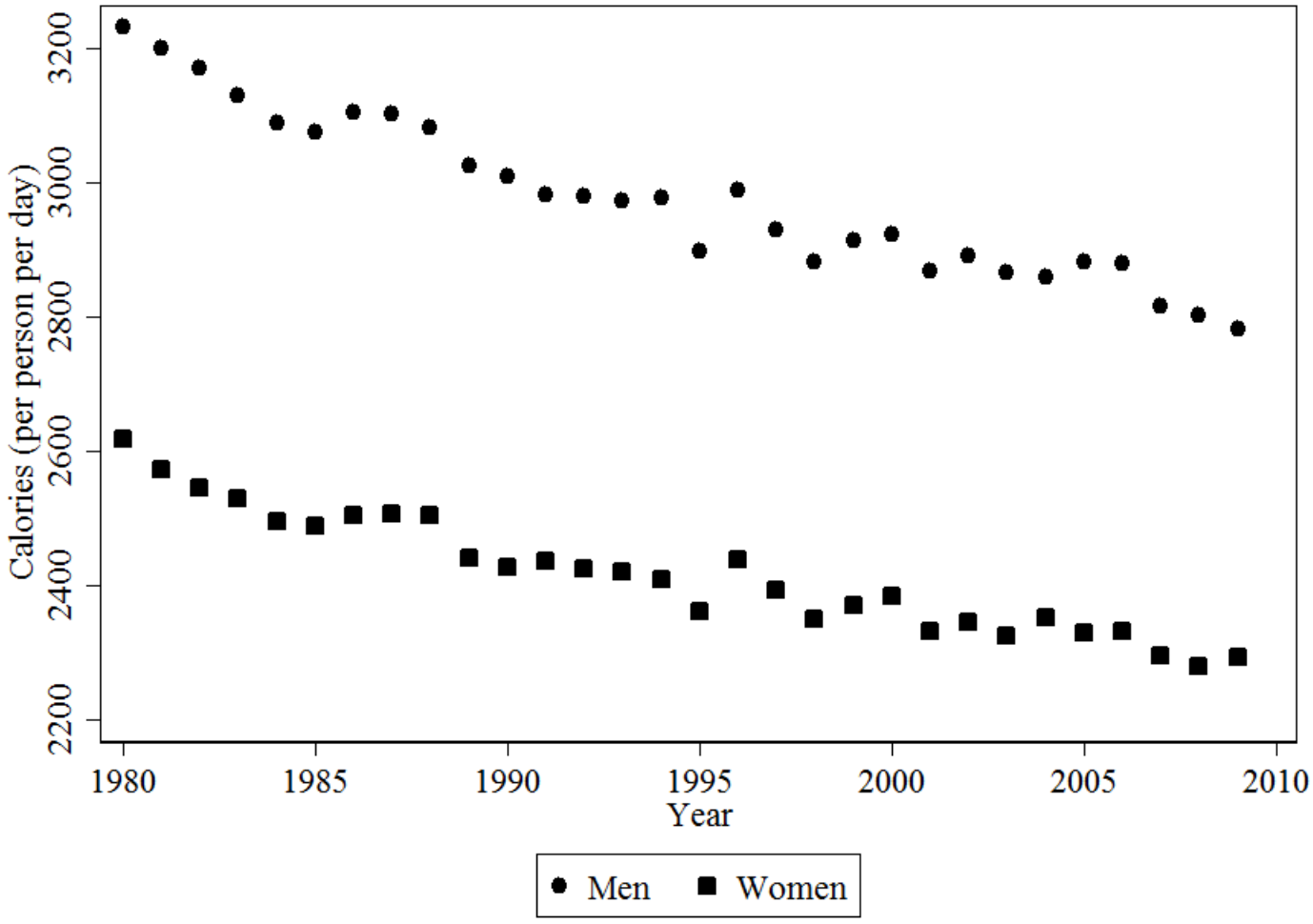
These data on calories from budget surveys seem puzzling given the observed weight gain. We show that the decline in calories is consistent with the observed weight gain when we also take account of changes in time use and activity levels. Before doing that we provide three other pieces of evidence that we think argue in favour of the budget survey data.

1. Previous studies based on UK household survey data used only information on food at home, this is accurately and consistently recorded over time, and the data clearly show that it has declined. However, information on snacks, soft drinks, alcohol and food out at restaurants have not been consistently recorded. These categories have increased in importance, as is clear from expenditure data. Could calories from these groups increased and compensated for the reduction in calories at home? We show that calories from the categories have increased, but not by an order of magnitude that would compensate the decline in food at home calories. We show some summary statistics below, and provide details in Appendix D.

2. We show that the decline in calories at home originates from plausible shifts in diet, away from calorie-rich food categories (butter, meat, jam etc) towards less calorie-dense foods, such as fruit and veg etc..

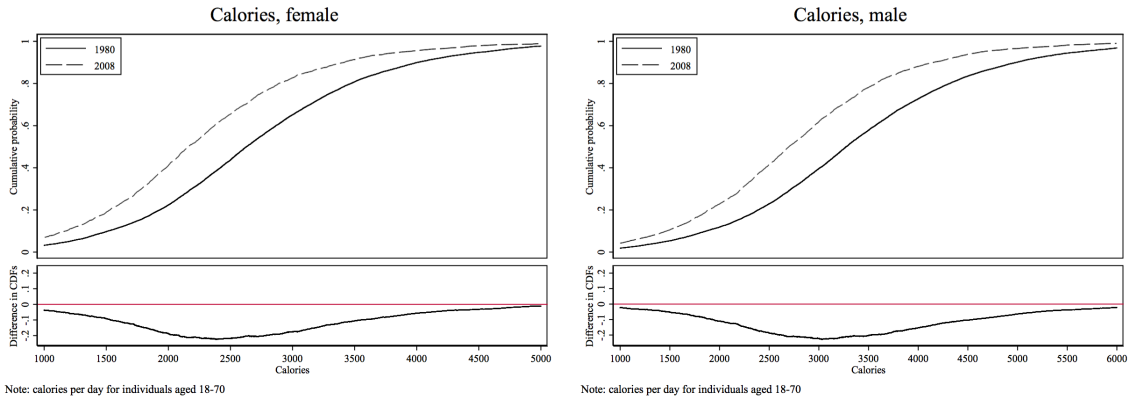
3. We discuss the likely measurement error in food availability data, which rely on macro-economic data of food production, exports and imports that are not well measured. We cite a number of papers that

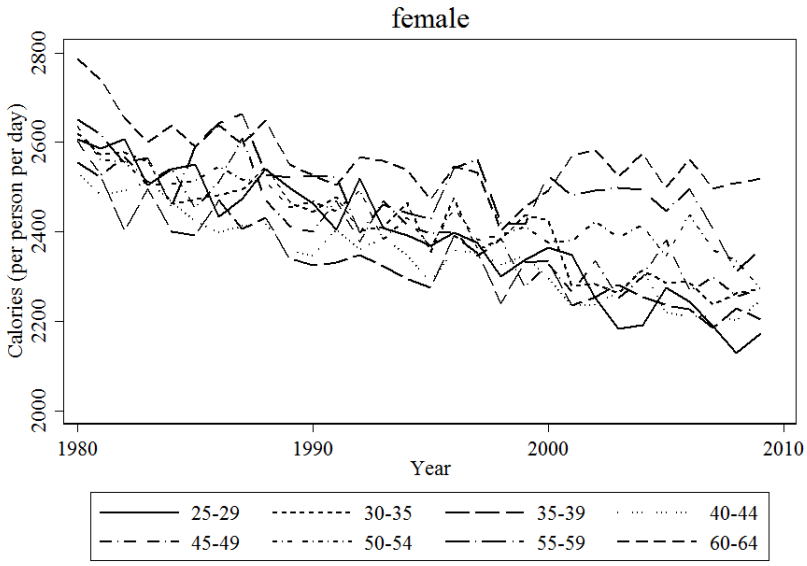
Figure 3.1: *Calories*

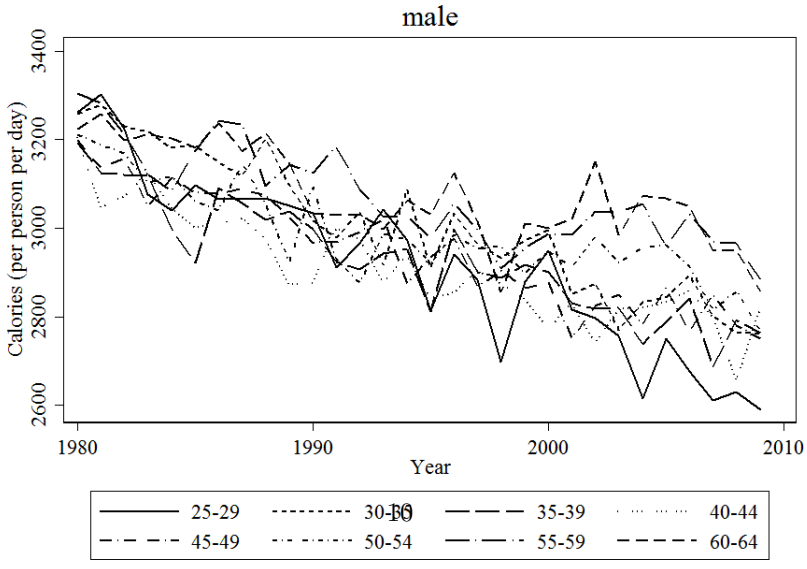


Notes: Individuals aged 25-64. Data from LCFS and imputed as described in Appendix **??**

Figure 3.2: *Calories*

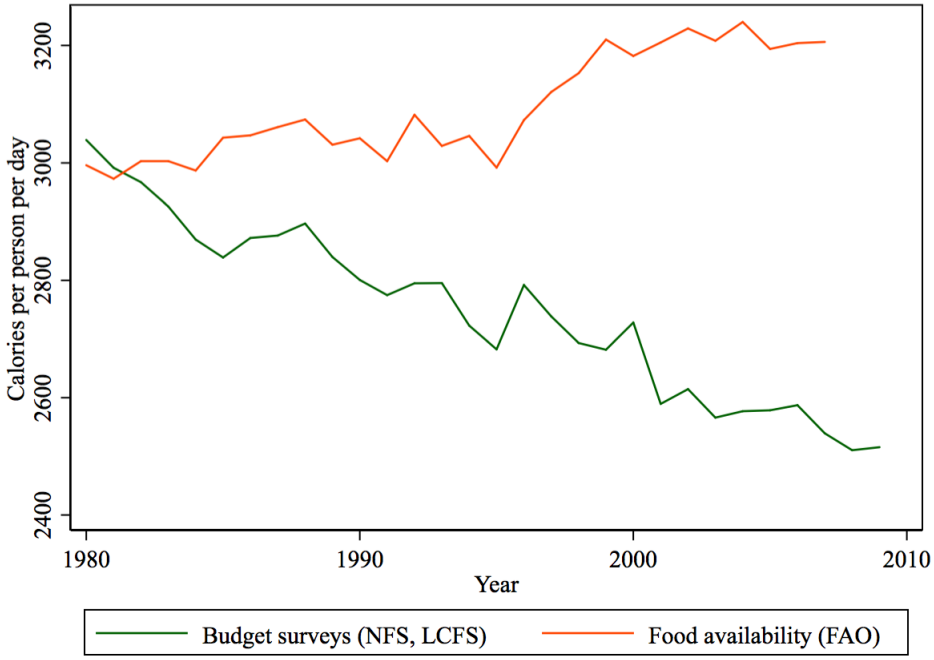






Sources: Individuals aged 25-64. Data from Health Survey for England (HSE) for 1992-2009; Health and Lifestyle Survey (HALS) for 1984-85 and 1991-1992; Dietary and Nutritional Survey of British Adults (DNSBA) for 1986-1987; National Heights and Weights survey for 1980.

Figure 3.3: *Calories per person per day in Food Availability data and Household Survey Data*



Source: xxx.

have raised concerns about these data in relation to other countries; further details are provided in Appendix C.2.

3.1 Calories from snacks, soft drinks, alcohol and food out at restaurants

One of the problems that has confounded research in this area has been the lack of historical data on calories. Current waves of the Living Costs and Food Survey (LCFS) record total food purchases and nutrient quantities for a representative sample of UK households. These data go back to 2001. Prior to that the National Food Survey (NFS) provides information on expenditure, quantities and nutrients for food that is purchased for consumption at home back to 1980 and data on expenditure on food consumed outside of the home and for alcohol (at home and consumed outside the home) are available back to 1980 in the Expenditure and Food Survey (EFS), but we do not have complete information on quantities and nutrients for eating out and alcohol from 1980-2000. We use information on expenditure on these categories, combined with multiple imputation by chained methods (MICE), to impute calories for these missing categories (Rubin (1987), Rubin (1996), Schafer (1997), Schafer (1999) and Royston and White (2011) amongst others; for applications of these methods see Attanasio and Pistaferri (2013) add others e.g. Moffitt?).

A unique feature of our analysis is that we obtain a measure that includes all calories, including food outside of the house and alcoholic drinks. Prentice and Jebb (1995) document the decline in calories from food purchased for consumption at home. However, food at home accounts only for two-thirds of expenditure. We use information on expenditure combined with imputation methods to obtain measures for the missing third of calories - calories from food eaten outside the home and alcoholic drinks.

3.2 Food categories accounting for decline in calories at home

3.3 Measurement error in food availability data

Food availability data records food production and trade and computes food for human consumption as a residual. A number of studies have shown that food availability data overestimates the time series trends in calorie consumption in many countries, e.g. in India (Deaton and Dreze (2010)), Japan (Dowler and Seo (1985)) and the US (Crane et al. (1992)). [1].

[1]See Appendix where we show the trends using microdata versus Food Agriculture Organization aggregate food balance sheets data and also find that levels in 1980 are similar, but trends are massively diverging with a positive FAO trend and a

Table 3.1: *Mean calories per person per day, food at home by food category*

	Males			Females		
	1980	2008	difference	1980	2008	difference
beverages	10	5	-5	12	6	-6
cereals	685	598	-88	820	711	-109
cheese	56	54	-2	66	65	-1
eggs	32	15	-17	38	18	-19
fandv	244	247	3	292	287	-6
fats	308	157	-151	362	186	-175
fish	18	26	9	21	30	9
meat	324	226	-97	388	284	-104
milk	240	159	-81	280	186	-94
other	33	72	39	39	82	43
sugars	203	55	-147	243	64	-179

Source: NFS/FES/LCFS.

4 Sloth?

There have been big changes in the way that people spend their time and the strenuousness of the activities they engage in. Energy expended in an activity depends on the time spent doing the activity, how strenuous it is, and the metabolism of the person engaged in the activity. Table 4.1 show broad patterns of time use in 1983 and 2005 for adult males and females from the Multinational Time Use Survey Data.

Table 4.1: *Time use, % of day spent on activity*

	Males		Females	
	1983	% change 1983-2005	1983	% change 1983-2005
sleep	29.5	14.2	30.6	13.3
work	15.4	-7.6	7.1	23.6
housework	7.5	5.3	15.9	-17.3
travel	4.3	20.5	4.1	16.7
other	43.3	-9.9	42.3	-8.7
of which:				
sports and exercise	1.4	18.0	0.7	64.9
tv	10.0	27.4	8.1	30.4

Source: Multinational Time Use Study, Versions World 5.5.3, 5.80 and 6.0 (released October 2012). Created by Centre for Time Use Research, University of Oxford.

To measure the strenuousness of activities we use the metabolic equivalents of tasks (METs) compiled by Tudor-Locke et al. (2011). METs are the metabolic rate while doing an activity relative to the metabolic rate while resting. A MET is defined as the activity metabolic rate divided by the resting metabolic rate, with lying or sitting quietly classified as 1 MET and a 3 MET activity requires three times the energy expenditure at rest.

Work is important because people spend a lot of time at work, and for many people it is the most strenuous activity they do. A sedentary job that involved sitting at a desk all day, such as a computer

negative NFS one in the UK

programmer, has a MET of 1.5, while an active job, such as a janitor or building cleaner, has a MET of 4.5; an equivalent size person working as a janitor would burn three times as many calories in an hour as if they worked as a computer programmer. For example, an average male working a 40 hour week will burn around 85kg per year in a strenuous occupation, but only 36kg per year in a sedentary occupation; an average female would burn 69kg in a strenuous occupation and around 30kg per year in a sedentary occupation.

4.1 Work

Work is an important use of time, and is one of the more strenuous activities that many individuals engage in, and activity levels have changed over the past three decades. Labour market participation and hours have changed Costa (2000), with females now more likely to be in, and stay in the labour force, and to be working longer hours, while males (at least in some cohorts) have retired early and now work shorter hours. We use data from the Labour Force Survey (LFS) to document these trends.

To obtain a consistent occupation classification coding for 1980-2009 we map detailed occupation codes across time as Nickell and Saleheen (2008) based on the three-digit SOC1990 codes (374 occupations); see Appendix B for details.To get a measure of the strenuousness of different occupations we map these occupation classifications to the occupations detailed in Tudor-Locke et al. (2011), which gives us the MET of each occupation. The 374 three-digit occupations fall into nine broad groups (Managers and Administrators, Professionals, Associate Professionals and Technical Occupations, Clerical and Secretarial Occupations, Craft and Related Occupations, Personal and Protective Services, Sales, Plant and Machine Operatives and Other).

Tables 4.2 and 4.3 shows change in the occupation mix of the workforce by broad occupation. The first column show the average strenuousness of occupations within that broad classification.

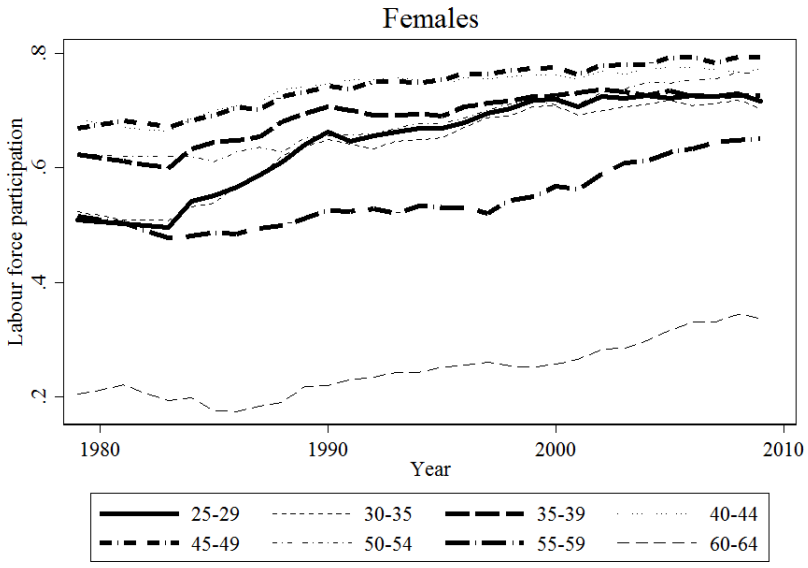
Figure 4.2 show the change in the strenuousness of work, driven by changes in occupation and labour force participation.

