Do healthier food baskets cost more?

Cherry Law[1,2]*, Andrés Sánchez Pájaro[3], Richard Smith[3], Laura Cornelsen[2]

[1] University of Reading, UK

[2] London School of Hygiene & Tropical Medicine, UK

[3] Center for Population Health Research, National Institute of Public Health, Mexico

[4] University of Exeter, UK

*corresponding author

Abstract

This paper uses purchase data from a large representative British household panel to explore the association between the healthiness and cost of food baskets. We classify items purchased that are high in fat, sugar and salt (HFSS) and use the share of calories obtained from these foods to measure the healthiness of the baskets. Our descriptive analysis reveals large variations in the healthiness of food baskets of similar costs. Our empirical results indicate a concave association between the healthiness and cost of food baskets. Buying a basket consisting predominantly of either non-HFSS energy or HFSS energy is likely to be less expensive than a mixed basket, challenging the commonly held view that healthier diets are more expensive than less healthy ones. However, although healthier baskets per se are not more expensive than a healthy basket, the 'distance' to move from predominantly HFSS to predominantly non-HFSS may entail increased costs as households move through the 'mixed basket' zone. Thus fiscal measures could help them to overcome the cost barriers in improving their diets over the short-term.

1. Introduction

Less healthy foods are often reported to be cheaper than healthy foods, leading to concerns over the financial barriers to healthy eating. The question of whether healthier diets cost more is of high health policy interest as it provides insight on how fiscal policies could be used to improve public health. High intake of fruit and vegetables can help reduce the risk of cardiovascular diseases and cancers (Aune et al. 2017). On the other hand, consumption of foods that are high in fat, sugar and/or salt is one of the main risk factors for obesity and a range of chronic conditions (Bechthold et al., 2017; Grosso et al., 2017; Malik et al., 2012; Rosenheck, 2008). In the UK, obesity and overweight are estimated to incur £19 billion health-care costs to the National Health Services and £16 billion costs of productivity loss to the wider society annually (Bell et al., 2023). The UK National Food Strategy demonstrated that highly processed foods, that tend to be unhealthy, are on average three times cheaper per calorie than healthier foods (Dimbleby, 2021)

However, price per calorie has limitations as a measure for comparing the cost of healthier/less healthy foods. It does not account for the fact that healthier foods tend to contain less protein, fat, sugar and salt and hence fewer calories per equivalent of weight. Furthermore, it does not consider other nutrients and foods that are a cheap source of calories may be an expensive source of other nutrients (Carlson & Frazão, 2012). An alternative measure is price per volume, which is equally not ideal for making cost comparisons due to the large variations in serving sizes across food groups (e.g. 100g of ham typically has more servings than 100g of beef burger). Price per serving has also been used to assess the relative cost of healthier foods. Currently, there is no standardised way to define serving size or average portion size. Cost analysis based on price per serving may be misleading as the amount of food people eat at a single sitting varies greatly across occasions and individuals. More importantly, comparing price across individual food groups does not address the question of whether healthier diets cost more as most people do not only consume one or two food groups but a variety of them.

To address this, some studies assess the cost of a diet that meets dietary recommendations. In the UK, Jones et al. (2017) find that meeting six or more government dietary recommendations leads to a cost up to 29% higher than a diet meeting no recommendations. While similar results have been found in the US (Rao et al., 2013) and Belgium (Vandevijvere et al., 2020), Lee et al., (2020) show that healthy diets would be 15-17% less expensive than current (unhealthy) diets in Australian cities. Incorporating the environmental consideration, Batis et al. (2021) find that the cost of the EAT–Lancet healthy reference diet was 21% lower than that of the Mexican dietary guideline's basket, and 40% lower than that of the basket reflecting current intake. These studies typically link dietary patterns to national food price databases or average prices collected from stores to estimate the dietary costs. These externally sourced prices however do not capture the large price variations caused by the quality difference across products within the same food groups. For instance, households could spend less on the same food group by buying non-branded products instead of the branded alternatives. As cost estimates of meeting dietary guidelines only reflect the national average, they are also limited in demonstrating the range of possible costs of healthy and less healthy food baskets that are consumed or purchased by households.

This paper uses an alternative approach, assessing the cost of healthier diets by using product-level household food expenditure data for at-home consumption. These data cover over 35million purchases made by a representative sample of over 30,000 British consumers in 2017 and include information on item prices and its nutritional content. This rich dataset not only provides the precise cost of food baskets purchased by households, but also allows us to evaluate the healthiness of entire food basket rather than focusing on a few food groups. For this, we employ the Nutrient Profiling Model, which has been used by the UK government in food policies, to classify each food item into foods with high fat, sugar and/or salt (HFSS) content and non-HFSS foods. We then compute the HFSS energy share of each monthly food basket, which is the proportion of total calories purchased from HFSS foodsand drinks.

Using this measure, we first document the large variations in the healthiness of food baskets of similar costs, which reflects that food baskets with lower HFSS energy share are not always more expensive. Next, we conduct an empirical analysis to explore how within-household variations in the overall food basket healthiness drive changes in basket costs. While household and postcode-month fixed effects are used to control for the substantial unobserved heterogeneity, biases can still rise from household-specific preference shocks that can have effect on food purchases and thus correlate with the nutritional quality of baskets. Following Dubois et al. (2014), we employ an instrumental variable (IV) approach to address this endogeneity through a household-specific measure of their external food environment.