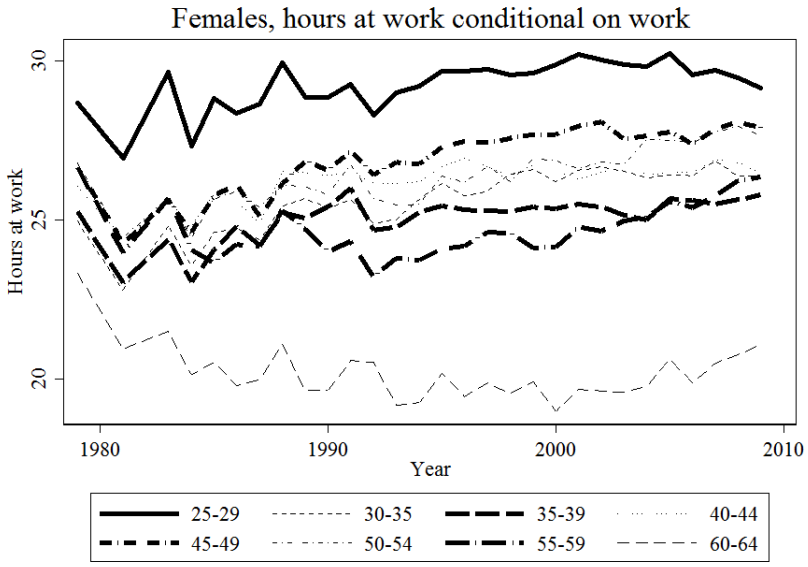
4.2 Other time use

Activities differ in their strenuousness. Lying or sitting quietly has a MET of 1. Housework activities have a MET of between 2.5-6.5, carrying a bag of groceries up a flight of stairs has a MET of 5.

Housework is also an important component of time use and physical activity. We include activities such as cooking, cleaning, shopping, car and home maintenance, gardening, DIY, etc. in housework. We compute

Figure 4.1: *Labour force participation and hours*





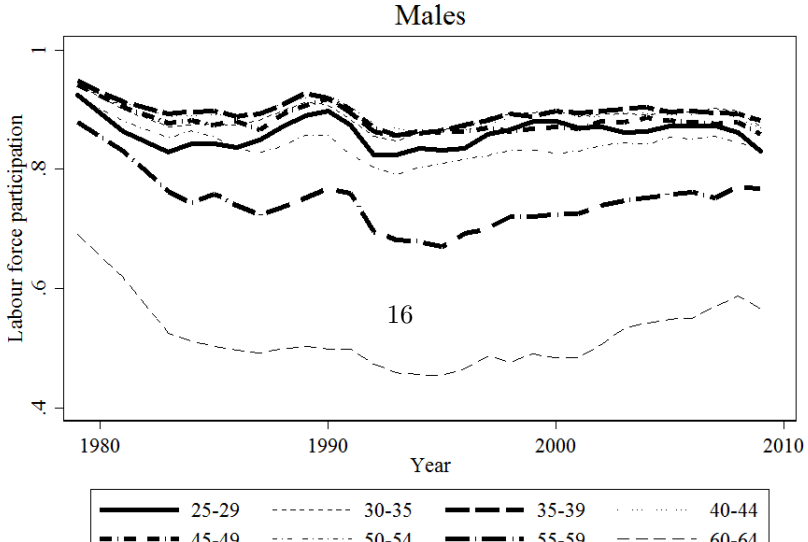


Table 4.2: *METs and share of workforce by occupation, females*

Occupation groups based on SOC1990		% of workforce				
	MET	1981	1990	2000	2008	change 1981-2008
1. Managers and Administrators	1.62	8.94	13.90	17.19	19.74	10.80
4. Clerical and Secretarial	1.72	35.61	30.92	24.10	20.43	-15.19
3. Assoc. Professional and Technical	2.15	4.21	6.11	9.08	10.83	6.62
2. Professional	2.23	5.95	8.83	12.10	13.06	7.11
7. Sales	2.34	8.98	8.49	8.26	6.88	-2.09
6. Personal and Protective Service	2.54	15.28	14.72	17.50	20.01	4.73
5. Craft and Related	2.81	5.65	4.30	2.05	1.49	-4.16
8. Plant and Machine Operatives	2.95	7.89	5.89	4.10	2.47	-5.42
9. Other	3.05	7.38	6.77	5.60	5.07	-2.31

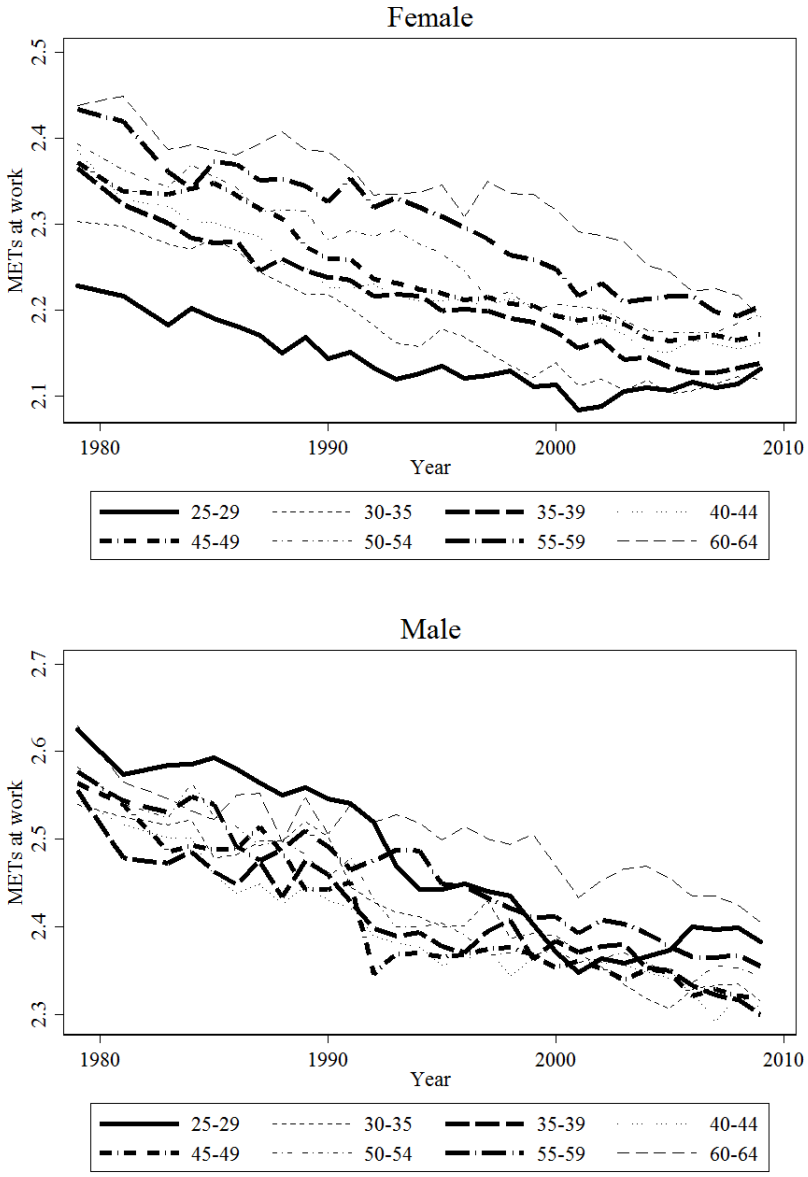
Source: xxx.

Table 4.3: *METs and share of workforce by occupation, males*

Occupation groups based on SOC1990		% of workforce				
	MET	1981	1990	2000	2008	change 1981-2008
1. Managers and Administrators	1.65	16.83	21.43	24.48	27.28	10.46
4. Clerical and Secretarial	1.72	7.29	6.08	5.16	5.09	-2.20
2. Professional	1.94	8.29	10.04	12.64	12.77	4.48
3. Assoc. Professional and Technical	2.14	5.83	7.36	8.80	10.32	4.49
7. Sales	2.14	4.74	4.24	3.74	3.21	-1.53
8. Plant and Machine Operatives	2.47	16.45	14.36	13.88	12.41	-4.03
6. Personal and Protective Service	2.59	5.60	4.75	5.70	5.93	0.33
5. Craft and Related	3.42	27.52	24.82	19.99	17.65	-9.87
9. Other	3.54	7.32	6.49	5.58	5.30	-2.02

Source: xxx.

Figure 4.2: *METs at work, conditional on work*



Notes: Individuals aged 25-64. Data from xxx

the strenuousness of housework based on Ainsworth et al. (2000) and detailed information on time spent on these tasks from the British Time Use Surveys from several years (1974-75, 1983-1984, 1987, 1995, 2000-2001 and 2005).

Due to the expansion of household appliances, housework has become less strenuous over time. For example, between 1991 and 2008 the percentage of households owning a dishwasher increased from 15 to 44% while the proportion using a microwave expanded from 56 to 92%. Women spend on average 3.2 hours per day in 2005 in home production which is manual work and thus strenuous. Young women below age 25 spend 2 hours less per day in home production relative to women aged 26+. Housework time increases substantially at prime child-bearing age and over and remains relatively constant thereafter.

Travel to work and out shopping is also an important activity because people do it regularly every day. Riding in a car has a MET of 1.3, riding a bike to work has a MET of 4, so it requires 4 times the energy expenditure at rest.

Exercise is not a big part of energy expended for most people. While it can be strenuous, for example running a 5-minute mile has a MET of 19, most people do not engage in physical exercise for a long amount of time. For example, jogging has a MET of 6.0, to burn an equivalent amount of energy to working a 40 hour week in exercise, the average person would have to jog for 8.6 hours a week relative to a sedentary job, and 20 hours a week relative to a strenuous job.

5 Analysis

In this section we show that the decline in calories is consistent with the observed weight gain when we also take account of changes in time use and activity levels. We use the first law of thermodynamics applies as an approximation to the way that humans operate syndrome and initial life conditions (rationing, etc.)].

Denote subscripts i: individual, t: time period, j: activity (e.g. occupation, travel mode, etc.) and variables H: hours, C: calories, W: weight, BMR, basal metabolic rate.

The following identity says that weight today equals weight yesterday plus energy consumed minus energy expended:

$$W_{it+1} = W_{it} + C_{it} - \sum_j MET_{ijt} H_{ijt} BMR_{it} \tag{5.1}$$

where

- W_{it+1}: weight in kilograms (kg) at beginning of period $t+1$ (end of period t), W_{it}: weight in kg at beginning of period t (end of period $t-1$);
- C_{it}: calories consumed (purchased) over t expressed in kg (e.g. 7716 calories converts to 1 kg);
- MET_{ijt}: the metabolic equivalent of activity j for individual of time i at time t, a MET is defined as the activity metabolic rate divided by the restÂ-ing metabolic rate, with lying or sitting quietly classified as 1 MET (e.g. a 3-MET activity requires 3 times the energy expenditure at rest);
- H_{ijt} share of hours individual i spent on activity j over t, $\sum_j H_{ijt} = 1$;
- BMR_{it}, basal metabolic rate is the amount of energy expended at rest, i.e. when activity level is $MET = 1$, expressed in kg per year (see appendix F).

What we want to consider is what would the reduction in activity levels need to be in order to justify the observe reduction in calories and weight gain.

We consider two uses of time, work ($j = w$) and other activities ($j = o$). We observe the share of hours spent at work, H_{iwt} and therefore hours in other activities are $H_{iot} = (1 - H_{iwt})$, and 5.1 becomes,

$$W_{it+1} = W_{it} + C_{it} - MET_{iwt}H_{iwt}BMR_{it} - MET_{iot}(1 - H_{iwt})BMR_{it}. \tag{5.2}$$

We observe energy intensity of work, MET_{iwt}, so we can solve for the unknown energy in other activities, MET_{iot},

$$MET_{iot} = \frac{W_{it} - W_{it+1} + C_{it} - MET_{iwt}H_{iwt}BMR_{it}}{(1 - H_{iwt})BMR_{it}} \tag{5.3}$$

$$= \frac{W_{it} - W_{it+1} + C_{it}}{(1 - H_{iwt})BMR_{it}} - \frac{MET_{iwt}H_{iwt}}{(1 - H_{iwt})}$$

$$= 1[H_{iwt} = 0]\left[\frac{W_{it} - W_{it+1} + C_{it}}{BMR_{it}}\right] + 1[H_{iwt} > 0]\left[\frac{W_{it} - W_{it+1} + C_{it}}{(1 - H_{iwt})BMR_{it}} - \frac{MET_{iwt}H_{iwt}}{(1 - H_{iwt})}\right]$$

Using data on C_{it}, W_{it}, and W_{it+1} we can compute the unknown MET, which gives us an index of the strenuousness of combined time use. However, we do not observe weight in period t and period $t+1$ for

the same individual, nor do we observe W_{it} and C_{it} for the same individual over this time period. We do observe these data in population surveys. Taking expectations we obtain (see Appendix E),

Taking expectations,

$$E\left[MET_{iot}\right] = E\left[\frac{W_{it} - W_{it+1} + C_{it}}{BMR_{it}}|H_{iwt} = 0\right] + E\left[\frac{W_{it} - W_{it+1} + C_{it}}{(1 - H_{iwt})BMR_{it}}|H_{iwt} > 0\right] - E\left[\frac{MET_{iwt}H_{iwt}}{(1 - H_{iwt})}|H_{iwt} > 0\right] \quad (5.4)$$

$$= E\left[\frac{W_{it}}{BMR_{it}}|H_{iwt} = 0\right] - E\left[\frac{W_{it+1}}{BMR_{it}}|H_{iwt} = 0\right] + E\left[\frac{C_{it}}{BMR_{it}}|H_{iwt} = 0\right]$$

$$+ E\left[\frac{1}{(1 - H_{iwt})}\frac{W_{it}}{BMR_{it}}|H_{iwt} > 0\right] - E\left[\frac{1}{(1 - H_{iwt})}\frac{W_{it+1}}{BMR_{it}}|H_{iwt} > 0\right] + E\left[\frac{1}{(1 - H_{iwt})}\frac{C_{it}}{BMR_{it}}|H_{iwt} > 0\right]$$

$$- E\left[\frac{H_{iwt}}{(1 - H_{iwt})}MET_{iwt}|H_{iwt} > 0\right]$$

where

$$E\left[\frac{W_{it+1}}{BMR_{it}}|H_{iwt} = 0\right] = E\left[W_{it+1}|H_{iwt} = 0\right]E\left[\frac{1}{BMR_{it}}|H_{iwt} = 0\right] + Cov\left[W_{it+1}, \frac{1}{BMR_{it}}|H_{iwt} = 0\right] \quad (5.5)$$

$$E\left[\frac{C_{it}}{BMR_{it}}|H_{iwt} = 0\right] = E\left[C_{it}|H_{iwt} = 0\right]E\left[\frac{1}{BMR_{it}}|H_{iwt} = 0\right] + Cov\left[C_{it}, \frac{1}{BMR_{it}}|H_{iwt} = 0\right] \quad (5.6)$$

$$E\left[\frac{1}{(1 - H_{iwt})}\frac{W_{it}}{BMR_{it}}|H_{iwt} > 0\right] = E\left[\frac{1}{(1 - H_{iwt})}|H_{iwt} > 0\right]E\left[\frac{W_{it}}{BMR_{it}}|H_{iwt} > 0\right] \quad (5.7)$$

$$+ Cov\left[\frac{1}{(1 - H_{iwt})}, \frac{W_{it}}{BMR_{it}}|H_{iwt} > 0\right]$$

$$E\left[\frac{1}{(1 - H_{iwt})}\frac{W_{it+1}}{BMR_{it}}|H_{iwt} > 0\right] = E\left[\frac{1}{(1 - H_{iwt})}|H_{iwt} > 0\right]E\left[W_{it+1}|H_{iwt} > 0\right]E\left[\frac{1}{BMR_{it}}|H_{iwt} > 0\right] \quad (5.8)$$

$$+ E\left[\frac{1}{(1 - H_{iwt})}|H_{iwt} > 0\right]Cov\left[W_{it+1}, \frac{1}{BMR_{it}}|H_{iwt} > 0\right]$$

$$+ E\left[W_{it+1}|H_{iwt} > 0\right]Cov\left[\frac{1}{(1 - H_{iwt})}, \frac{1}{BMR_{it}}|H_{iwt} > 0\right]$$

$$+ E\left[\frac{1}{BMR_{it}}|H_{iwt} > 0\right]Cov\left[\frac{1}{(1 - H_{iwt})}, W_{it+1}|H_{iwt} > 0\right]$$

$$E\left[\frac{1}{(1-H_{iwt})}\frac{C_{it}}{BMR_{it}}|H_{iwt}>0\right]=E\left[\frac{1}{(1-H_{iwt})}|H_{iwt}>0\right]E\left[C_{it}|H_{iwt}>0\right]E\left[\frac{1}{BMR_{it}}|H_{iwt}>0\right] \tag{5.9}$$

$$+E\left[\frac{1}{(1-H_{iwt})}|H_{iwt}>0\right]Cov\left[C_{it},\frac{1}{BMR_{it}}|H_{iwt}>0\right]$$

$$+E\left[C_{it}|H_{iwt}>0\right]Cov\left[\frac{1}{(1-H_{iwt})},\frac{1}{BMR_{it}}|H_{iwt}>0\right]$$

$$+E\left[\frac{1}{BMR_{it}}|H_{iwt}>0\right]Cov\left[\frac{1}{(1-H_{iwt})},C_{it}|H_{iwt}>0\right]$$

$$E\left[\frac{H_{iwt}}{(1-H_{iwt})}MET_{iwt}|H_{iwt}>0\right]=E\left[\frac{H_{iwt}}{(1-H_{iwt})}|H_{iwt}>0\right]E\left[MET_{iwt}|H_{iwt}>0\right] \tag{5.10}$$

$$+Cov\left[\frac{H_{iwt}}{(1-H_{iwt})},MET_{iwt}|H_{iwt}>0\right]$$

Our calculated implied METs are as shown in Table 5.1. Starting in 1980, we find very similar implied levels of activity for men and women of about 1.8, i.e. activity that is factor 1.82 (w) to 1.85 (m) more energy-intensive than resting. Over time, activity levels decrease for both, men and women, in similar fashion to around 1.41 (w) to 1.44 (m), or by ca. 22%. When we additionally consider our data on activity at work, and compute the implied METs in non-work activities (such as leisure, housework and travelling), we find discrepancies between women and men: due to low labour force participation in the 1980s, the relatively higher strenuousness at work does not lead to a reduction in non-work METs in the 1980s for women. However, for men, we find that high and strenuous activity at work falls by almost 10% over time. Again, trends are somewhat similar for both men and women.[2] Our resulting implied METs given the trends in weight, calories, activity at work and sleep show strongly falling trends over time. They fall - for both men and women- by 22% over time.

A quick comparison of these numbers with the qualitative trends in Table 4.1 indicates potential components of such decline. Women show a large reduction in time spent on (strenuous) housework since 1983. This time is mostly crowded out by additional work hours at (on average comparatively) lower strenuousness

[2]As METs vary due to occupational changes in our sample, these drops are influenced by changes in occupation-specific LFP.

levels, and by travel (which has a low and decreasing MET over time due to the rise of travel by car or public transport). For men, we see an even larger increase in sedentary travel time, which is consistent with our trend in non-work METs.