Both standard panel and IV results show that the cost of food baskets increases with the HFSS energy share at a decreasing rate. This indicates that whether increasing the energy share of non-HFSS foodsincreases the cost of diets depends on the composition of current food basket. Small substitutions towards non-HFSS energy are likely to increase the cost if the current food basket has a high HFSS energy share, in which case only a substantial change would lead to an equivalent costing healthy basket. This non-linear linkage between the cost of food baskets and the associated HFSS energy share is observed among households across various socioeconomic groups.

This paper contributes to the literature on the complexity of food choices. Causal evidence on dietary quality of US households show that differences in neighbourhood food environment (i.e. availability and price of healthy foods) and household food spending only explain a part of the gap in dietary quality between high and low SES groups (Allcott et al., 2019; Hastings et al., 2021). Dubois et al. (2014) show that prices and attributes of food products do not fully explain the observed differences in nutritional composition across countries. Their result highlights the importance of preferences and eating habits in making food purchases. Indeed, our results show that it is possible to have a much healthier basket without spending more if households are willing to make dramatic changes in their diets. However, making big dietary changes are challenging (Hut & Oster, 2022). This implies that while fiscal measures could help households with poor diets to overcome the cost barriers in making small step changes toward a healthier diet, they are not sufficient to deliver healthier populations on their own. Complementary policies are therefore needed to address the non-financial barriers to changing diets.

The rest of the paper is structured as follows. The next section describes our data and explains the construction of our measure of food basket healthiness. Sections 3 and 4 present our conceptual framework and empirical approach respectively. In section 5 we share our results. A final section summarises the results and discusses policy implications.

2. Data

Purchase data

We use household-product level data on food and drink purchases in Great Britain (GB), obtained from the Kantar Fast Moving Consumer Goods (FMCG) panel. Kantar, a consumer insight company, operates a multi-year, open-cohort design panel and collects information on British household purchases of consumer goods. The panel maintains ~30,000 households selected via stratified sampling, with quotas set for GB regions, household size, age of main shopper, number of children and occupational socio-economic status (SES). Food and beverage purchases for take-home consumption are recorded throughout the year by participants using provided hand-held barcode scanners. Non-barcoded items like loose fruits are recorded using pre-specified barcode lists. Nutritional information of products purchased (energy, sugar, salt, protein, fibre, fat and saturated fat content) is collected by Kantar through direct measurements in places of purchase twice a year. In cases where direct information is not available, nutritional values are either copied across from similar products or an average value for the product category is calculated and used instead (between 11.0% and 19.6% of products on average, depending on category (Berger et al., 2019)).

Socio-demographic data on the panel include information on the postcode area of residence[1], household size and presence of children, age and highest education of the main shopper and the SES based on the occupation of the main income earner in the household. Households of respondents with higher and intermediate managerial, administrative or professional occupations are classified as high SES. Middle SES households are those with main shoppers who are skilled manual workers and those with supervisory or clerical and junior managerial, administrative or professional occupations. Households are categorised as low-SES if the respondents are semi-or unskilled manual workers, state pensioners, causal or lowest grade workers or who are unemployed with state benefits (Cornelsen, Berger, et al., 2019). In our analysis, we use data on purchases made by 31723 households in 2017. Summary statistics on their socio-demographic characteristics is provided in Appendix A1.

HFSS energy share (HES)

Since we are interested in the overall dietary pattern rather than daily purchases, we aggregate the food and drink purchases reported to monthly level and refer that as monthly food baskets. The healthiness of these food baskets is measured by its share of energy from HFSS foods, which is computed in two steps. First, we apply the Nutrient Profiling Model developed by the Food Standards Agency (FSA) which scores food and drink products based on their nutritional content (Department of Health, 2011). This model has been used in a number of policies in England (e.g. advertisement restrictions of HFSS foods for children, advertisement restrictions on HFSS foods in Transport for London network; restrictions on location promotions and a proposed ban on volume-based price promotions). Points are given for ‘A’ (energy, saturated fat, total sugar and sodium), and ‘C’ (fruit, vegetables and nut content, fibre and protein) nutrients. A score for the product is then obtained by subtracting points of ‘C’ nutrients from ‘A’ nutrients. Food items with a score of 4 or more points, and drinks with a score of 1 or more points, are classified as HFSS foods. All alcoholic beverages are considered as unhealthy (i.e. HFSS) as the model does not specify how to classify them. Next, we compute the total calorie content from HFSS foodsand drinks purchased by the household in each month, and divide it by the total calorie content purchased from all food and drinks during the same period. This gives us the share of calorie from HFSS foodsin the monthly food baskets, which is referred to as HFSS energy share hereafter. The higher value of this measure is, the less healthy is the monthly food basket as it contains relatively more calories from HFSS food.

In addition to its policy relevance, HFSS energy share has a few advantages over the existing approaches used in assessing the costs of healthy diets. First, it relies on the relative contribution of various foods to the diet instead of the actual calorie content, mitigating the issue that healthy foods are typically low in calorie and is therefore inherently more expensive when price per calorie is used to compare costs across foods. Second, it measures the healthiness of entire food baskets purchased by households, reflecting a more comprehensive view of diets than cost comparison of individual items. Third, this measure captures a more accurate and realistic profile of food baskets as it does not rely on a pre-determined list of food items that meets dietary guidelines. Instead, it assesses the healthiness of food baskets on a continuous scale based on actual purchases made by households.

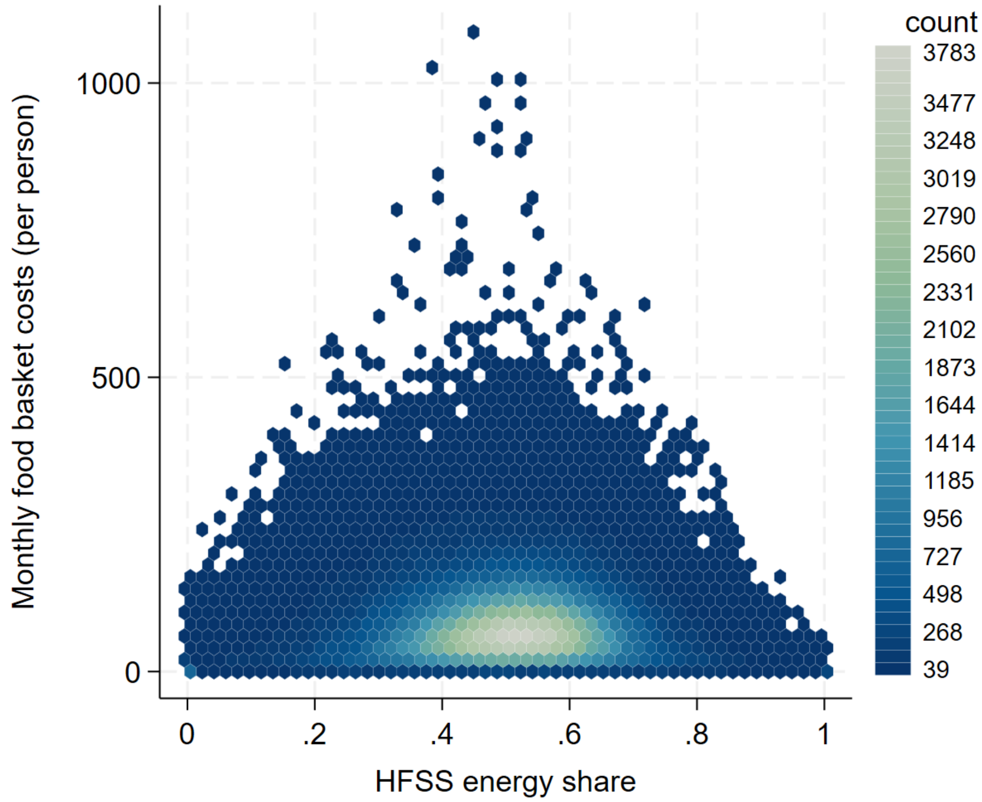
[1] Postcode area is represented by the initial 2–4 characters of the alphanumeric UK postcode.

.

Cost and healthiness of food baskets

Figure 1 shows the distribution of HFSS energy share across all food baskets purchased by our household sample in 2017. The cost of these baskets is the sum of expenditure on the foods and drinks in the basket divided by household size.[2] The colour gradient of the hexagon markers indicates the density of observations. As expected, the healthiness of food baskets is highly heterogeneous across households. The lighter markers indicate that most households purchased baskets with around 40%-60% of HFSS energy share, in other word, a mixture of HFSS and other healthier products. The cost of these food baskets varies from mostly less than £200 to some over £600. Figure 1 also illustrates that food baskets of similar costs can contain predominantly healthy foods with a low HFSS energy share or vice versa, suggesting that the linkage between cost and the healthiness of diets may be more complicated than the widely-held view of healthy foods being more expensive.

Figure 1 Cost and healthiness of take-home food baskets for British households in 2017