6 Summary and conclusion

[to be written]

Table 5.1: *Implied METs overall (not including sleep), work METs and implied non-work activities given imputed calories*

Year	Females			Males		
	overall METs (excl. sleeping)	non-work METs	work METs	overall METs (excl. sleeping)	non-work METs	work METs
1980	1.82	1.80	2.36	1.85	1.52	2.58
1981			2.34			2.54
1982						
1983			2.31			2.52
1984	1.69	1.68	2.31	1.75	1.43	2.53
1985			2.31			2.51
1986	1.68	1.65	2.31	1.73	1.43	2.50
1987			2.29			2.50
1988			2.28			2.48
1989			2.28			2.50
1990			2.26			2.48
1991	1.61	1.55	2.26	1.62	1.25	2.47
1992	1.59	1.55	2.24	1.62	1.32	2.44
1993	1.58	1.54	2.24	1.61	1.31	2.43
1994	1.58	1.53	2.23	1.61	1.32	2.43
1995	1.53	1.48	2.23	1.55	1.22	2.41
1996	1.59	1.54	2.22	1.61	1.30	2.41
1997	1.55	1.49	2.22	1.57	1.24	2.41
1998	1.51	1.45	2.21	1.53	1.19	2.40
1999	1.52	1.47	2.20	1.54	1.20	2.40
2000	1.52	1.46	2.20	1.55	1.22	2.39
2001	1.49	1.43	2.18	1.51	1.19	2.37
2002	1.50	1.43	2.18	1.52	1.22	2.38
2003	1.47	1.41	2.17	1.50	1.19	2.38
2004	1.48	1.41	2.17	1.48	1.16	2.37
2005	1.47	1.40	2.16	1.50	1.20	2.36
2006	1.47	1.41	2.16	1.50	1.21	2.36
2007	1.43	1.35	2.16	1.45	1.12	2.35
2008	1.41	1.33	2.16	1.44	1.11	2.35

Source: Own calculations.

A Data on Weight

Information on the distribution of weight in the population comes from three surveys.

National Heights and Weights Survey, 1980 The National Heights and Weights Survey is a representative sample of adults in Great Britain aged 16-64 in 1980 with around 10,000 respondents. It was commissioned by the then Department of Health and Social Security.

Health and Lifestyle Survey (HALS) 1984/85, 1991/92 The first HALS was sponsored by the Health Promotion Research Trust and draw a sample of the adult population (aged 25 or more) in England, Wales and Scotland in 1984-1985. Its aim was to collect information on self-reported health, physiological measures, psychological and cognitive measures, dietary and exercise habits as well as social and working conditions. A second wave of the survey collecting information on surviving respondents of the first wave was carried out in 1991-1992. About 9,000 individual were interviewed in the first wave and 5,000 re-interviewed in the second wave. **Dietary and Nutritional Survey of British Adults 1986/87** A sample of the population of British adults aged 16 to 64 in 1986-1987, the Dietary and Nutritional Survey of British Adults collects information on physical measures, blood pressure, dietary and exercise habits as well as personal and household characteristics. Besides having a short interview, respondents are asked to fill a diary for seven weeks with their intakes of food and beverages consumed in and out of home. The sample consists of about 2,200 respondents. **Health Survey for England (HSE) for 1991 to 2010** Sponsored by the Information Centre for Health and Social Care the Health Survey for England started collecting information on health and factors that affect health among the English population in 1991. A cross-sectional annual survey, it draws a yearly sample of between 10,000 and 20,000 individuals representative of the English population living in private households. Respondents have a face-to-face interview and have clinical and physical measures taken. The survey contains a core questionnaire that is repeated every year and focuses on different health issues every year such as cardiovascular disease and associated risk factors, accidents, disability, physical activity and fitness, long-term health conditions, among others.

B Data on occupation

The Labour Force Survey (LFS) is the largest household survey in the UK and provides official figures for employment and unemployment. It samples between 40,000 and 60,000 households, annually from 1979-1991,

quarterly from 1992-2009.

Due to the long sample period, there are several changes in the classification of occupations in the LFS (CODOT for 1980-1990, SOC1990 for 1991-2000 and SOC2000 from 2001). Since no unique mapping between these is available for all codes, we map the detailed occupation codes across time in the following way: In the LFS 1991, data were dual-coded to CODOT and to the Standard Occupational Classification (SOC) code. The resulting mapping is for some codes not unique. We follow a similar strategy as Nickell and Saleheen (2008) and calculate gender-specific transition probabilities for all code pairs and then randomly assign a match based on the generated gender-specific probability distributions to assign a SOC 1990 code in all years (100-999). We follow a similar method for the mapping between the 1990s and the 2000s. Since no LFS dataset is dual-coded for SOC 1990 and SOC2000, we use the BHPS (period 2001-2005) to get the universe of code mappings. Based on these, we again determine the transition probabilities and draw SOC1990 for the SOC2000 codes from the resulting gender-specific probability distributions. This gives us a consistent occupation classification coding for 1980-2009 based on the three-digit SOC1990 codes and also allows us to generate consistent broad occupation classifications based on the 1st digit. There are 374 occupations identified which fall into nine broad groups (Managers and Administrators, Professionals, Associate Professionals and Technical Occupations, Clerical and Secretarial Occupations, Craft and Related Occupations, Personal and Protective Services, Sales, Plant and Machine Operatives and Other).

C Data on calories

The main types of data on calories include:

- *Household budget surveys*: Budget surveys collect information on expenditure and quantities purchased. Households are usually required to record daily purchases in a diary over a one or two week period. Calorie and nutrient density of food products is matched in from standard dietary reference tables. These data are currently collected as part of the Living Costs and Food Survey (LCFS) in the UK and the Consumer Expenditure Survey (CEX) in the US.

- *Individual intake surveys*: Intakes studies ask individuals to record details of daily food consumption, typically over a 24 hour period. Nutrient conversion factors are assigned to food items to obtain measures of nutrient intakes.

- *Market research scanner data*: Kantar
- *Aggregate food balance sheets*: Collected by the United Nations Food and Agriculture Organization (FAO) but also the US Department of Agriculture, they measure food availability for human consumption at the country level. The FAO reports quantities as well as calories, protein and fats available for human consumption at the country level in a given time period.

C.1 Intake data

It has been shown for many countries that there is large underreporting in intake surveys, especially for women (e.g. Bingham et al. 1995, Briefel et al. 1997, Rennie et al. 2007). Table C.1 compares daily average calorie intakes from British Nutrition Surveys and our calorie purchase data with Estimated Average Requirements for energy from the Department of Health. It shows large underreporting in intake surveys, especially for women, relative to average energy requirements. Average calorie *purchases* - on the contrary- are above estimated average energy requirements.

The calorie purchase data makes allowance for cooking losses and waste from inedible foods, but not for other food waste. According to Defra (2007) about 11% of calories are wasted.[3] Incorporating these waste estimates, average calorie purchases are very close to EARs.

Calorie purchases are likely less prone to systematic underreporting since the diary does not reveal their eating, just their spending behaviour. An additional source of anonymity is that a respondent records food spending for the household and not his or her individual spending.

C.2 Food Availability Data

The United Nations Food and Agriculture Organization's (FAO) food balance sheet. It covers production, trade, feed and seed, waste and food availability for roughly 180 countries since 1961. The FAO reports quantities as well as calories, protein and fats available for human consumption at the country level in a given time period. FAO's food balance sheets data have been used to study the contribution of energy intake to the obesity problem in developed and developing countries (See for instance: Cutler et al. (2003) in the US and Bleich et al. (2008)). Due to lack of long term series data from household surveys, researchers have used

[3]The data do not account for visitors to the household. Free food is included until 2000, and estimated afterwards (4.3% of calories from all foods per person per day)

Table C.1: *Calorie purchases versus intakes*

Year	Women			Men		
	Nutrition survey	NFS/EFS	EAR	Nutrition survey	NFS/EFS	EAR
1986/87	1,680	2,145[a]	1,940	2,450	2,703[a]	2,550
2000/01	1,635	2,267	1,940	2,308	2,842	2,550
2008/09	1,645	2,186	1,940	2,255	2,702	2,550
Incorporating 11% of calorie wastage						
1986/87	1,680	1,909[a]	1,940	2,450	2,405[a]	2,550
2000/01	1,635	2,018	1,940	2,308	2,529	2,550
2008/09	1,645	1,945	1,940	2,255	2,587	2,550

Source: DoH (1991). [a]: only food at home.

FAO's food balance sheet data as proxies for calorie intakes. But there are a number of potential problems with FAO's data that are worth mention.

Researchers have used food available for human consumption as a proxy for food consumption. But FAO estimate food availability as a residual and thus its error, either level or sign, is not quantifiable. According to FAO's food balance sheet handbook: "...At a minimum, this means the quantity of food available for human consumption would have to be estimated independently based on other existing statistical sources of information. One such form would be household survey which collects quantities of food items consumed or acquired". A number of studies (see below) have tried to compare data from household surveys and food balance sheets but no general explanation for the difference between the two data sources has been reached. Differences in levels are expected but household surveys and FAO's food balance sheets should result in similar trends over time. That is not the case for several countries, in particular India (Deaton and Dreze (2010)), Japan (Dowler and Seo (1985)), the US (Crane et al. (1992)) and the UK (see Figure **??**).

From the household budget survey data we estimate a decline of XX% in calories per person between 1980 and 2009, FAO reports an increase of 10%. Deaton and Dreze (2010) find similar differences for India. They comment that *"the reliability of FAO data is at best uncertain"*. Moreover, when referring to FAO's "food supply" figures, Deaton and Dreze (2010) argue that are *"derived from rather speculative "balance sheets" of national production and utilisation, instead of household surveys, which are often more reliable"*.

Crane et al. (1992) study the trends in macronutrients per capita estimated from food supply data and the trends from intake household surveys for the US in the period 1965 to 1988. They use data from several intake household surveys and food supply data from the US Department of Agriculture and find that the trend in macronutrients from the two sources expressed in grams diverge. While they report an increase in fats, carbohydrates and proteins from food supply data, the trend is declining when using household survey data. They also find an increase in calories per capita according to food supply and a decline when measured by intakes. Dowler and Seo (1985) show similar trends for Japan between 1953 and 1975.

Several studies have compared results from the FAO's food balance sheets and household surveys for given years. Pomerleau et al. (2003) compare FAO data on fruit and vegetables with reported intakes. In general they find that FAO overestimates intake in 14 out of 15 countries included in the study with the overestimation ranging from 5 to 270%. It is expected that calories data from FAO overestimate calorie intakes as FAO reports availability and does not take into account waste or losses. Other studies that arrive to similar findings are Rodrguez-Artalejo et al. (1996), Serra-Majem et al. (2003) and Rodrigues et al. (2007).

In a review commissioned by the FAO to understand the limitations and propose improvements to the food balance sheets data, Jacobs and Sumner (2002) state that food stock, feed and other uses are particularly prone to error. Food for human consumption is computed as a residual and thus if any of its components is measured with error, food supply is also going to be measured with errors. Take the example of cereals, they represent between 21% and 26% of total calories during 1980 to 2009. Cereals for feed represent 48% of total food supply implying that if we are interested in food supply, errors in the measurement of feed should not be taken lightly.

Food availability is measured as the residual of food supply (of each commodity) minus measurable uses, such as farm inputs (feed and seed), exports, ending stocks, and industrial uses.

Data on supplies of processed food is not very good; they provide loss-adjusted data: this accounts for loss at three levels (farm to retailer, retailer, and consumer). Until recently this was very poorly documented, and there seem to be a lot of assumptions involved in the process

Data is "Average daily per capita calories from the U.S. food availability, adjusted for spoilage and other waste" taken from: `http://www.ers.usda.gov/data-products/food-availability-%28per-capita%29-data-system/food-availability-documentation.aspx#construction`

The data is constructed as:

- the US food availability data are collected by the USDA's Economic Research Service (ERS) and reflect the amount of food available for human consumption in the US
- ERS calculates the residual of a commodity's total annual available supply after subtracting measurable uses, such as farm inputs (feed and seed), exports, ending stocks, and industrial uses
- food availability data measure the use of basic commodities, such as wheat, beef, and shell eggs for food products at the farm level or an early stage of processing
- total annual available supply of each commodity consists of the sum of production, imports, and beginning stocks. These three components are either directly measured or estimated by government agencies using sampling and statistical methods.
- the food supply is then given by:

  available commodity supply (product + imports + beginning stocks) - measurable nonfood use (farm inputs + exports + ending stocks, etc.) = total annual food supply of a commodity

- measurable nonfood uses are farm inputs (feed and seed), exports, ending stocks, and industrial uses.
- in a few cases, supplies for human food use are measured directly and one of the other use components becomes the residual. This is the case for wheat, in which flour production is measurable and available from manufacturers' reports on flour milling and therefore, livestock feed use becomes the residual.
- per capita food availability is calculated by dividing the annual total food supply during a specific time period by the U.S. total resident population plus Armed Forces overseas in a given year. Yearly population estimates are from the U.S. Census Bureau. For commodities not shipped overseas in substantial amounts, such as fluid milk and cream, ERS uses the resident population as the base. No adjustments are made for changes in the demographic makeup of the population.
- there is separately available loss-adjusted food availability data (see below)

The USDA highlights some limitations and sources of error:

- "little data exist on supplies of more highly processed foods. In short, relatively good data exist for many food ingredients (for example, flour, sugar, or eggs) but not for common food products (for example, bread, cookies, or beef hot dogs)."

- the data are aggregates for the United States and no data are available for use as a proxy for consumption within States or regions or by socioeconomic or demographic categories (such as those of the consumer who ultimately ate the food)
- food use is generally estimated as the residual of the balance sheets, food use data are subject to the various types of errors present in each of the balance sheet components. Primary sources of error are incomplete reporting, inaccurate conversion factors, and inappropriate estimation techniques.
- data on stocks are not available for some commodities

Loss-adjusted food availability data:

- Food loss represents the edible amount of food, postharvest, that is available for human consumption but is not consumed for any reason. It includes cooking loss and natural shrinkage (for example, moisture loss); loss from mold, pests, or inadequate climate control; and food waste.
- Food waste is a component of food loss and occurs when an edible item goes unconsumed, as in food discarded by retailers due to color or appearance and plate waste by consumers.
- Historically, the methods use to calculate these were not very transparent, so the ERS has worked to update and document the loss assumptions, but they comment that "ERS is in the process of determining the next steps for the data series to fill in remaining data gaps. Hence, the data series is considered to be preliminary."
- There are three general types of losses:
    1. losses at the primary level (e.g. farm to retail weight):
        - ERS and the University of Minnesota's Food Industry Center (TFIC) compiled revised agricultural conversion factors from farm to retail.
        - Loss estimates are sometimes called conversion factors, particularly when describing how a farm commodity is transformed into a consumer-ready product (for example, fresh chicken to boneless fresh chicken). Through information from a series of industry interviews, TFIC updated the conversion factors for the main categories of meats and poultry, as well as for several fruits and vegetables.

2. losses at the retail level (e.g. supermarkets, but not other retail outlets, but not restaurants or other food service outlets)
   - in September 2007, ERS obtained updated food loss estimates at the retail/institutional level to the consumer level (for example, from supermarkets) for fresh fruits, vegetables, meat, poultry, and seafood through a competitive grant with the Perishables Group, Inc. (PG).
   - PG compared supplier shipment data with point-of-sale data from stores in large, national supermarket retail chains to identify loss percentages.
   - new assumptions were close to the old ones
3. losses at the consumer level. 2 components: loss of non-edible share of food (collected from National Nutrient Database for Standard Reference, compiled by USDA's Agricultural Research Service (U.S. Department of Agriculture, 2007)) and edible share of food i.e. plate waste
   - RTI International used a numerical estimation method to calculate consumer-level food loss estimates using Nielsen Homescan data (food purchase data) and the dietary intake component of the National Health and Nutrition Examination Survey (NHANES) (food consumption data).
   - ERS then analyzed how the loss-adjusted food availability per capita data would change if the proposed RTI estimates of consumer-level food loss were incorporated into the data series.

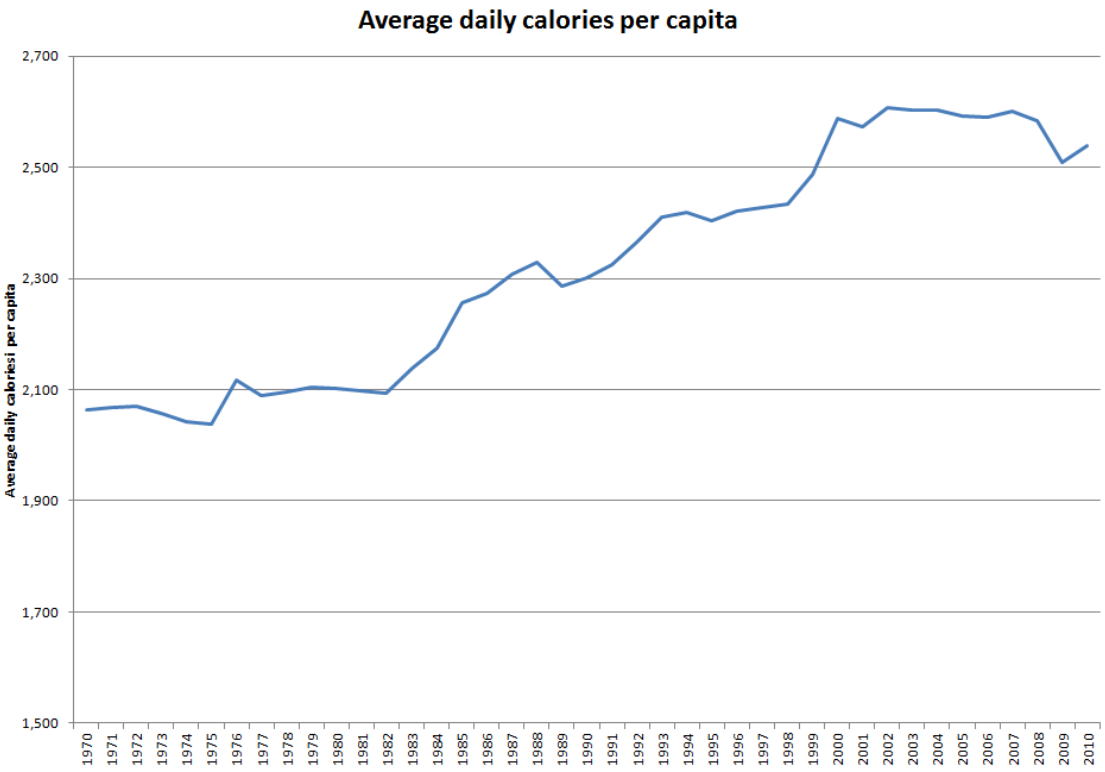
Figure C.1 shows the average calories per capita per day in the FAO data since 1970. We can also look at the trends in calories from the different food groups, see Figure C.2. The groups that increased a lot are grains and added fats and oils.