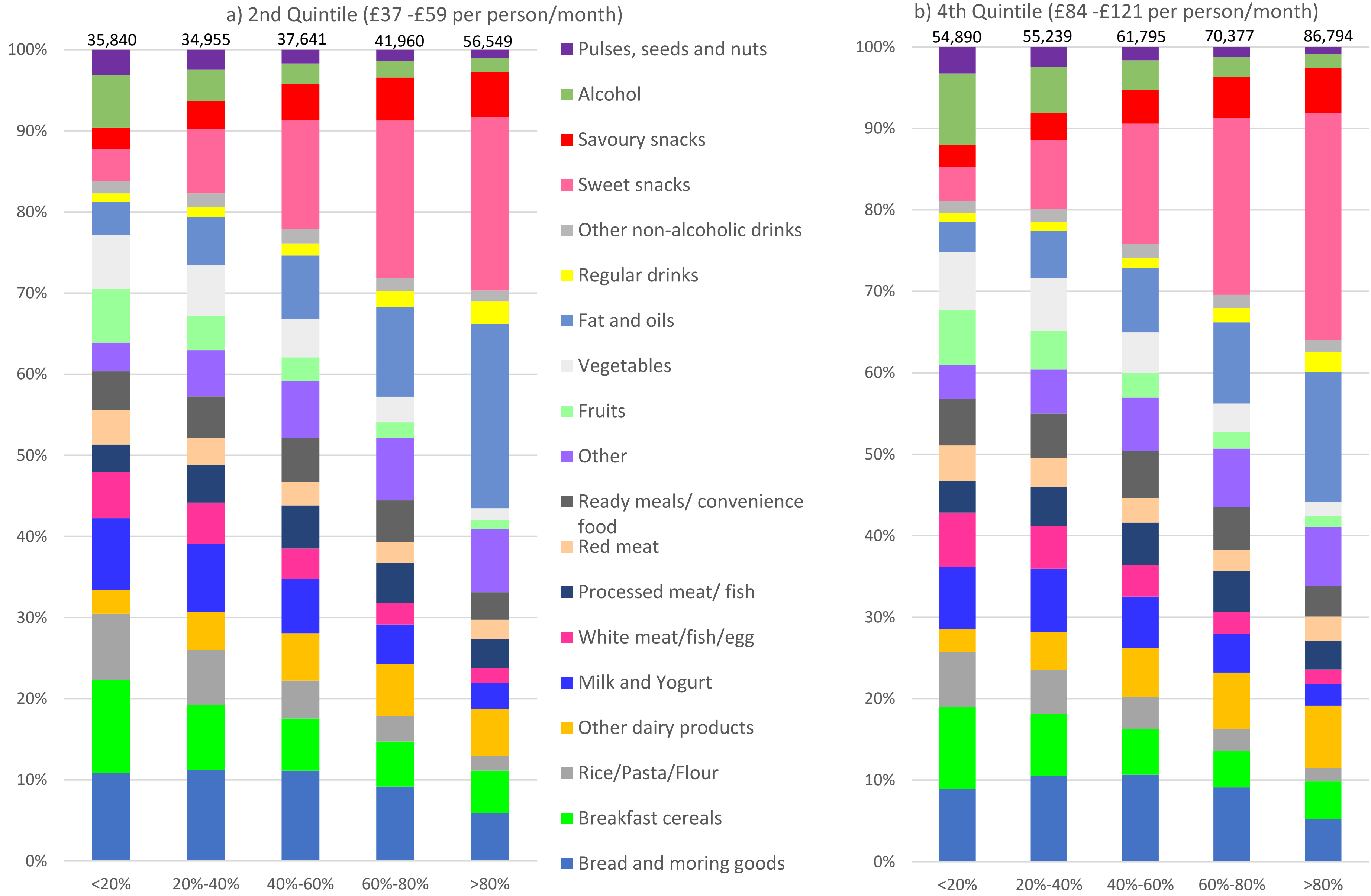


This characterisation however raises the question of how baskets with high HFSS energy share differ from the ones with lower share. Figure 2 displays the composition of food baskets that have similar costs but contain different levels of HFSS energy share. Specifically, we sort all monthly food baskets according to their cost into quintiles and report the share of calories from 19 food groups[3] of those in the 2nd and 4th quintiles of basket cost (i.e. £37-£59 and £84-£121 per person respectively). It is striking that the relative contribution of the food groups to total calories displays similar patterns for both quintiles. For example, breakfast cereals and staples such as bread, rice, pasta, take up around 25-30% of the total calories among baskets with HFSS energy share lower than 40% in both quintiles but only 12-18% in the less healthy baskets with HFSS energy share over 60%. For latter baskets, sweet snacks and, fat and oils, are the top calorie

[2]Expenditure is adjusted to inflation using consumer price index from the Office of National Statistics. As we do not know gender and age of all household members we do not make further adjustments to household composition.
[3] The list of food products included in each food group is provided in the Appendix table A2.

contributor regardless of the quintiles. Overall, the share of calories from each of the food groups is comparatively much more balanced in the healthier baskets (<40% HFSS energy share) and a clear pattern of bigger share from pulses, seeds and nuts, fruits, vegetables, dairy products, and white meat, fish and eggs. Processed meats and ready meals have a similar contribution across the range of baskets and quintiles while the contribution of alcohol to calories is relatively higher in healthier baskets as well as in the 4^{th} quintile. What the figure demonstrates, is the substantial variation in food composition across baskets based on HFSS energy shares but relative similarity when comparing between quintiles. Thus healthier baskets are not necessarily more expensive. In the following sections we will empirically examine the linkage between the cost and healthiness of monthly food baskets purchased by British households.

Figure 2 Calorie distribution of food groups within baskets of similar costs



Note: Figures on the top of each bar give the corresponding average total calorie content of baskets.

3. Conceptual framework

We present a simple demand model, adopted from Dubois et al. (2014), to link the characteristics of food baskets to their cost.[4] Consider that a household i with demographics η_i chooses from N food products, where product n is characterized by C characteristics $\{a_{n1}, \dots, a_{nC}\}$. The household decides the quantity of product n, denoted by y_{in}, and of the outside good x_i to purchase. $\boldsymbol{z}_i$ represents a $C\,x\,1$ vector of characteristics of foods purchased by household i and hence z_{ic}= $\sum_{n=1}^{N} y_{in} a_{nc}$. Let $\gamma_i$be the marginal utility of income, the utility maximisation problem of household i is expressed as

$$\max_{x_i, y_i} U\,(x_i, \boldsymbol{z}_i, \boldsymbol{y}_i; \eta_i) = \prod_{n=1}^{N} (f_{in}(y_{in}))^{u_{in}} \prod_{c=1}^{C} h_{ic}(z_{ic})\ \exp(\gamma_i x_i) \tag{1}$$

where $f_{in}(y_{in})$ and $h_{ic}(z_{ic})$ are individual specific functions that determine utility from foods and utility from food characteristics respectively. Let p_o and p_n be the price of the outside good x_i and of one unit of y_{in}. With an income of I_i, the budget constraint of household i is given as

$$\sum_{n=1}^{N} y_{in} p_n + p_o x_i \leq I_i; \quad x_i, y_{in} \geq 0, \tag{2}$$

The first order condition can be obtained by maximising utility subject to budget constraints for a given n

$$\mu_{in} \frac{f'_{in}(y_{in}) y_{in}}{f_{in}(y_{in})} + \sum_{c} a_{nc} y_{in} \frac{h'_{ic}(z_{ic})}{\sum_l h_{ic}(z_{ic})} = \gamma_i \frac{p_n}{p_0} y_{in} \tag{3}$$

where $h_{ic}(z_{ic})$ is a Cobb-Douglas utility function ie. $h_{ic}(z_{ic}) = z_{ic}^{\beta_c}$ and $f_{in}(y_{in})$ takes the functional form of constant elasticity of substitution, $f_{in}(y_{in}) = \lambda_{in} y_{in}^{\theta_{in}}$. Rearranging the terms and simplifying yields the following expression:

$$p_n y_{in} = p_0 \frac{\mu_{in}\theta_{in}}{\gamma_i} + \sum_{c} p_0 \frac{\beta_c}{\gamma_i} a_{nc} y_{in} \tag{4}$$

Given our focus on the healthiness of the entire food basket, we sum the above expression over n, i.e. all food products within the basket purchased by household i:

$$\sum_{n} p_n y_{in} = p_0 \frac{\sum_n \mu_{in}\theta_{in}}{\gamma_i} + \sum_{c} p_0 \frac{\beta_c}{\gamma_i} \sum_{n} a_{nc}\, y_{in} \tag{5}$$

Let W_i= $\sum_n p_n y_{in}$, the cost of food basket purchased by household i, and $z_{ic} = \sum_n \alpha_{cn} y_{in}$ the characteristics of the corresponding food basket. This yields:

[4] Dubois et al. (2014) study whether differences in food purchases across countries can be attributed to differences in prices and nutritional characteristics. They develop a demand model for food products and nutrients, which nests commonly used models in characteristics space (e.g. discrete choice model and hedonic price model). In this paper, we adopt their model to study to the extent to which differences in the nutritional quality of food baskets can explain differences in basket costs.

$$W_i = p_0 \frac{\mu_i \theta_i}{\gamma_i} + \sum_c p_0 \frac{\beta_c}{\gamma_i} z_{ic} \quad (6)$$

4. Estimation approach

We assess how HFSS energy share is associated with basket costs by leveraging on within-household variations in food baskets over time. To derive the estimating equation, we introduce a subscript t to indicate that the baskets purchased vary over time and a subscript r to allow price heterogeneity across postcode areas[5]. Following Dubois et al (2014), we assume one of the food characteristics ($c = 1$) to be unobserved and normalise $p_0 = 1$ and $\gamma_i = 1$. The unobserved heterogeneity of household preferences and food environment is modelled as a combined error term by letting $p_0 \frac{\mu_i \theta_i}{\gamma_i} + p_0 \frac{\beta_1}{\gamma_i} z_{i1} = \delta_i + \tau_{rt} + \varepsilon_{it}$. The cost of basket ($W_{it}$), is measured by per person expenditure on all foods and drinks purchased by household i in month t, to account the difference in household size. Our main estimation equation can thus be derived as follows:

$$W_{it} = \sum_c \beta_c z_{ict} + \delta_i + \tau_{rt} + \varepsilon_{it} \quad (7)$$

Where z_{ict} indicates the healthiness of the food baskets, which is measured as the HFSS energy share of the basket (HES_{it}). Given the non-linearity observed in Figure 1, we also include the quadratic term of HES_{it} to examine the linkage between the healthiness and the cost of monthly food baskets.

Household fixed effects (δ_i) capture the heterogeneity in food preferences that are driven by differences in time-invariant household characteristics. This includes not only demographic differences like income, education, age, cultural background and occupation, but also differences in underlying preferences that are often not observed by researchers, such as eating out patterns or adherence to specific diets such as vegetarian or vegan etc. Month-postcode area fixed effects (τ_{rt}) are included as food availability is likely to vary across time and areas. Finally, the effects of unobserved time-varying household characteristics and preference shocks are captured by ε_{ip}. This set up allows us to examine how within-household changes in the healthiness of food baskets affect the per person basket cost.