Some papers that use these data [should include Cutler papers here]

"Increased food energy supply is more than sufficient to explain the US epidemic of obesity" - Swinburn, Sacks and Ravussin (2009)

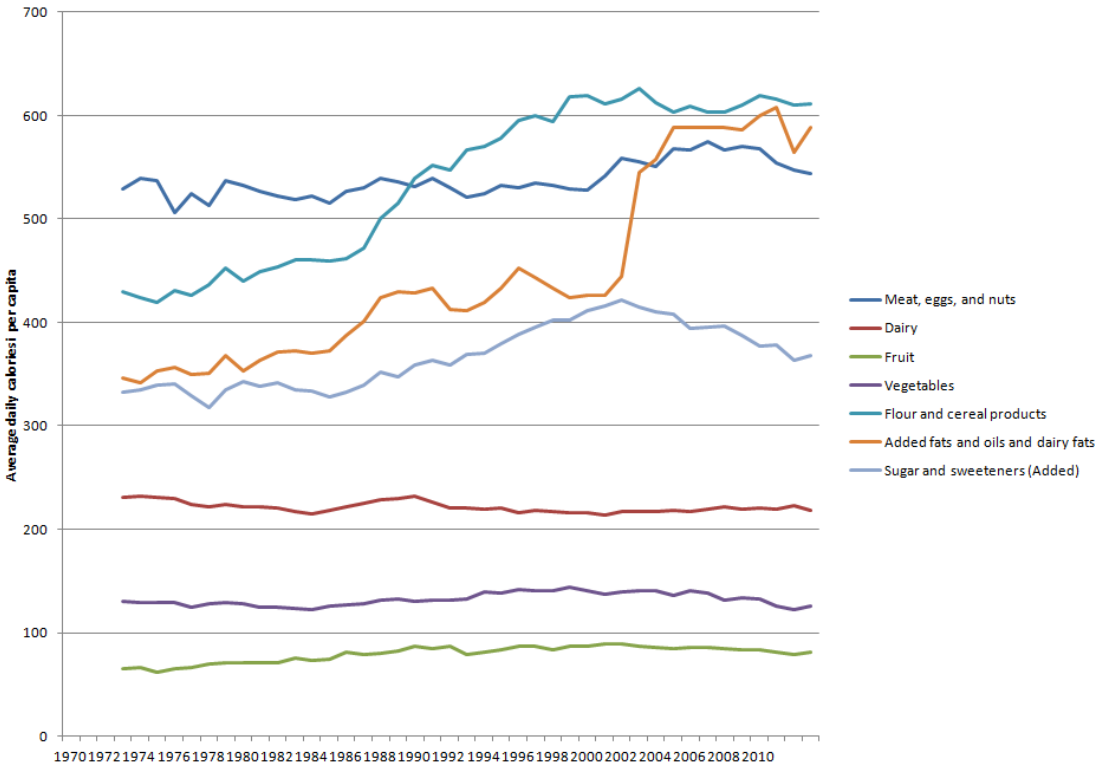
- predict changes in weight from the changes in estimated energy intake in US children and adults from 1970s to 2000s using equations that relate energy intake to weight gain
- food supply data for 1970s-2000s: adjusted for food loss; represent the food available for consumption

Figure C.1: *Average daily calories per capita since 1970*



Source: FAO.

Figure C.2: *Average daily calories per capita since 1970, by food group*



Source: FAO.

- argue that "the USDA states that the food supply data provide a useful indicator of trends in consumption over time because many of the errors in the data are likely to be systematic, thus cancelling out when changes over time are examined "
- use the NHANES to measure the average weights in the population
- the 1970 and 2000 US census data were used to derive the child:adult population proportions, and the doubly labeled water studies were used to derive data for energy intake proportions between children and adults (by using the appropriate mean age and weight from the census and NHANES data)
- **Result**: increased energy intake appears to be more than sufficient to explain weight gain in the US population

"Increased energy intake entirely accounts for increase in body weight in women but not in men in the UK between 1986 and 2000" - Scarborough et al (2011)

- use the NDNS to measure body weight height and age in 1986-7 and 2000-1, just for adults
- measures of total energy intake at the two time points were taken from food availability data, collected from food balance sheets prepared by the FAO; adjusted for loss using methods developed by the USDA (although this does not adjust for wastage at the retail level - outdated, see below). Assume that food wastage within each broad food group has remained constant
- estimates of total energy intake for men and women separately were derived by applying the ratio of energy intake in men compared with women derived from the 19867 Dietary and Nutrition Survey to the adjusted food availability data (the ratio in 19867 (143) was virtually identical to the similar ratio recorded in 20001 (140)).
- **Results:** For men, only part of the increase in body weight can be ascribed to the increase in energy intake, suggesting that reductions in physical activity levels have also played a role. For women, increases in energy intake explain all of the increase in body weight, suggesting that this was the dominant causal factor for increases in body weight over this time period.

- however, do report intake from dietary surveys and from purchase data on food in and show that between 1986 and 2000 mean total energy 'intake' fell, but they basically argue that they believe biases in food availability data are more stable over time
- there is quite a good discussion about why their results might not be accurate

C.3 UK Budget Surveys

National Food Survey (NFS) for 1980 to 2000

The National Food Survey (NFS) was originally set up in 1940 by the then Ministry of Food to monitor the adequacy of the diet of urban households during wartime, but it was extended in 1950 to become representative of British households. It is a repeated cross-section that samples about 8,000 households per year. The information is collected continuously during the year. The person who does most of the food shopping is in charge of keeping records of all the food purchased and brought home during seven days. For each item, the diary keeper has to record in the diary: a short description, quantity purchased and the cost. Free food is recorded only at the time of use. Depending on the year there are between 180 and 240 food items. Together with food purchases, demographic characteristics of the household members are obtained through a face-to-face interview.

Expenditure and Food Survey (EFS) and Living Costs and Food Survey (LCFS) for 2001 to 2010

The last wave of the NFS was conducted in 2000. From April 2001, the NFS has been merged with the Family Expenditure Survey to form the new Expenditure and Food Survey (EFS). Later in 2008 and due to small changes in the questionnaire, the EFS was re-named as the Living Costs and Food Survey (LCFS). The EFS/LCFS collects household expenditure not only on food and beverages but also on other durable and non-durable goods. All household members aged 16 or more keep a personal diary of daily expenditure for 14 days. Simplified diaries are kept by members of the household aged between 7 and 15. Contrary to the NFS, the EFS/LCFS collects not only information on all the food and beverages brought home but also on food and drinks that do not enter the house such as restaurant meals and school meals. For household purchases the quantities are collected in the diaries and also through receipts. This is not possible for eating out, take-aways and some instances of free food. For these items the Department for

Environment, Food and Rural Affairs (DEFRA) use estimated portion sizes to calculate weight/volume and in turn nutrient intakes. Where possible, whole meals eaten out are split into food components. There are in total around 500 food items reported in the EFS/LCFS of which 240 are of food brought home. Another difference with the NFS is with respect to free food. While free food is recorded in the diary in the NFS, it is obtained through an interview in the EFS/LCFS and most categories are estimated by assigning standard portion sizes. Expenditure on regular items, such as utilities, mortgages payments or rents, as well as infrequent expenditures are obtained through a face-to-face interview. Demographic characteristics of household members are also obtained through a face-to-face interview.

Consistency: NFS, EFS and LCFS

We made two changes to the original expenditure data in order to make the categories in the different surveys consistent over time.

Grouping: Although the surveys are similar in its structure, there are few changes in the available items over time. In general, there is a higher level of aggregation in the expenditure items available in the NFS than in the EFS/LCFS. Not considering alcohol, confectionery and soft drinks and eating out categories, there are roughly between 180 and 200 items between 1974 and 1984, 215 since 1985 and 240 since 2001. In order to make the items consistent over time we have to grouped some of them. A detailed list of items is available from the authors upon request.

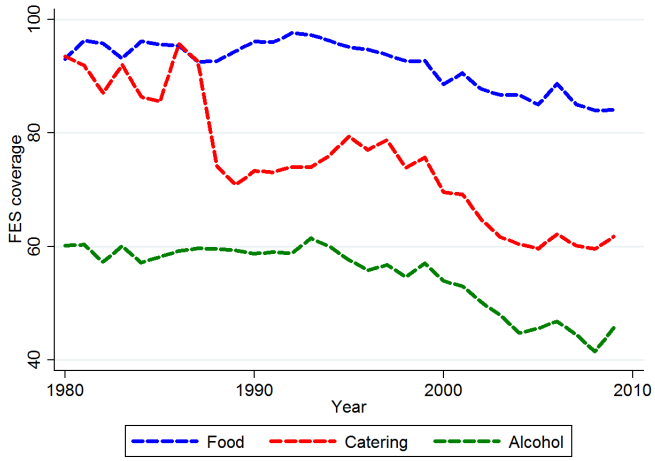
Adjustment factors: According to DEFRA, for some types of food, expenditure estimates from the NFS (1980-2000) are substantially lower than those from the EFS (from 2001). In order to make the data comparable DEFRA suggest to apply an adjustment factor. Take the example of Yoghurt (code 1301). The adjustment factor proposed by DEFRA is 1.14732982499301. Then, if average purchased quantities of yoghurt per person per day in the NFS is 28.9 ml, DEFRA suggests to adjust that value as 28.9 times Adjustment factor=33.1 ml. We follow this procedure for all the items. A detailed explanation of the estimation of the NFS adjustment factors can be found in DEFRA (2011).

C.4 Expenditure: Household surveys versus National Accounts

The coverage of total expenditure in household budget surveys relative to consumption in the National Accounts has been studied, among others, by Deaton (2005), Attanasio et al. (2006) and Brewer and O'Dea (2012). According to Attanasio et al. (2006) the coverage problem is more important in the US than in the UK. The EFS/FES in the UK has had historically followed quite closely consumption in the National Accounts but coverage problems have recently been detected by Brewer and O'Dea (2012). National Accounts data is not exempted from problems but it is a good benchmark to which compare household level data.

We use data from Brewer and O'Dea (2012) to assess the coverage of food expenditure in the FES and subsequent surveys during the period covered in our study. Figure C.3 shows the percentage of expenditure captured in the FES relative to the National Accounts for our three expenditure categories: food at home, alcohol and food out (catering). The level of coverage of food expenditure is relatively high but has declined over time, particularly from the mid-1990s. While the FES captured 93% of food expenditure in the National Accounts in 1980, coverage has declined to 84% in 2009. The cases of alcohol and food out are more worrisome. Alcohol coverage has declined from 60% in 1980 to 46% in 2009 and food out coverage has declined from 94% to 62% during that period.

Figure C.3: *Expenditure FES coverage compared to National Accounts (in %)*



Source: Brewer and O'Dea (2012)

Given the declining coverage and assuming that the National Accounts are correct, what would be the trend in calories assuming that expenditure coverage is the same as for calories? Figures C.4, C.5, C.6 and

C.7 show coverage adjusted calories for the total and our three categories: food at home, alcohol and food out. The declining pattern of total calories over time has not change even after the adjustment by expenditure coverage. This is because 80% of calories are from food at home; category for which the coverage problem is not that important. The decline in average household calories purchases after accounting for expenditure coverage is -28%, compared with -39% when using unadjusted calories.

Figure C.4: *Total calories per household: our estimates and corrected for FES expenditure coverage*

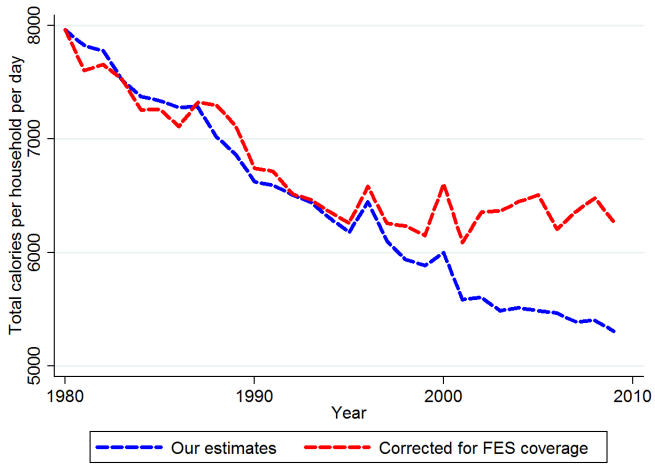


Figure C.5: *Calories per household from food in: our estimates and corrected for FES expenditure coverage*

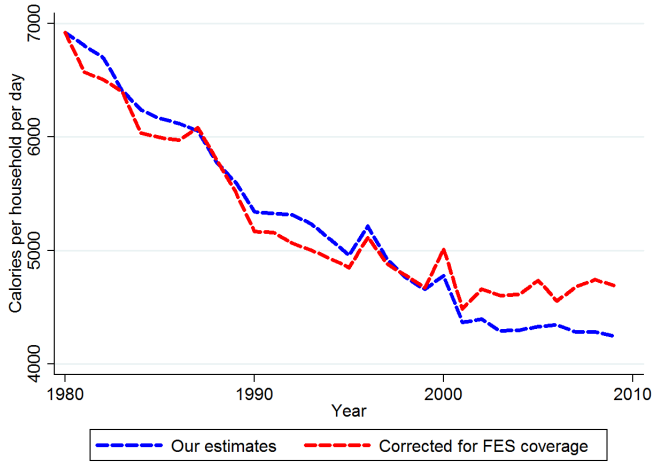


Figure C.6: *Calories per household from eating out, takeaways, soft-drinks and confectionary: our estimates and corrected for FES expenditure coverage*

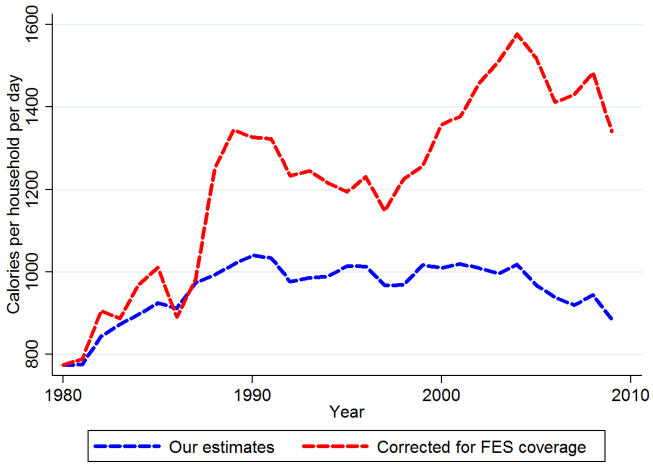
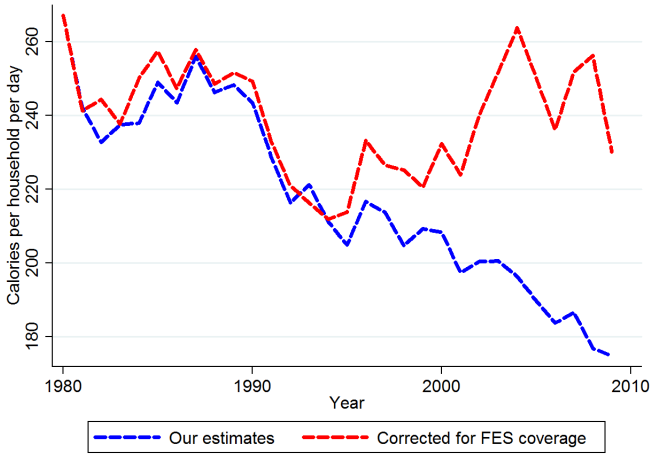


Figure C.7: *Calories per household from alcohol: our estimates and corrected for FES expenditure coverage*



C.5 From household to individual calories

Data on food purchases are reported at the household level. The composition of households changes substantially over time. For example, the average household size is 2.99 in 1974 and falls to 2.36 by 2009. In 1974 the average number of children is 0.93, while by 2009 it falls to 0.47. Comparing trends over time at the household-level could thus be misleading.

For our main estimates we convert household-level calories into individual-level calories assuming that the household shares all calories out amongst household members in accordance to their relative estimated average calorie requirements. The Estimated Average Requirements (EAR) for energy by age and gender from the Department of Health (DoH, 1991) are:

Table C.2: *Estimated Average Requirements (EAR)*

Age	Female	Male
4-6	1545	1715
7-10	1740	1970
11-14	1845	2220
15-18	2110	2755
19-50	1940	2550
51-59	1900	2380
60-74	1900	2330
75 plus	1810	2100

Source: DoH (1991).

This is similar in spirit to Chesher (1998), who non-parametrically estimates calorie profiles over age using the NFS data. We also replicate Chesher's method using a parametric version. As described by Chesher in Section 5 in **?**, average household calorie purchases can be expressed as the sum of the average calorie purchases of its members. If average calorie purchases for each household member aged A could be expressed as $\alpha_M + \beta_M A$ if male and $\alpha_F + \beta_F A$ if female, we can estimate average calorie purchases by age and gender by estimating the following equation by Ordinary Least Squares:

$$x_{it} = \alpha_M NM_{it} + \alpha_F NF_{it} + \beta_M TM_{it} + \beta_F TF_{it} + \epsilon_{it}, \tag{C.1}$$

where i is an index for households, t is time, x represents household calorie purchases, NM is the number of males in the household, NF is the number of females in the household, TM is the sum of the ages of males in the household and TF is the sum of the ages of females in the household. We estimate equation C.1 for each year between 1980 and 2009 in order to obtain estimates of the average calorie purchases for males and females over time. Figure C.8 shows the results using a parametric Chesher method and our optimal sharing rule for males and females.

D Imputing data on snacks, drinks, food out and alcohol

We want a measure of the amount of calories (x) that households purchase over time. This is measured as the quantity purchased times a calorie conversion factor:

$$x_{it} = q_{it}\eta_t \tag{D.1}$$

where

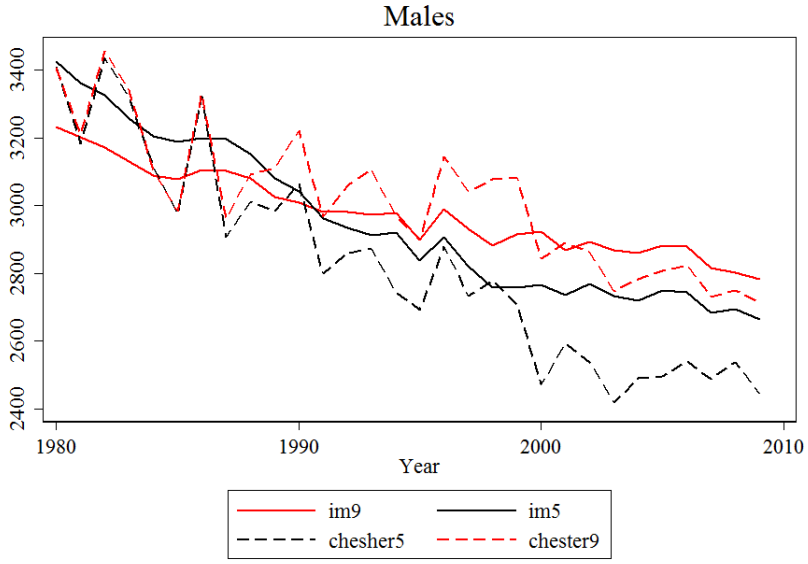
- i index household,
- $t = 1980, ..., 2009$, index time,
- q_{it}: grams purchased by household i at time t,
- η_t: calories per gram.