Threats to identification

Through the household and postcode-month fixed effects, our estimation approach enables us to mitigate the bias from household heterogeneity that are time-invariant within a year and the bias from time-varying food characteristics that are fixed within areas. However, it remains possible that the error term is correlated with our key variable of interest if the household-specific preference shocks affect food choices. As large dietary changes can be rare, this potential endogeneity bias might be rather small and to be less of a concern. Study from the US shows that household diets appear to be on average unresponsive to illness or other household circumstances like childbirth, marriage, divorce, job loss, retirement and income changes (Hut and Oster, 2022).

To address this possible issue, we additionally employ an instrument variable approach which exploits the variation of food characteristics offered in local areas due to exogeneous reasons. Specifically, we generate a

[5] Dubois et al. (2014) uses NUTS3, which are upper-tier authorities or groups of lower-tier authorities in the UK. This information is not available in our dataset, which collects the postcode areas of where households reside instead.

measure of the healthiness of food environment faced by each household, based on the food baskets purchased by other households who live in the same local area and shop in the same retailer chain r in the same month, as household i. These households are referred as the reference group $h(i)$. Our instrument variable, $\omega_{h(i)t}$, is given as follows:

$$\omega_{h(i)t} = \sum_{r} \frac{F_{irt}}{D_{it}} \overline{HES}_{h(i)rt} \quad (8)$$

where $\overline{HES}_{h(i)rt}$ denotes the average HFSS energy share of baskets purchased by reference group $h(i)$ from retailer chain r in month t. This can be thought as a measure of the healthiness of food available from the retailer within the local area. As households often purchased foods from more than one retailer chain, we weigh $\overline{Hfss}_{h(i)rt}$ by the relative frequency of household i shopped in each retailer r, which is the number of days household i shopped in each retailer chain r in month t, F_{irt}, divided by the total number of days they shopped for foods in that month, D_{it}.[6] The quadratic term of $\omega_{h(i)t}$ will also be used to ensure that the number of instruments is no less than the number of endogenous variables. As we show below, our instruments are highly correlated with HES_{it} and its quadratic term. The identifying assumption is that the variation in $\omega_{h(i)t}$ is independent from the error term conditional on the household and month-postcode area fixed effects.

Another potential threat to identification is reverse causality rising from the concern that food choices are limited by the financial ability of households. They may purchase less healthy foods in order to keep the cost of food baskets within their budget. The inclusion of household fixed effect in our estimation accounts for the income and wealth of the households, the main drivers of their financial constraints. Furthermore, Cherchye et al. (2020) find that only a small portion of within-individual variation in food choices made by British individuals from 2005-2011 is explained by prices and budgets along with advertising and weather. There could remain a concern over reserve causality if households face time-varying financial needs and foods take up a substantial share of their expenditure. This is however unlikely to be the case in the UK as foods and non-alcoholic drinks only took up less than 14% of the total household spending for British households on average in 2019 (Office for National Statistics, 2020).

5. Results

Table 3 presents the OLS estimates of our main specification (7). Column 1 shows the results of the model with only linear term of HES_{it} and household fixed effect. The HFSS energy share is shown to be positively and significantly associated with the cost of food basket. This implies that the greater share of unhealthy foods in the baskets, the more the households spend on it. We consider the non-linear association between the healthiness and cost of food baskets in column 2. The coefficient of $HES_{it} \times HES_{it}$ is found to be negative and significant. Together with the positive coefficient of HES_{it}, they indicate that the cost of food baskets increases with the HFSS energy share at a decreasing rate. This relationship continues to hold when seasonality is controlled via month fixed effects in column 3. Column 4 shows the estimates from our preferred specification, consisting of both household and month-postcode area fixed effects. The

[6] Our instrument follows the same identification approach as the one used in Dubois et al. (2014) by approximating the quality of foods available to each household based on the purchases made by households in a reference group. For each food category, Dubois et al. (2014) define the reference group as other households in the local area who do their shopping in the retail chain where households shop most frequently. In our dataset, households often purchased their foods from several retailers at the same frequency in a month. We therefore use a weighted average HFSS energy share of food bought by the reference groups from different retailers to better capture the household-specific food environment.

coefficients of HES_{it} and $HES_{it} \times HES_{it}$ remain statistically significant and consistent with the concave relationship between the cost and healthiness of baskets found in the previous models.

These results illustrate that whether buying more healthy foods increases the cost of diets depends on the composition of the current food basket. In columns 2-4, we also estimate the level of HES^* at which the basket cost is perceived to be most expensive (i.e. the local maximum of the linkage between cost of food baskets and HES energy share). Let β_1 and β_2 be the coefficient of HES_{it} and $HES_{it} \times HES_{it}$ respectively then $HES^* = -\frac{\beta_1}{2\beta_2}$. Based on our preferred specification (column 4), a basket with 52.8% of calorie from HFSS foods is associated with the highest cost. In other words, buying relatively more unhealthy foods will only lower the basket cost if the current basket already contains predominately unhealthy foods (i.e. HES >52.8%).

Table 1 OLS estimates

	(1)	(2)	(3)	(4)
HES_{it}	29.169***	192.357***	187.273***	186.833***
	(0.898)	(3.989)	(3.936)	(3.892)
$HES_{it} \times HES_{it}$		-170.320***	-177.289***	-176.895***
		(3.584)	(3.706)	(3.659)
Household fixed effect	Yes	Yes	Yes	Yes
Month fixed effect	No	No	Yes	No
Month-postcode area fixed effect	No	No	No	Yes
Observations	328,249	328,249	328,249	328,249
HES^* (local maximum)	NA	0.565	0.528	0.528
		(0.003)	(0.002)	(0.002)

Note: HES^*: HFSS energy share of monthly food baskets. Standard error clustered at postcode areas in parentheses *** p<0.01, ** p<0.05, * p<0.1 HES^*:

Instrumental variable estimates

Considering the potential endogeneity in our measure of the healthiness of baskets, we employ IV regression below to assess the robustness of the concave relationship identified. Table 2 presents the IV estimates using the household-specific food environment measure (i.e. $\omega_{h(i)p}$) as the instrument. Column 1 controls for household heterogeneity while column 2 further controls for seasonality using month fixed effects. Column 3 presents the IV results of our preferred model which include both household and month-postcode area fixed effects. We first check the relevance of our instrument to the potentially endogenous regressor through the first stage results reported in panels A and B. As expected, they are positively associated to the HFSS energy share, which indicates that the more unhealthy the food environment the higher the share of calories from HFSS foods in the households' food baskets. The F-statistics presented suggest that both instruments are reasonably strong individually as they satisfy Staiger and Stock (1997)'s rule-of-thumb (i.e., F > 10). They are also above 104.7, the threshold under which standard errors would need to be corrected (Lee et al., 2022).

Next, we examine the second stage estimates in panel C, which continue to show a concave relationship between the cost and healthiness of food baskets. The positive coefficient of HES_{it} indicates that the cost of food basket increases with its HFSS energy share while the negative coefficient of $HES_{it} \times HES_{it}$ shows that the rate of such increase decreases as the basket contains more energy from HFSS foods. We use the Cragg-Donald Wald F statistic and the Kleibergen-Paap Wald rk F statistic to examine whether equation 7 is weakly identified by both instruments jointly. When there are two endogenous variables and two instrumental variables, the critical value for the instruments to have sufficient strength to eliminate at least 90% of the bias in the OLS regressions is 7.03 (Stock and Yogo, 2005). Both test statistics are much larger than 7.03 across all models, suggesting that our IV analysis is unlikely to suffer from weak instrument problem. Additionally, the p-values of Kleibergen-Paap *rk* LM statistic for underidentification are smaller than 0.01, rejecting the null hypothesis of no correlation between the endogenous variables and the instrumental variables at a 1% significance level. These test statistics confirm the strength of our instruments.

Table 2 IV estimates

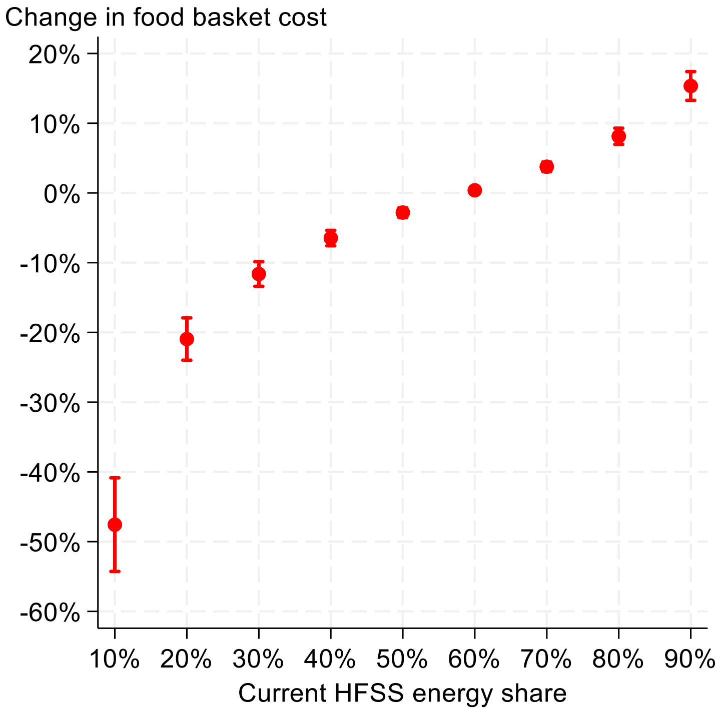
Specification	(1)	(2)	(3)
Panel A: First stage results (Dependent variable: HES_{it})			
$\omega_{h(i)t}$	-0.418***	-0.915***	-1.072***
	(0.113)	(0.115)	(0.101)
$\omega_{h(i)t} \times \omega_{h(i)t}$	0.835***	1.115***	1.246***
	(0.110)	(0.114)	(0.101)
F-statistics	855.83	295.51	182.46
Panel B: First stage results (Dependent variable: $HES_{it} \times HES_{it}$)			
$\omega_{h(i)t}$	-1.197***	-1.704***	-1.883***
	(0.123)	(0.133)	(0.125)
$\omega_{h(i)t} \times \omega_{h(i)t}$	1.609***	1.900***	2.052***
	(0.123)	(0.134)	(0.126)
F-statistics	792.91	262.73	167.64
Panel C. Second stage results (Dependent variable: Cost of food baskets)			
HES_{it}	897.885***	885.715***	908.039***
	(112.332)	(80.371)	(165.386)
$HES_{it} \times HES_{it}$	-770.552***	-809.360***	-789.894***
	(92.150)	(69.088)	(129.408)
Cragg-Donald Wald F statistic	1225.22	388.39	269.25
Kleibergen-Paap Wald rk F statistic	165.36	138.67	90.95
Kleibergen-Paap rk LM statistic	<0.001	<0.001	<0.001
Household fixed effect	Yes	Yes	Yes
Month fixed effect	No	Yes	No
Month-postcode area fixed effect	No	No	Yes
Observations	328,249	328,249	328,249
HES^{*} (local maximum)	0.657	0.557	0.563
	(0.009)	(0.008)	(0.008)