The information that we have is shown in Table D.1.

We use information on the distribution of expenditure on food out and alcohol from the EFS across demographics and time to impute expenditure for the sample of households observed in the NFS prior to 2001. The samples for the EFS and NFS are drawn from the same population and in the same manner (after 2000 the surveys were merged into the EFS/LCFS to reduce duplication of data collection).

We use a multi-step imputation method. We impute real expenditures $\tilde{q}_{it}$ for the NFS population using information on expenditure and demographics from the EFS. We obtain a backward forecast of calories per pound (in December 2005 prices), $\tilde{\eta}_{it}$, using information from the NFS for 2001-2010. We combine these to obtain an imputed measure of nutrient purchases x_{it} for the period 1980-2000.

Figure C.8: *Alternative methods of getting individual from household*



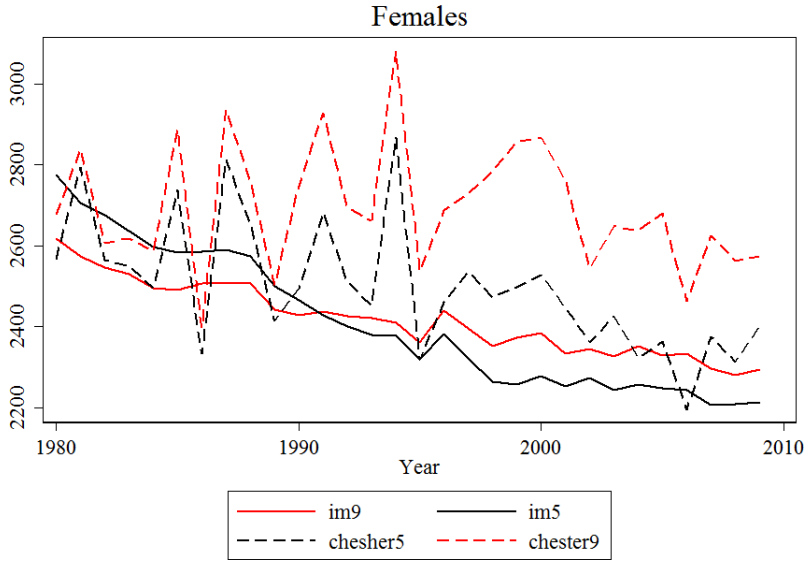


Table D.1: *Food spending and nutrition information in the National Food Survey*

		Food home		Food out		alcohol	
		1980-2001	2001-2010	1980-2001	2001-2010	1980-2001	2001-2010
Calories per gram	η_t	NFS	LCFS		LCFS		LCFS
Food quantity (in grams)	q_{it}	NFS	LCFS		LCFS		LCFS
Food expenditure (in GBP)	y_{it}	NFS, EFS	LCFS	EFS	LCFS	EFS	LCFS
Household demographics	D_{it}	NFS, EFS	LCFS	NFS, EFS	LCFS	NFS, EFS	LCFS
Food Prices	p_t	ONS	ONS	ONS	ONS	ONS	ONS

$$x_{it} = q_{it}\frac{y_{it}}{q_{it}p_t}\eta_{it}\frac{p_t}{\frac{y_{it}}{q_{it}}} = \frac{y_{it}}{p_t}\frac{\eta_{it}p_t}{up_{it}} = \tilde{q}_{it}\tilde{\eta}_{it} \tag{D.2}$$

where $up_{it} = \frac{y_{it}}{q_{it}}$ is the unit price (or value) of a food group, e.g. the amount paid per unit (in grams) and $\tilde{\eta}_{it}$ is the nutrient conversion factors η_{it} expressed per-real-December 2005 GBP.

D.1 Step 1: imputing real expenditure in the NFS ($\tilde{q_{it}}$)

In the first step we follow a multiple imputation procedure to obtain $\tilde{q_{it}}$. Multiple imputation is a statistical method to deal with missing values. We specify a model to predict missing values, once we have this we predict values multiple times, then we combine these multiple imputations to reflect the uncertainty in the imputation process.

We use multiple imputation by chained equations (MICE) (Rubin (1987), Rubin (1996), Schafer (1997), Schafer (1999) and Royston and White (2011), Horton and Kleinman (2007) amongst others). We are imputing two missing values (food out and alcohol) and we observe a third (food in); we use information on their joint distribution. The MICE algorithm in our case starts by regressing a transformation of real expenditure on alcohol on demographics D_{it}, real expenditure on food at home and real expenditure on food out. The missing values for $\tilde{q_{it}}$ are replaced by simulated values from the posterior predictive distribution of $\tilde{q_{it}}$ to obtain $\hat{\tilde{q_{it}}}$. In order to draw values from the observed values, PMM uses the closest neighbour based on the predictive value of a linear model. After obtaining imputed values for real expenditure on alcohol, we regress real expenditure on food out on demographics, real expenditure on food at home and real expenditure

on alcohol.

The key step in multiple imputation is specifying the model to predict the missing values. We tackle two particular problems.

First, a large number of households report zero expenditure on food out and alcohol, so we use the predictive mean matching (PMM) to impute the missing values (Vink et al. (2014).

Second, we observe nominal expenditures y_{it} and prices p_t. We can obtain $\tilde{q_{it}}$ either by directly imputing $\widehat{\left(\frac{y_{it}}{p_t}\right)}$ or by first imputing current expenditure and then dividing by prices, $\frac{\hat{y}_{it}}{p_t}$. We do both and choose the method for which the imputed values are closest to the observed ones where we observe both (2001-2010); the one that works best is to directly impute real expenditures $\tilde{q}_{it}$, so that is what we describe here.

The quantity index has a skewed distribution; we follow Royston and White (2011) and use a shifted log transformation. A real number γ is estimated and subtracted from the original variable to obtain a zero skewness transformed variable. The transformation is:

$$k_{it} = ln(\pm\left(\frac{y_{it}}{p_t}\right) - \gamma) \tag{D.3}$$

where γ and the sign are chosen so that the skewness of k_{it} is zero.

We then estimate the following linear regression model (for each of food out and alcohol):

$$k_{it} = \alpha + \beta D_{it} + \lambda_1 D_{it}\omega_m + \lambda_2 D_{it}\mu_t + \omega_m + \mu_t + \delta k'_{it} + \epsilon_{it} \tag{D.4}$$

where:

- D_{it} are household demographics (and interactions with time)
- ω_m are month dummies
- μ_t are year dummies
- k'_{it} is a vector of terms in real expenditure; for food out it includes real expenditure on food at home and on alcohol, and for alcohol it includes real expenditure on food at home and food out; the real expenditure variables are transformed as described by equation (D.3).

To obtain the simulated values from the posterior predictive distribution we use PMM, a procedure in

which imputed values are draws from the observed values of the imputed variable. In our case, imputed values for k_{it} on each expenditure group in the NFS for $t = 1980, ..., 2000$ are sampled from the observed values in the EFS. PMM draws from the closest neighbour based on the predictive value of a linear model. We repeat this procedure for 3 imputations.

The estimate of a household's real food expenditure is given by the mean of the 3 imputations,

$$\bar{k}_{it} = \frac{1}{3}\sum_{j=1}^{3} \hat{k}_{ijt}. \tag{D.5}$$

The variance of $\bar{k}_{it}$ is

$$\sigma_{total} = \sigma_{WI} + \frac{4}{3}\sigma_{BE}, \tag{D.6}$$

where the within-imputation variance $\bar{\sigma}_{WI} = \frac{1}{3}\sum_{j=1}^{3}\sigma_j$, and the between-imputation variance $\bar{\sigma}_{BE} = \frac{1}{3}\sum_{j=1}^{3}\hat{k}_{ijt} - \bar{k}_{it}$.

Data

We use EFS data on expenditure and ONS data on prices to construct a quantity index, $\left(\frac{y_{it}}{p_t}\right)$, for 1980-2009. This is measured in real December 2005 GBP.

Table D.2: *Food categories*

Food category	Definition of category	Availability
food consumed at home	milk products, cheese, meat, fish, eggs, oil and fats, sugar and preserves (e.g. jams etc.), fruit and vegetables, cereals, cakes, buns and pastries, other (incl. ready meals)	all periods
food consumed outside the home	any food eaten at restaurants, cafes, bars, bistros, fast food outlets, takeaways, snacks and confectionery (in and out), soft drinks (in and out)	2001-2009
alcohol	any beer, cider, wine, spirits, alcopops (consumed in and outside the home)	2001-2009

The household demographics are:

Imputed real expenditure

Table D.3: *Household demographics*

hht_i	household type dummies (single male, single female, lone parent, couple, family and other)
$work_i$	dummy variable that takes the value 1 if the head of household is working and 0 otherwise
age_i	dummy variables that take the value 1 if the head of household age group is equal to (less than 35, 35-50, 51-64, 65+)
A_i	number of adults in the household
K_i	number of kids in the household
E_i	number of income earners in the household
interactions:	
	total expenditures and household type dummies
	total expenditures and work status
	total expenditures and age of head (group dummies)
	age group of head and household type
	household type and number of kids in different age groups

In order to assess the validity of our imputation procedure we compare the distribution of the imputed values with the observed values at different points in time. The imputation procedure does very well in capturing the distribution of the observed values; the density functions of the imputed values are similar to those of observed values for all of our categories.

We are also interested in assessing how our imputation procedure does in estimating the mean over time. For both alcohol and eating out our imputation procedure does a very good job in capturing the mean trend over time. Figure D.3 shows the mean of observed real expenditure on alcohol across households over time in the EFS and the imputed values in the NFS. It also shows the 95% confidence interval computed using the equation (D.6). Figure D.4 shows the same variables for eating out. These figures show that our imputed values follow the observed ones very closely. Moreover, the observed values generally lie within the 95% confidence interval.

D.2 Step 2: backcasting calorie values $\tilde{\eta}_{it}$

In step 1 we obtained an imputed quantity. We also need to know the calorie density. We backcast $\tilde{\eta}_{it}$, the amount of calories per pound spent on food consumed outside the home or on alcohol. We use information

Figure D.1: *Real expenditure on alcohol*

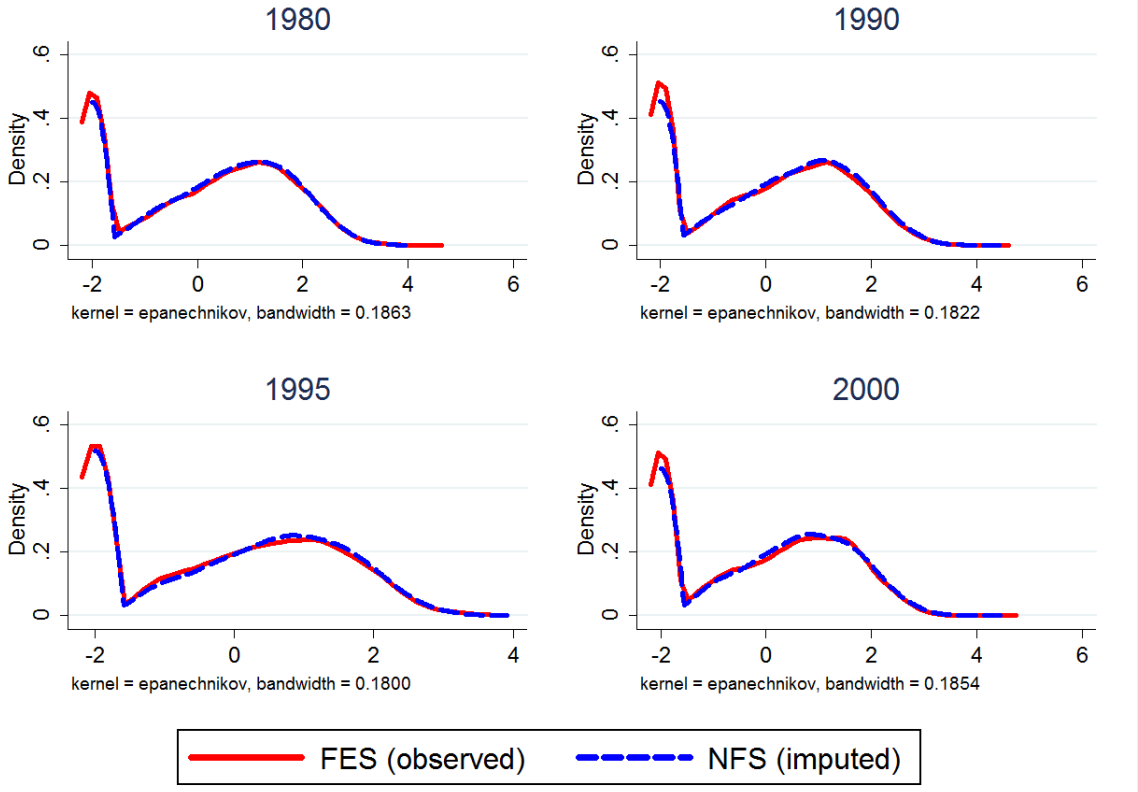


Figure D.2: *Real expenditure on food out*

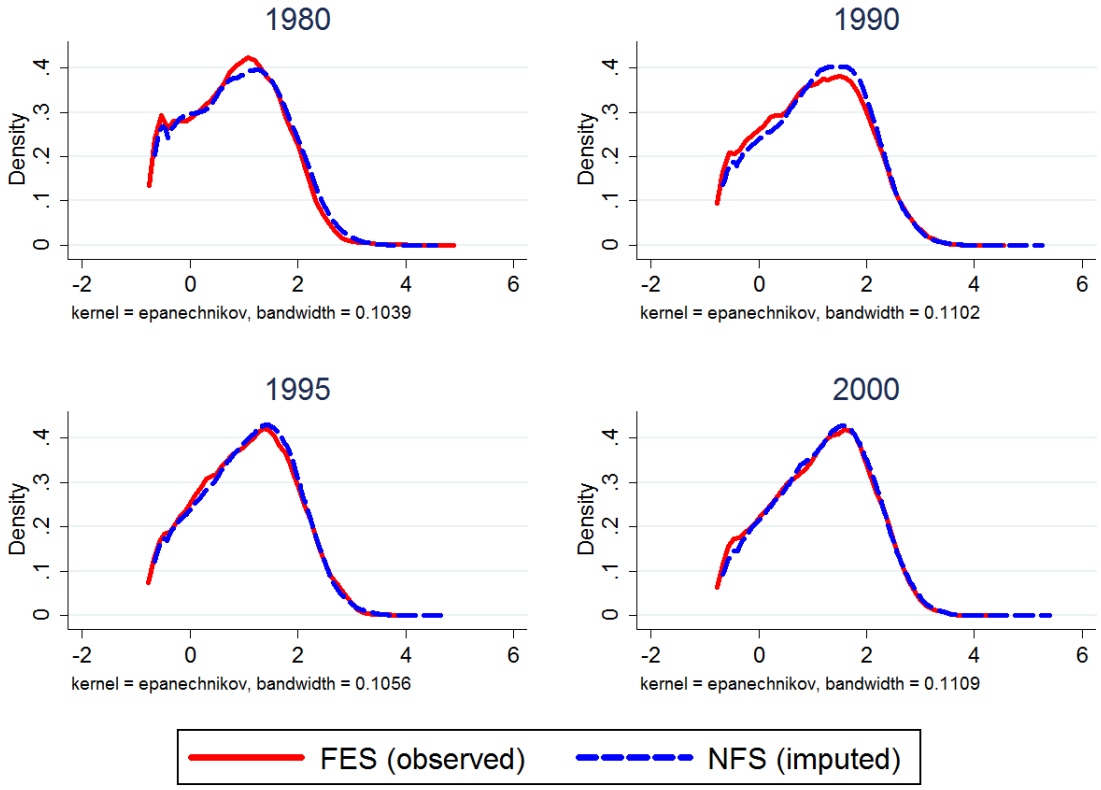


Figure D.3: *Real expenditure on alcohol*

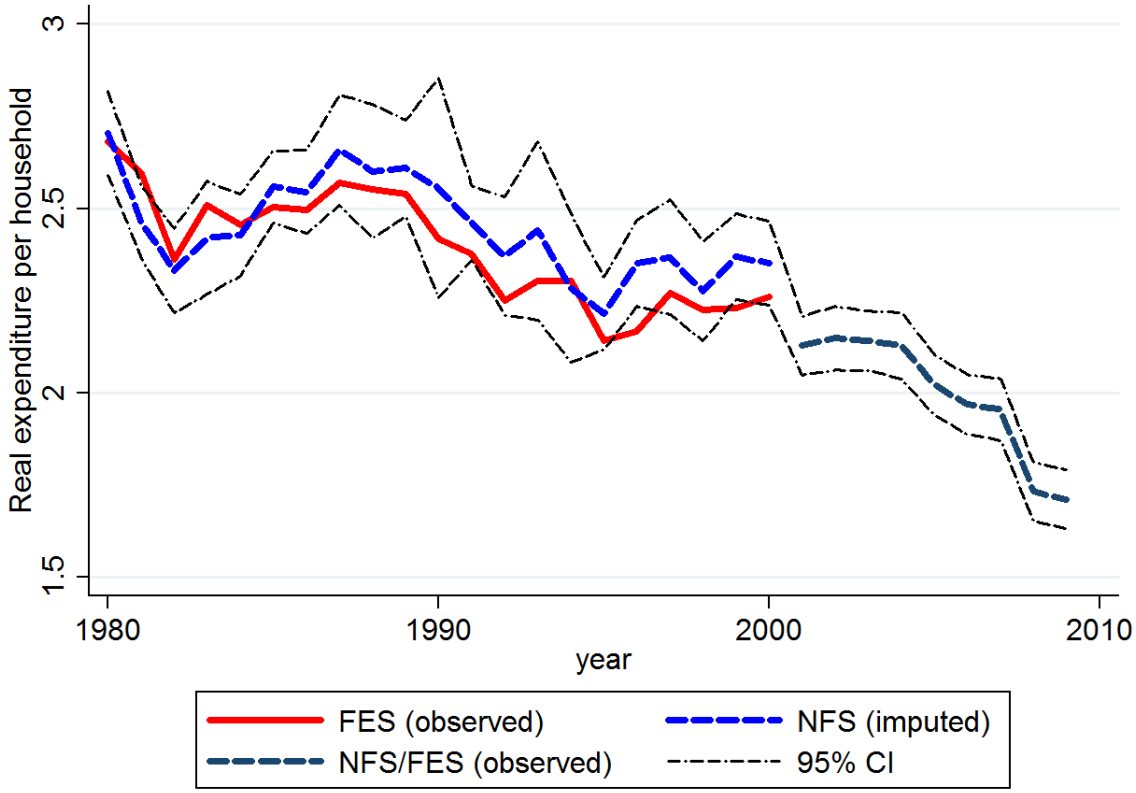
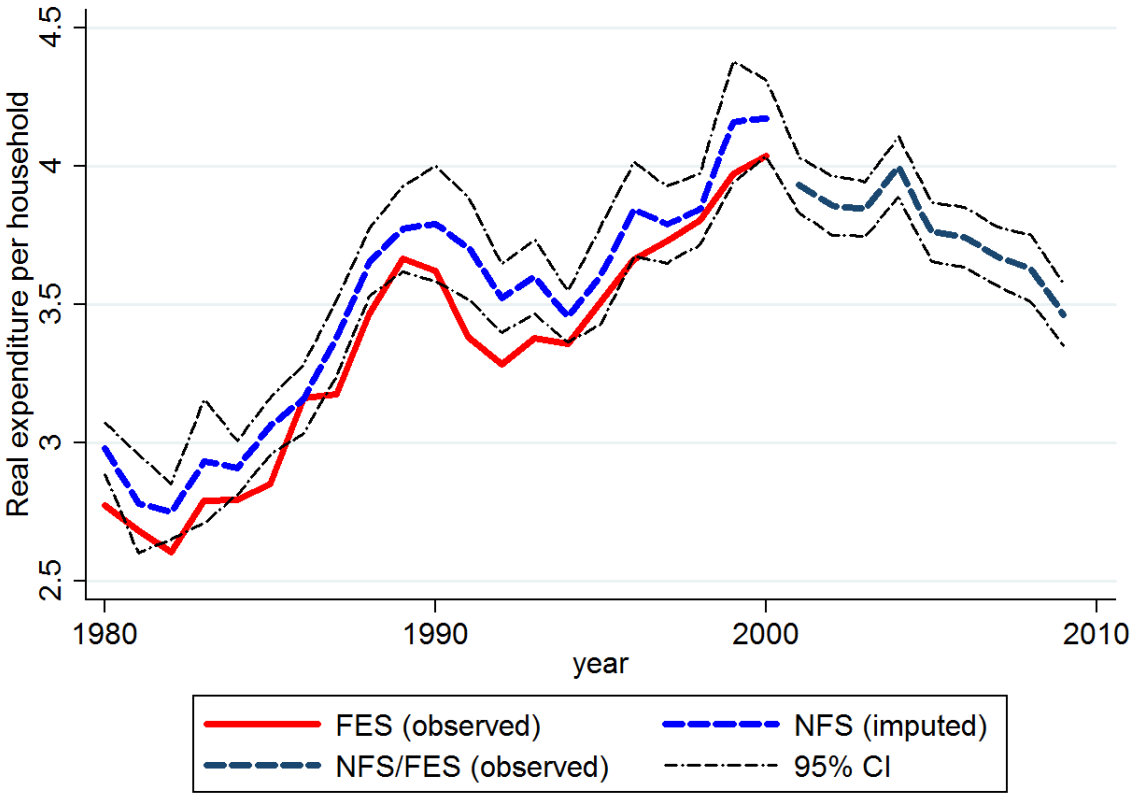


Figure D.4: *Real expenditure on food out*



in the NFS between 2001 and 2009 to do this. We consider a number of alternatives, our preferred model is one that simply uses the mean of $\tilde{\eta}_{it}$ that we observe 2001-2009 and assume that the value was the same in all time periods. Alternatively, we relate observable characteristics to the nutrient values and use this to backcast.

Nutrients per pound are $\tilde{\eta}_{it} = \frac{\eta_{it} p_t}{\frac{y_{it}}{q_{it}}}$, which we observe 2001-2009 . We compute unit values as the weighted sum of unit values at a more disaggregated level s, $up_{it} = \frac{y_{it}}{q_{it}} = \frac{\sum_s y_{ist}}{\sum_s q_{ist}}$, and calories per GBP as a weighted average: $\eta_{it} = \sum_s w_{ist}\delta_{st}$, where each weight, $w_{ist} = \frac{q_{ist}}{q_{it}}$, is the percentage of the overall quantity consumed (in an aggregate category) that household i consumes of sub-item s. Note that δ_{st} does not vary across households, as it is the nutrient amount contained in 1 kilogram of detailed food group s, while η varies by household due to the household-specific consumption of goods contained in an aggregate food group.

In sum, we construct nutrients per pound spent (in 2005 prices) between 2001 and 2009 as:

$$\tilde{\eta}_{it} = \eta_{it}\frac{p_t}{up_{it}} = \sum_s w_{ist}\delta_{st}\frac{p_t}{up_{it}} = \sum_s \frac{q_{ist}}{q_{it}}\delta_{st}\frac{p_t}{up_{it}} \tag{D.7}$$

During the period 2001-2009, $\delta_{st} \sim \delta_s$, i.e. the sub-item level measure of calories per gram, does not vary much. Furthermore, for food at home, we find only limited time variation in the nutrient conversion factors at the disaggregated level over the long time period between 1980 and 2009. The variation in $\tilde{\eta}_{it}$ thus originates from cross-sectional and time series variation in prices and, even more, in the the weights w, which reflect the composition of food demand in the aggregate categories.

We model $\tilde{\eta}_{it}$ as a function of detailed household characteristics, prices and expenditure patterns to capture food demand and diet changes which drive changes in $\tilde{\eta}_{it}$:

$$\tilde{\eta}_{it} = \alpha + \beta D_{it} + \boldsymbol{\gamma p_t} + \zeta y_{mit} + \zeta y_{it} + y_{it}D_{it} + \epsilon_{it} \tag{D.8}$$

where

$\boldsymbol{p_t}$ is a vector of prices (food in, eating out and alcohol),

D_{it} denote household characteristics,

y_{mit} is (real) total food expenditures (and expenditures squared), and

y_{it} is (real) expenditure on the food group (eating out or alcohol).

In estimation, especially for the category alcohol, we need to account for the frequency of zeros, due to abstinence or purchase infrequency. We estimate the model using the Generalised Linear Model (GLM) estimator which maximizes the log-likelihood:

$$Q(\theta) = \sum_{i=1}^{N} [a(m(x_i, \beta)) + b(y_i) + c(m(x_i, \beta))] \tag{D.9}$$

where $m = E(y|x)$ is the conditional mean of y, $a(.)$ and $c(.)$ correspond to different members of the linear exponential family and $b(.)$ is a normalizing constant. We use a log-link function, assuming that $a(m(x_i, \beta)) = \ln(E(y|x)) = x\beta$, using the Newton Raphson maximum likelihood method and assuming a Gaussian distribution of $\tilde{\eta}_{igt}$.

Figure D.5 shows fitted versus observed calories per GBP (in 2005 prices); they suggest that the model does a good job of predicting the trends in the observed period. Calories per pound spent on eating out is largely flat through time, with the exception of the period between the mid 1990s and 2000, in which calories per pound spent on eating out fall distinctively (for the average from 360 calories to 300 calories per pound spent). This is mainly due to an increase in the price for eating out during these years.

Figure D.6 shows equivalent results for alcohol. We match the trends and levels reasonably well; we under-predict by around 5 calories per household per pound spent. It can be seen in Figure D.6 that calories per pound spent on alcohol items increases strongly over the last decade, by around 20% between 2001 and 2009. This reflects a shift from alcohol consumed out to alcohol consumed at home and from beer to wine and spirits, both patterns associated with cheaper calories.

Figures D.8 and D.8 show that, while our imputed values (necessarily) have a lower variance than the observed data, since we cannot capture unobserved heterogeneity across households in our imputation method, we nonetheless capture skewness and mean quite well.

D.3 Food out, takeaways, soft-drinks and confectionery

In this section we present more details on the composition of this aggregate category for the last 10 years of data. Note that we observe calories directly from the survey for the period 2001-2009 so we can look at each of the categories separately: food out, takeaways and soft-drinks and confectionery. Calories from the three

Figure D.5: *Calories per GBP spent (in 2005 prices) for food out)*

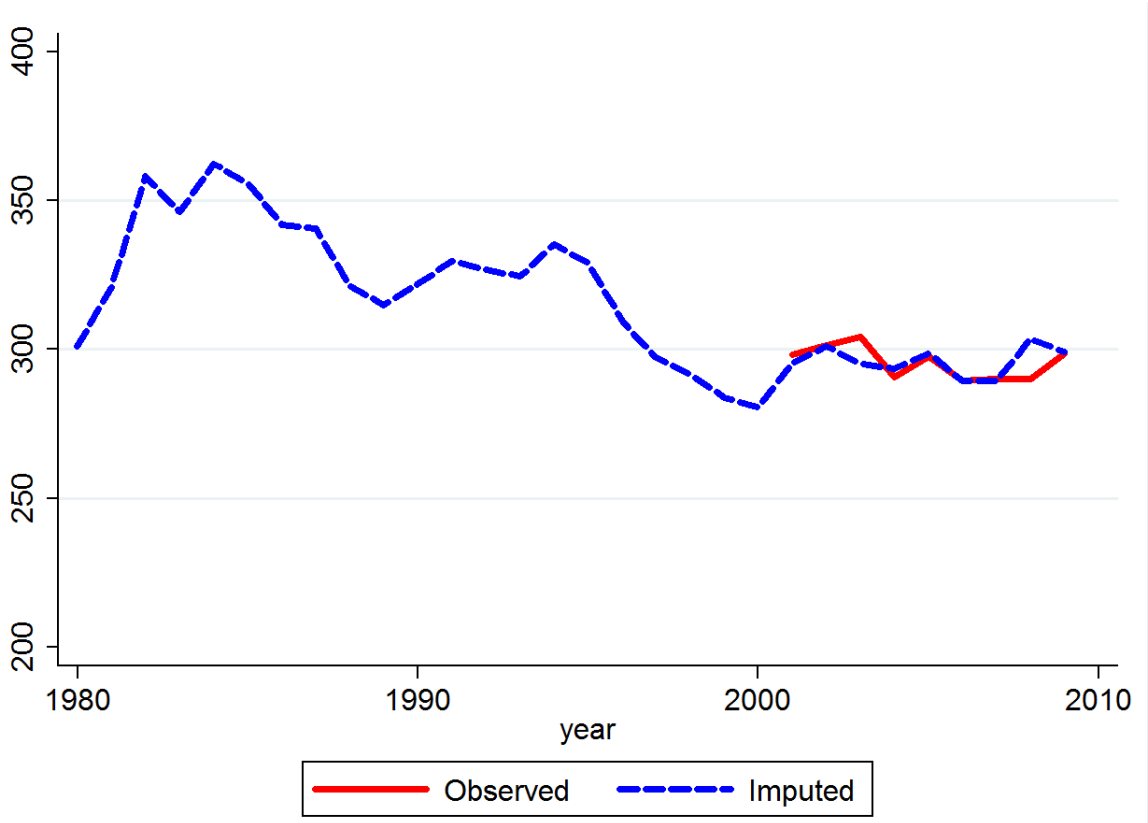


Figure D.6: *Calories per GBP spent (in 2005 prices) for alcohol*

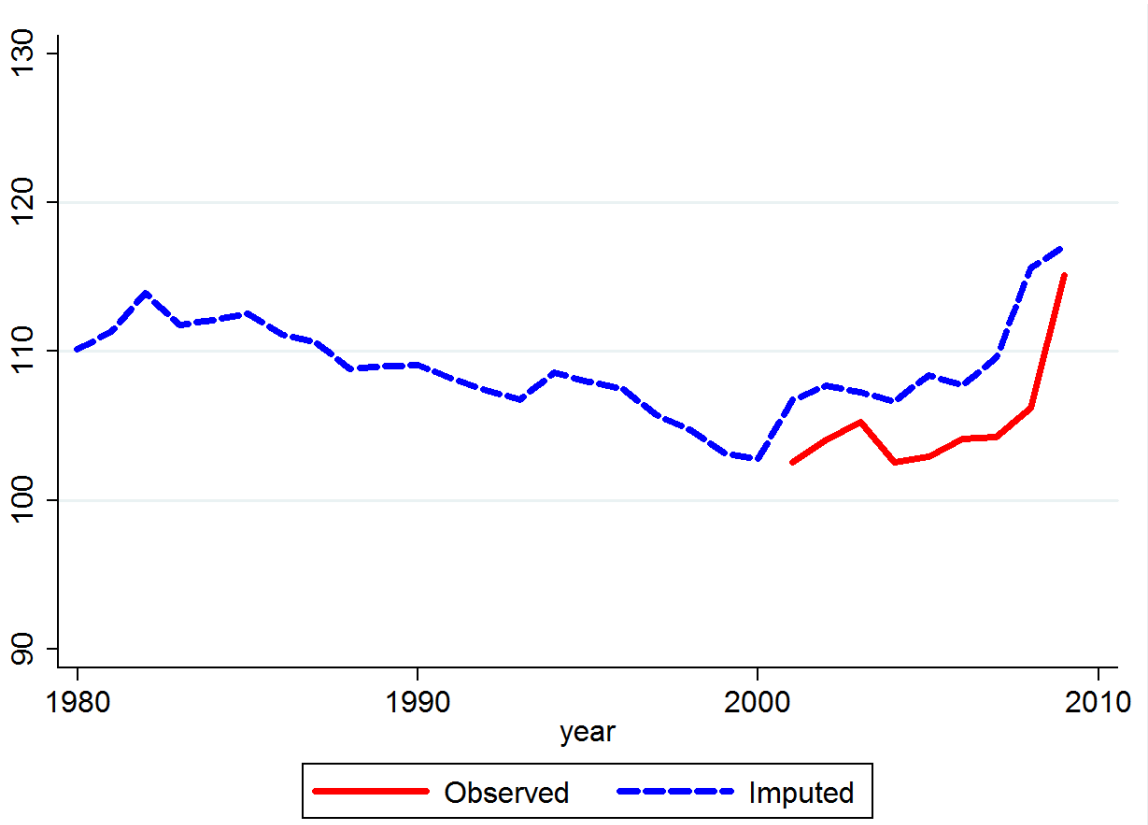


Figure D.7: *Calories per GBP spent (in 2005 prices) for alcohol*

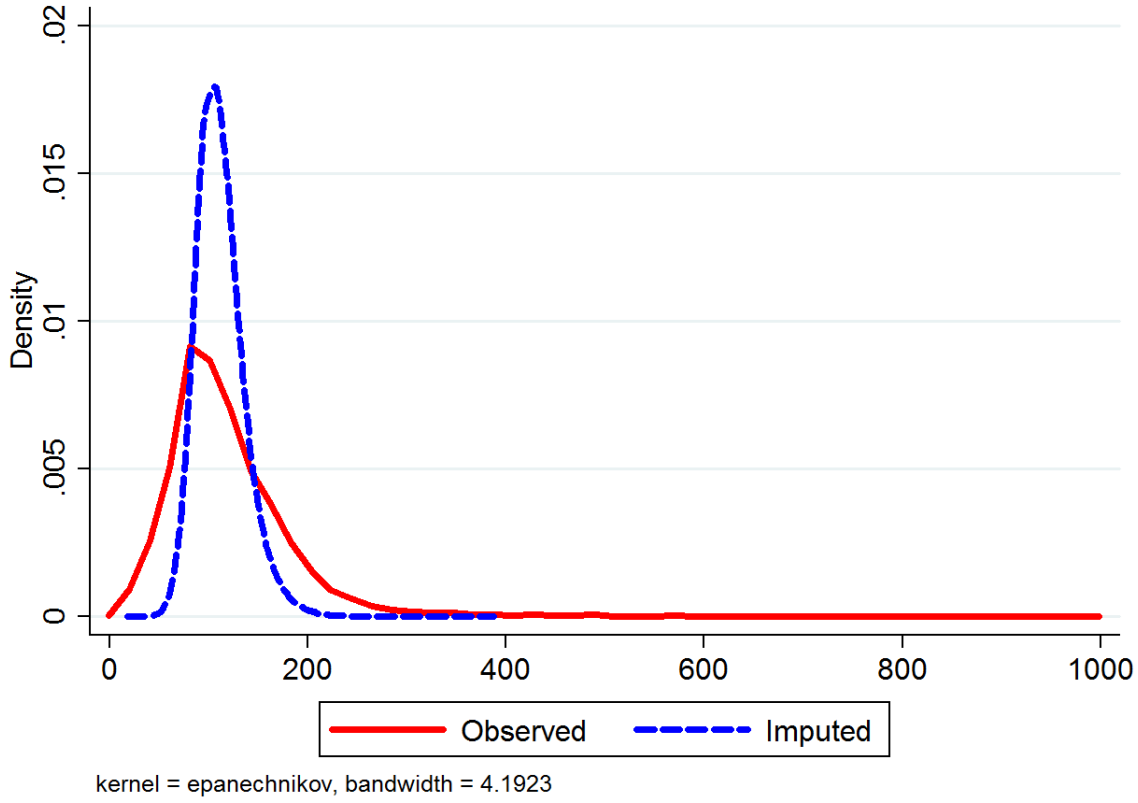
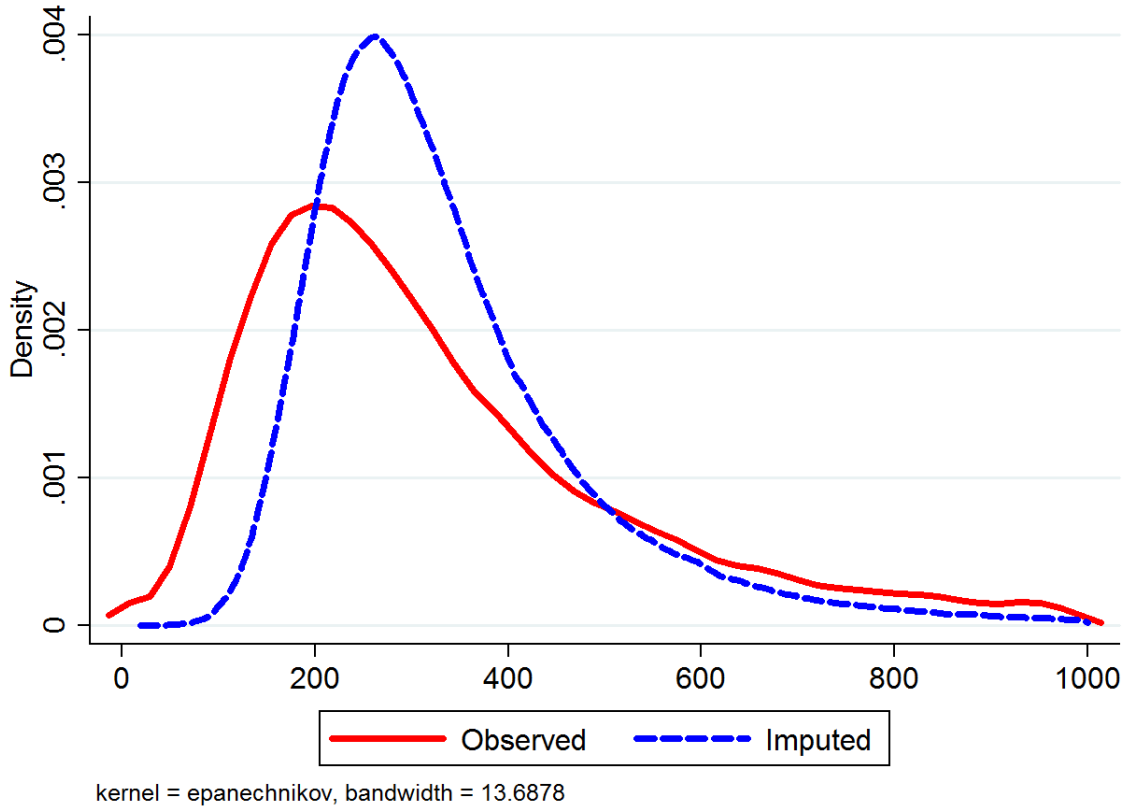


Figure D.8: *Calories per GBP spent (in 2005 prices) for food out*



categories declined between 2001 and 2009: food out declined by 18%, takeaways by 17% and soft-drinks and confectionery by 7%.

Figure D.9: *Calories per household per day: Food out, takeaways, soft-drinks and confectionery (2001-2009)*

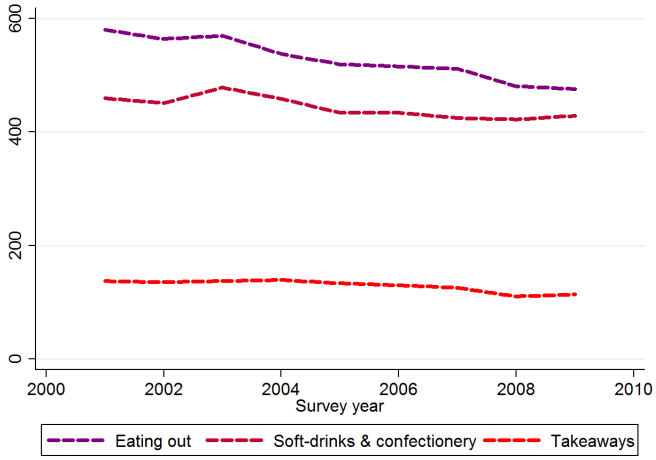


Figure D.10 shows average calories from eating out over time from our two imputation methods and the observed values during the period 2001-2009. Our imputation under-predict calories from eating out by between 50 to 100 calories per household per day.

Figure **??** shows a similar picture by household type. Figure D.11 shows the distribution.

D.4 Alcohol

Figure D.12 shows average calories from alcohol over time from our two imputation methods and the observed values during the period 2001-2009.

Figure **??** shows a similar picture by household type. Figure D.11 shows the distribution.

There is a decline in calorie purchases from alcohol over the period 1980-2009. We also show observed calorie purchases for beer and wine & spirits during the period 1992-2009 and expenditure data from the Family Expenditure Survey (FES) to give more details on the expenditure changes during the period 1980 to 2009.

While data on calories from alcohol consumed at home are available from 1992, data for alcohol consumed out are available from 2001. Figure D.14 shows daily calorie purchases per household for beer and wine & spirits depending on whether consumption is at home or out. The overall decline in calories from alcohol is

Figure D.10: *Calories from food out*

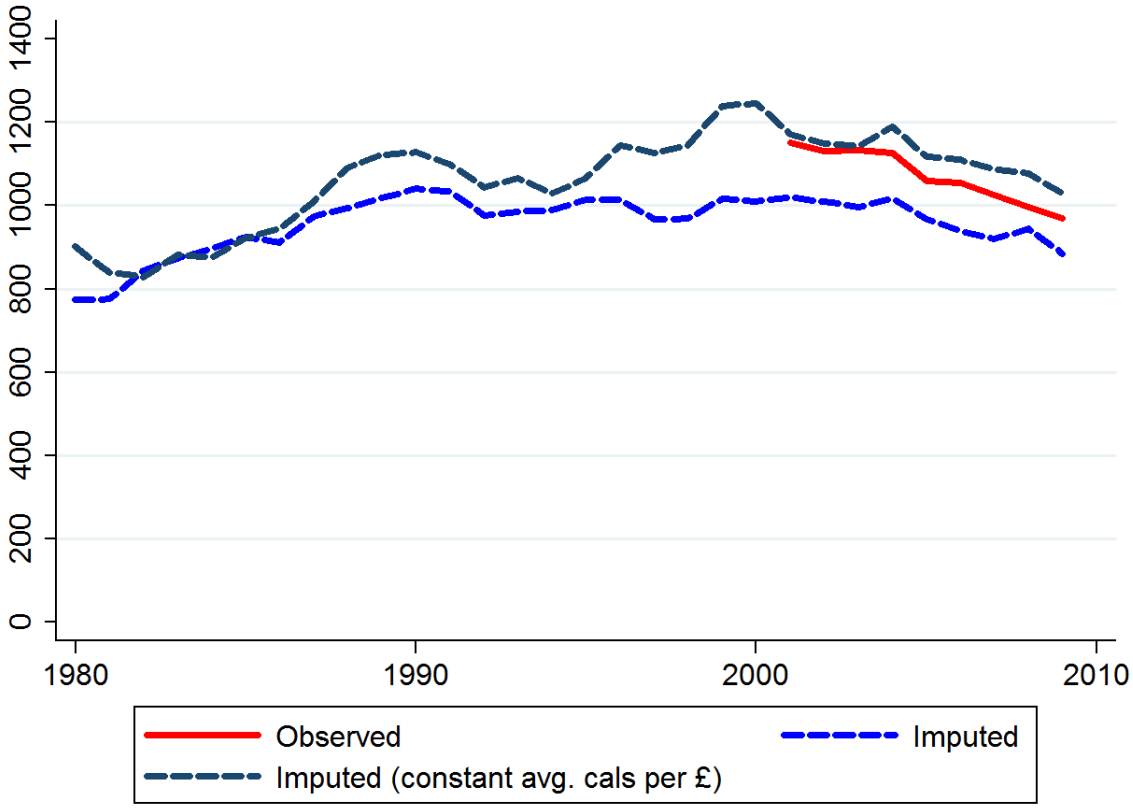


Figure D.11: *Calories from food out*

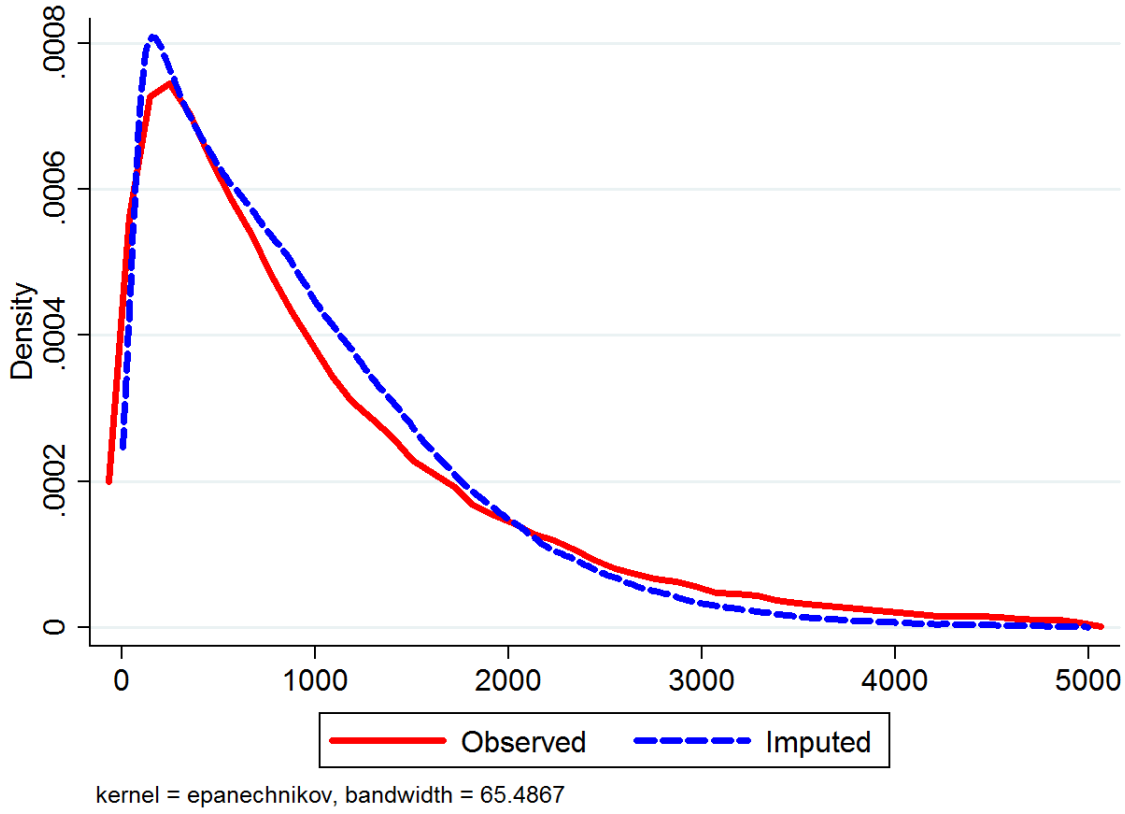


Figure D.12: *Calories from Alcohol*

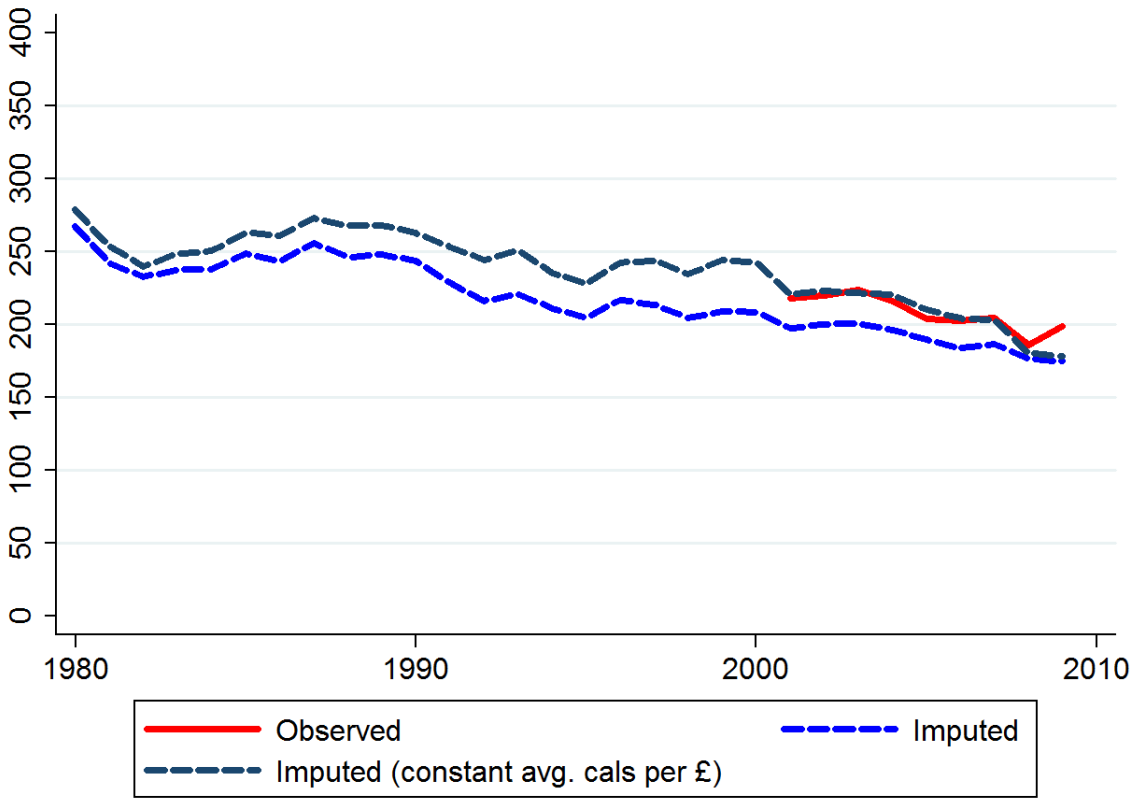
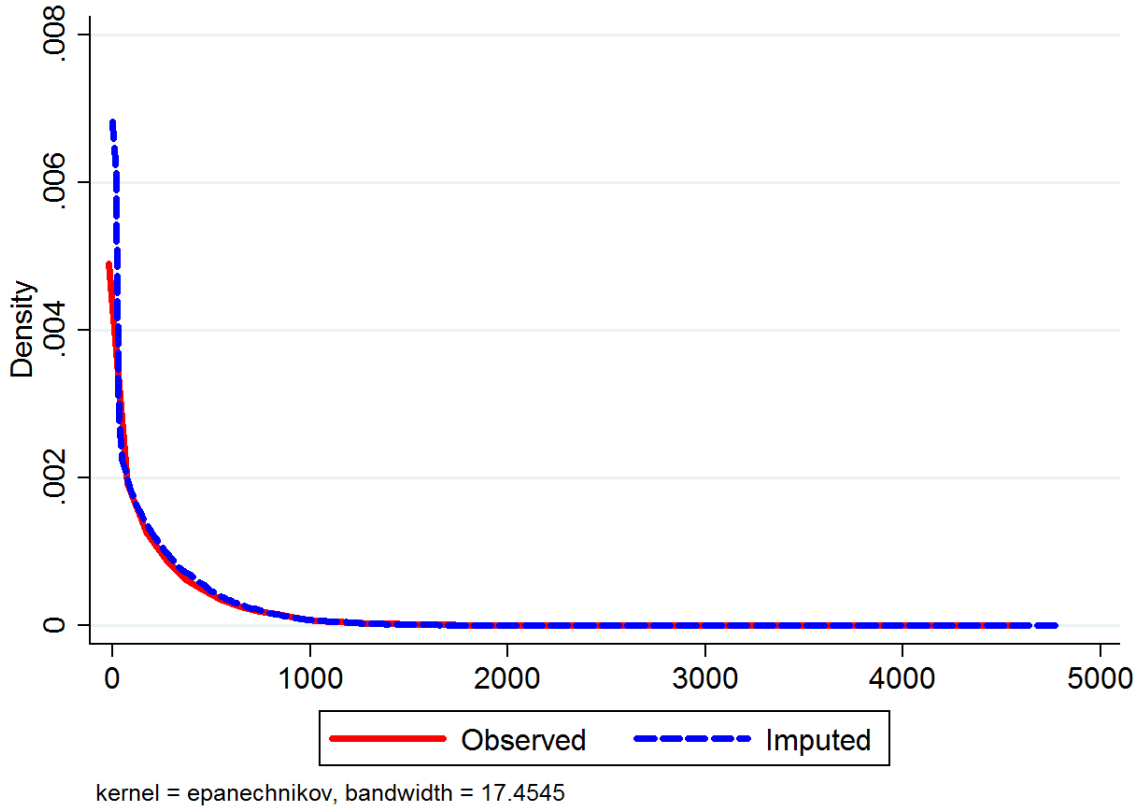


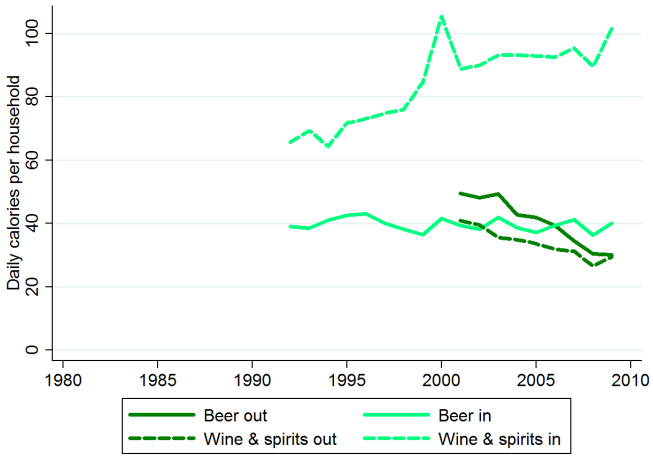
Figure D.13: *Calories from Alcohol*



due to the decline in consumption out: beer consumed on premises declined 39% and wine & spirits declined 29% between 2001 and 2009. On the other hand, while calories from beer consumed at home remained relatively flat beetween 1992 and 2009, calories from wine & spirits consumed at home increased by 55%. While alcohol consumed out represented 41% of total calories from alcohol, of which 23% were from beer and 19% from wine &spirits, by 2009 the figure is just 30%. Calories from wine & spirits consumed at home are now 51% of total calories from alcohol.

The composition of calories from alcohol has changed: households are substituting from beer to wine and from consuming alcohol out of the home to drinking at home. While the former cannot be explained by prices (and potentially reflects changes in the availability of wine and/or preference shifts), the latter coincides with strong increases in the price of alcohol consumed outside the home over time, while the price of alcohol for home consumption has remained comparatively flat since 1990. While this has led to a very small increase in calories per pound spent due to the lower price of alcohol bought for home consumption, quantities of (liquid containing alcohol, not of alcohol itself) consumed have decreased. For example, if the choice between a standard sized glass of wine and a pint of lager is made in favor of the wine, consumed calories fall from 200 per pint of beer to 160 per 0.2l wine glass.

Figure D.14: *Daily calories from alcohol: by type (per household)*

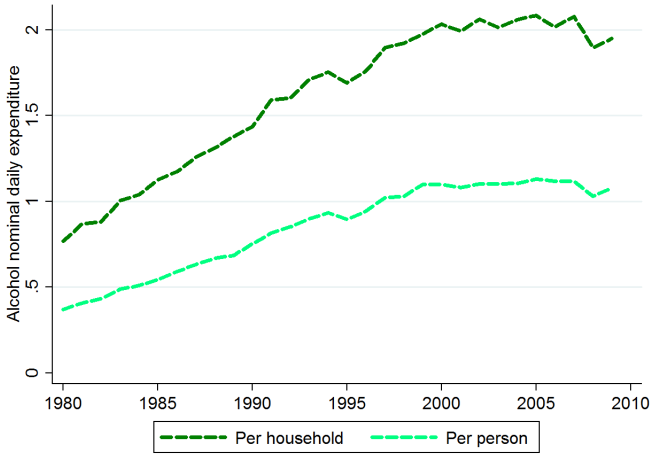


Source: National Food Survey-Expenditure and Food Survey-Living Cost and Food Survey

This pattern is also present when looking at expenditure. While alcohol nominal expenditure has steadily grown until 2000, both average alcohol nominal expenditure per household and per person has remained con-

stant during the last decade (see Figure D.15). This is explained by two features: the decline of expenditure on beer and the increase of expenditure on wine and spirits (see Figure D.16) and the decline of expenditure on alcohol consumed out and the increase of expenditure on alcohol consumed at home (see Figure D.17). We observe then a shift from beer to wine and spirits and from consumption out to consumption at home. Finally, Figure D.18 shows that the decline of expenditure on alcohol consumed out is explained by a sharp decline in expenditure on beer that is not offset by the increase in wine and spirits consumed out.

Figure D.15: *Total alcohol nominal daily expenditure (per household and per person)*



Source: Family Expenditure Survey

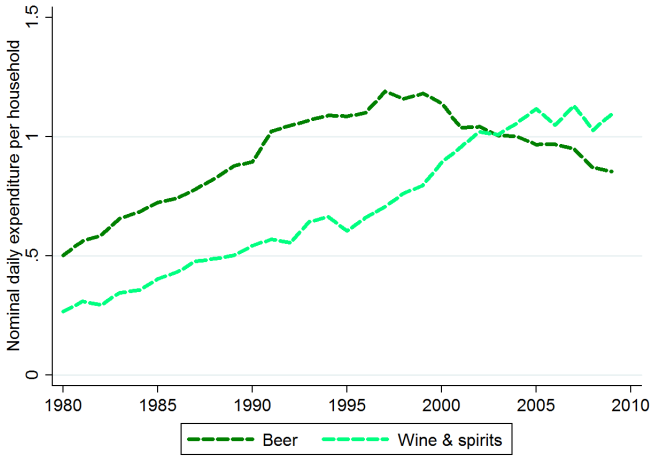
A second robustness check for our estimation of the quantities purchased of alcohol can be performed using annual alcohol clearances by Her Majesty Revenue and Customs (HMRC). HMRC reports historic series of the amount of good cleared for consumption as well as the amount of duty collected on alcohol since 1980. Figure D.19 shows an index of the amount of litres of alcohol cleared by HMRC between 1980 and 2009 together with an index of real alcohol expenditure from the FES. The base year for the index is 1980.

E Expectations

The expected value of the product of two random variables is,

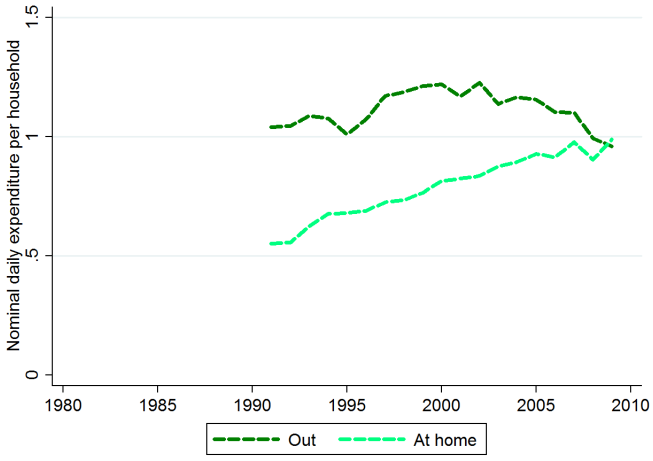
$$E(XY) = E(X)E(Y) + Cov(X, Y) \tag{E.1}$$

Figure D.16: *Nominal daily expenditure: beer and wine & spirits (per household)*



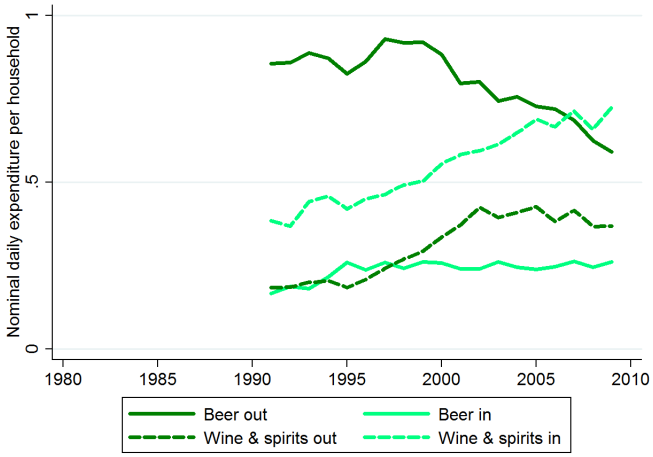
Source: Family Expenditure Survey

Figure D.17: *Nominal daily expenditure: alcohol at home and out (per household)*



Source: Family Expenditure Survey

Figure D.18: *: Nominal daily expenditure: alcohol at home and out by type (per household)*



Source: Family Expenditure Survey

Figure D.19: *Alcohol: Real expenditure in the FES and amount of good cleared by HMRC (1980=100)*

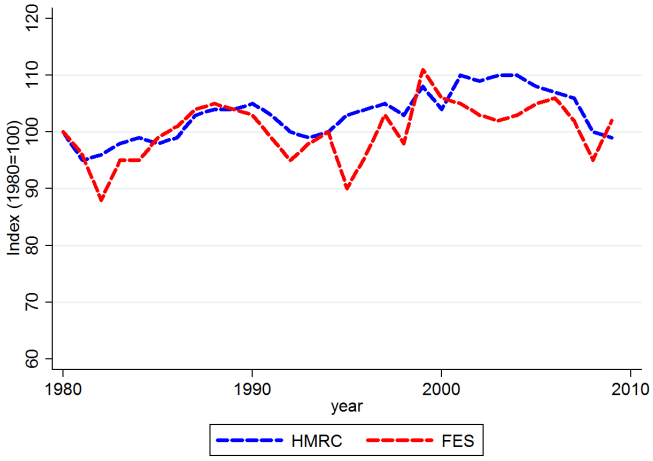
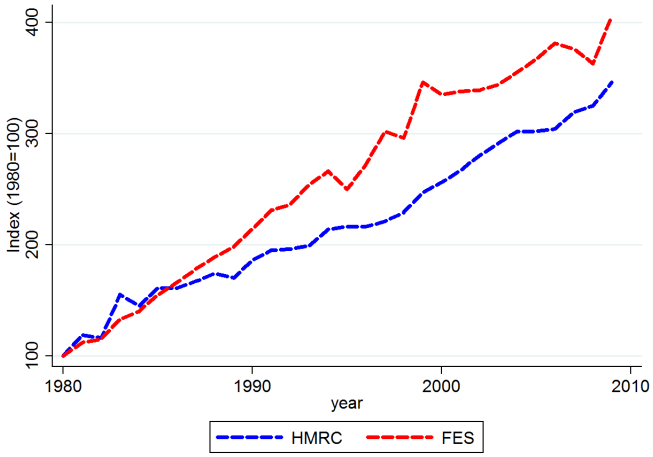


Figure D.20: *Alcohol: Expenditure in the FES and duty collected by HMRC (1980=100)*



and

$$E\left(\frac{X}{Y}\right) = E(X)\frac{1}{E(Y)} + Cov(X, \frac{1}{Y}) \tag{E.2}$$

or, per Taylor approximation:

$$E\left(\frac{X}{Y}\right) \approx \frac{E(X)}{E(Y)} - \frac{Cov(X,Y)}{E(Y)^2} + \frac{Var(Y)E(X)}{E(Y)^3} \tag{E.3}$$

The expected value of the product of three random variables (assuming multivariate normality) is,

$$E(XYZ) = E(X)E(Y)E(Z) + E(X)Cov(Y,Z) + E(Y)Cov(X,Z) + E(Z)Cov(X,Y) \tag{E.4}$$

To show this define:

$$U = YZ \tag{E.5}$$

Substitute:

$$E(XU) = E(X)E(U) + Cov(X, U)$$

$$E(XYZ) = E(X)E(YZ) + Cov(X, YZ)$$

$$E(XYZ) = E(X)\left[E(Y)E(Z) + Cov(Y, Z)\right] + Cov(X, YZ)$$

The covariance of a product of random variables (Bohrnstedt and Goldberger (JASA, 1969)) who use two products (i.e. 4 RVs):

$$\begin{aligned} Cov(XY, UV) = {} & E(X)E(Y)C(Y, V) + E(X)E(V)C(Y, U) \\ & + E(Y)E(U)C(X, V) + E(Y)E(V)C(X, U) \\ & + E(\Delta X \Delta Y \Delta U \Delta V) \\ & + E(X)E(\Delta Y \Delta U \Delta V) + E(Y)E(\Delta X \Delta U \Delta V) \\ & + E(U)E(\Delta X \Delta Y \Delta V) + E(V)E(\Delta X \Delta U \Delta Y) \\ & - Cov(X, Y)Cov(U, V) \end{aligned}$$

Under multivariate normality, all third moments vanish:

$$Cov(XY, UV) = E(X)E(Y)C(Y, V) + E(X)E(V)C(Y, U)$$

$$+ E(Y)E(U)C(X, V) + E(Y)E(V)C(X, U)$$

$$+ E(\Delta X \Delta Y \Delta U \Delta V)$$

$$- Cov(X, Y)Cov(U, V)$$

and according to Anderson(1958):

$$E(\Delta X \Delta Y \Delta U \Delta V) = Cov(X, Y)Cov(U, V) + Cov(X, U)Cov(Y, V) + Cov(X, V)Cov(Y, U) \tag{E.6}$$

so that the above simplifies to:

$$Cov(XY, UV) = E(X)E(Y)C(Y, V) + E(X)E(V)C(Y, U) \tag{E.7}$$

(E.8)

$$+ E(Y)E(U)C(X, V) + E(Y)E(V)C(X, U)$$

(E.9)

$$+ Cov(X, U)Cov(Y, V) + Cov(X, V)Cov(Y, U)$$

Now with only one product, not two, we adapt to our case: For Y=1:

$$E(Y) = 1 \tag{E.10}$$

$$Cov(\cdot, Y) = 0 \tag{E.11}$$

$$\Delta Y = 0 \tag{E.12}$$

This simplifies to:

$$Cov(X, UV) = E(U)Cov(X, V) + E(V)Cov(X, U) \tag{E.13}$$

In our notation:

$$Cov(X, YZ) = E(Y)Cov(X, Z) + E(Z)Cov(X, Y) \tag{E.14}$$

The expected value of the product of three random variables is,

$$E(XYZ) = E(X)\left[E(Y)E(Z) + Cov(Y, Z)\right] + E(Y)Cov(X, Z) + E(Z)Cov(X, Y) \tag{E.15}$$

So the expectation of a product of RVs is the product of the expected values of each RV plus all combinations of expectation times the covariance:

$$E(XYZ) = E(X)E(Y)E(Z) + E(X)Cov(Y, Z) + E(Y)Cov(X, Z) + E(Z)Cov(X, Y) \tag{E.16}$$

E.1 Single index of time use

Consider a single index of time use ($j = 1$), in which case $H = 1$ and 5.1 becomes,

$$W_{it+1} = W_{it} + C_{it} - MET_{it}BMR_{it}. \tag{E.17}$$

We can solve for the unknown MET_{it},

$$MET_{it} = \frac{W_{it} - W_{it+1} + C_{it}}{BMR_{it}}. \tag{E.18}$$

Using data on C_{it}, W_{it}, and W_{it+1} we can compute the unknown MET, which gives us an index of the strenuousness of combined time use. However, we do not observe weight in period t and period $t + 1$ for the same individual, nor do we observe W_{it} and C_{it} for the same individual over this time period. We do

observe these data in population surveys. Taking expectations we obtain (see Appendix E),

$$E[MET_{it}] = E\left[\frac{W_{it}}{BMR_{it}}\right] \tag{E.19}$$

$$- E[W_{it+1}]E\left[\frac{1}{BMR_{it}}\right] + Cov\left[W_{it+1}, \frac{1}{BMR_{it}}\right]$$

$$+ E[C_{it}]E\left[\frac{1}{BMR_{it}}\right] + Cov\left[C_{it}, \frac{1}{BMR_{it}}\right]$$

Data sources are:

- HSE: $E[W_{it}/BMR_{it}]$, $E[W_{it+1}]$, $E[1/BMR_{it}, COV[W_{it}/BMR_{it}]$ which we use as $COV[W_{it+1}/BMR_{it}$ $E[1/BMR_{it}]$
- LCFS/NFS/EFS: $E[C_{it}]$
- Kantar: $Cov[C_i, BMR_i]$

All expectations are taken over gender and decade born in, except Kantar which is over gender.

F Basal Metabolic Rate (BMR)

see

`http://en.wikipedia.org/wiki/Basal_metabolic_rate`

Mifflin St Jeor equation for BMR:

$$P = (10weightkg + 6.25heightcm + 5ageyears + x) \tag{F.1}$$

where $x = 5$ for men and $x = -161$ for women. This is in kcal per day so we need to divide by 7716 to make in kg and multiply by number of days in t.

Mifflin, MD; St Jeor, ST; Hill, LA; Scott, BJ; Daugherty, SA; Koh, YO (1990). "A new predictive equation for resting energy expenditure in healthy individuals". The American journal of clinical nutrition

51 (2): 241–7

G References

References

Ainsworth, B., W. Haskell, M. Whitt, M. Irwin, A. Swartz, S. Strath, W. O'Brien, D. Bassett, K. Schmitz, P. Emplaincourt, D. Jacobs, and A. Leon (2000). Compendium of physical activities: An update of activity codes and met intensities. *Medicine and Science in Sports and Exercise 32*(Suppl), 498–516.

Attanasio, O., E. Battisin, and A. Leicester (2006). From micro to macro, from poor to rich: Consumption and income in the UK and the US. *presented at the University of Michigan National Poverty Center conference, Consumption, Income and the Well-Being of Families and Children*.

Attanasio, O. and L. Pistaferri (2013). What do we know about consumption inequality? some evidence using the new psid consumption measure. *mimeo*.

Baum II, C. L. and S.-Y. Chou (2011). The soci-economic causes of obesity. *NBER Working Paper 17423*.

Bleich, S., D. Clutler, C. Murray, and A. Adams (2008). Why is the developed world so obese? *Annual Review of Public Health 29*, 273–295.

Brewer, M. and C. O'Dea (2012). Measuring living standards with income and consumption: evidence from the UK.

Brunello, G., P. Michaud, and A. S. de Aldeano (2009). The rise of obesity in europe: an economic perspective. *Economic Policy*, 553–596.

Chesher, A. (1998). Individual demands from household aggregates: time and age variation in the composition of diet. *Journal of Applied Econometrics 13*(5), 505–524.

Chou, S., M. Grossman, and H. Saffer (2004). An economic analysis of adult obesity: results from the behavioral risk factor surveillance system. *Journal of Health Economics 23*, 565–587.

Costa, D. (2000). From mill town to board room: The rise of women's paid labor. *Journal of Economic Perspectives 14*(4), 101–122.

Crane, N., C. Lewis, and E. Yetley (1992). Do time trends in food supply levels of macronutrients reflect survey estimates of macronutrient intake? *American Journal of Public Health 82*(6), 862–866.

Cutler, D. M., E. L. Glaeser, and J. M. Shapiro (2003). Why Have Americans Become More Obese? *The Journal of Economic Perspectives 17*(3), 93–118.

Deaton, A. (2005). Measuring poverty in a growing world (or measuring growth in a poor world). *The Review of Economics and Statistics 87*(1), 1–19.

Deaton, A. and J. Dreze (2010). From calorie fundamentalism to cereal accounting. *Economic and Political Weekly 45*, 87–92.

DEFRA (2011). Adjustments to National Food Survey estimates.

Dobbs, R., C. Sawers, F. Thompson, J. Manyika, J. Woetzel, P. Child, S. McKenna, and A. Spatharou (2014). Overcoming obesity: an initial economic analysis. *McKinsey Global Institute Discussion Paper*.

Dowler, E. A. and Y. O. Seo (1985). Assessment of energy intake: Estimates of food supply v measurement of food consumption. *Food Policy 10*(3), 278–288.

Duffey, K. J. and B. M. Popkin (2011, 06). Energy density, portion size, and eating occasions: Contributions to increased energy intake in the united states, 1977-2006. *PLoS Med 8*(6).

Ewing, R., T. Schmid, R. Killingsworth, Z. Zlot, and S. Raudenbush (2003). *American Journal of Health Promotion 18*, 47–57.

Finkelstein, E. A. and L. Zuckerman (2008). *The Fattening of America How The Economy Makes Us Fat, If It Matters and What To Do About It*. New Jersey, USA: Wiley.

Horton, N. and K. Kleinman (2007). Much ado about nothing: a comparison of missing data methods and software to fit incomplete data regression models. *The American Statistician 61*(1), 79–90.

Jacobs, K. and D. A. Sumner (2002). The food balance sheets of the food and agriculture organization: A review of potential ways to broaden the appropriate uses of the data. Technical report, Manuscript. Review sponsored by FAO.

Ladabaum, U., A. Mannalithara, P. Myer, and G. Singh (2014). Obesity, abdominal obesity, physical activity, and caloric inake in us adults: 1988-2010. *The American Journal of Medicine*.

Lakdawalla and Philipson (2009). The growth of obesity and technological change. *Economics & Human Biology 7*(3), 283–293.

Lakdawalla, D., T. Philipson, and J. Bhattacharya (2005). Welfare-enhancing technological change and the growth of obesity. *AEA Papers and Proceedings 95*(2).

Nickell, S. and J. Saleheen (2008). The Impact of iImmigration on Occupational Wages: Evidence from Britain. *Federal Reserve Bank of Boston Working Paper No. 08-6*.

Philipson, T., R. Posner, et al. (2003). The long-run growth in obesity as a function of technological change. *46*(Suppl. 3), s87–s107.

Pomerleau, J., K. Lock, and M. McKee (2003, 6). Discrepancies between ecological and individual data on fruit and vegetable consumption in fifteen countries. *British Journal of Nutrition 89*, 827–834.

Prentice, A. and S. Jebb (1995). Obesity in britain: gluttony or sloth? *British Medical Journal*.

Rashad, I. (2006). Structural estimation of caloric intake, exercise, smoking, and obesity. *The Quarterly Review of Economics and Finance 46*(2), 268–283.

Rodrguez-Artalejo, F., J. Banegas, A. Graciani, R. Hernndez-Vecino, and J. del Rey-Calero (1996). Food supply versus household survey data: nutrient consumption trends for spain, 1958-1988. *European Journal of Epidemiology 12*(4), 367–71.

Rodrigues, S., C. Lopes, A. Naska, A. Trichopoulou, and M. Almeida (2007). Comparison of national food supply, household food availability and individual food consumption data in portugal. *Journal of Public Health 15*(6), 447–455.

Royston, P. and I. R. White (2011, 12). Multiple imputation by chained equations (mice): Implementation in stata. *Journal of Statistical Software 45*(4), 1–20.

Rubin, D. B. (1987, June). *Multiple Imputation for Nonresponse in Surveys* (99 ed.).

Rubin, D. B. (1996). Multiple Imputation After 18+ Years. *Journal of the American Statistical Association 91*(434), 473–489.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. Monographs on statistics and applied probability ; 72. Chapman & Hall.

Schafer, J. L. (1999). Multiple imputation: a primer. *Statistical Methods in Medical Research 8*(1), 3–15.

Seale, J., A. Regmi, and J. Bernstein (2003). The international evidence on food consumption patterns. *USDA Technical Bulletin Number 1904*.

Serra-Majem, L., D. MacLean, L. Ribas, D. Brul, W. Sekula, R. Prattala, R. Garcia-Closas, A. Yngve, M. Lalonde, and A. Petrasovits (2003). Comparative analysis of nutrition data from national, household, and individual levels: results from a who-cindi collaborative project in canada, finland, poland, and spain. *Journal of Epidemiology & Community Health 57*(1), 74–80.

Sturm, R. and R. An (2008). Obesity and economic environments. *118*(531), 1385–1417.

Swinburn, B., G. Sacks, and E. Ravussin (2009). Increased food energy supply is more than sufficient to explain the us epidemic of obesity. *The American Journal o fClincial Nutrition 90*(6), 1453–1456.

Tudor-Locke, C., B. Ainsworth, T. Washington, and R. Troiano (2011). Assigning metabolic equivalent values to the 2002 census occupational classification system. *Journal of Physical Activity and Health 8*(4), 581–586.

Vink, G., L. Frank, J. Pannekoek, and S. van Buuren (2014). Predictive mean mathcing imputation of semicontinous variables. *Statistica Neerlandica 68*(1), 61–90.