Note: HES_{it}: HFSS energy share of monthly food baskets. Standard error clustered at postcode areas in parentheses *** p<0.01, ** p<0.05, * p<0.1

Comparing to the OLS estimates in table 1, the coefficients of HES_{it} and HES_{it}*HES_{it} under the IV analysis are larger, depicting a possible download bias of OLS estimates. Despite these differences in the magnitude of coefficients, the HES^{*} estimated under the IV and OLS analyse are very close. Based on our preferred specification, the IV estimate of HES is 56.3% (Column 3 in table 2) while the corresponding OLS estimate is 52.8% (Column 4 in table 1).

To give more context to the findings, we estimate the percentage change in the cost of baskets when the HFSS energy share is reduced by 5%. Figure 3 shows the cost changes using IV estimates from our preferred specification across a range of value of HFSS energy share (from 10% to 90%). Conditional on household and month-postcode area fixed effects, if the current food basket contains 80% calories from HFSS foods, a 5% reduction in HFSS energy share will increase the basket cost by over 10%. In contrast the same percentage decrease in HFSS energy share will lower the basket cost by around 20% if the current basket contains only 20% calories from HFSS foods.

Figure 3. Predicted percentage changes in food basket costs with a 5% reduction in HFSS energy share



Note: The changes are estimated using the IV estimates in table 2 column 3.

Heterogeneity

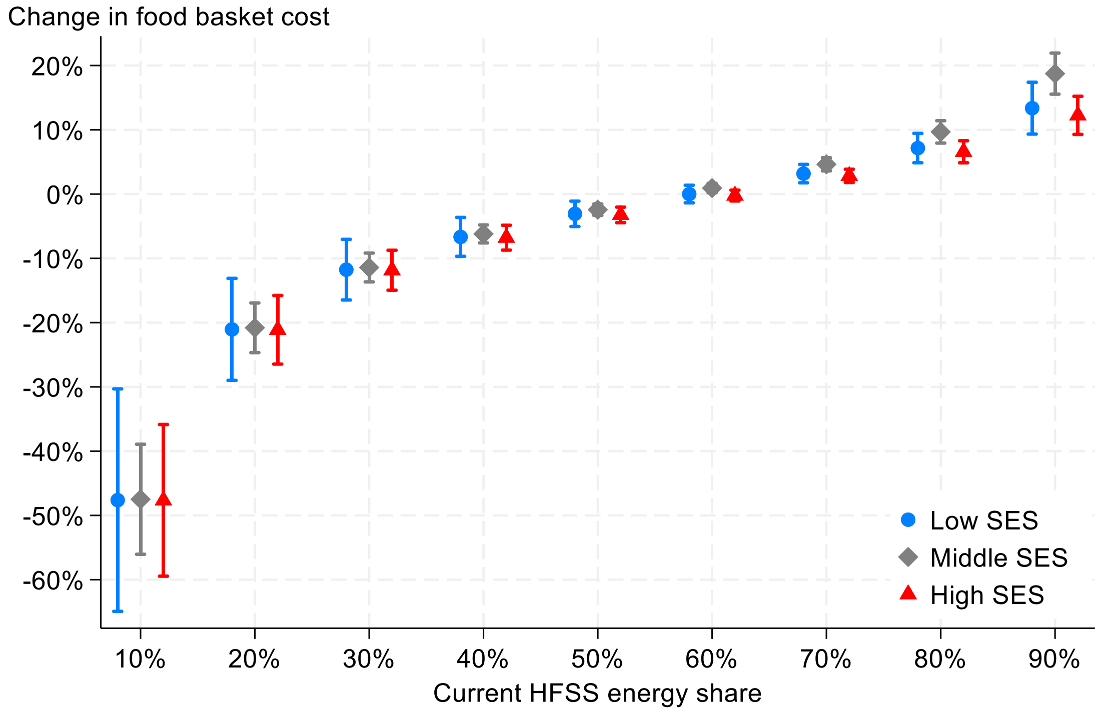
In this section, we explore whether the concave linkage between food basket costs and HFSS energy share is heterogeneous in subsamples by social-demographic characteristics of households.[7] In the UK, there has been rising policy interest in social inequality in diets. The National Food Strategy has stressed the urgency to help low income families eat well as they are more likely to suffer from diet-related conditions (Dimbleby, 2021). We first repeat the IV estimation of equation 7 separately for households of different occupational SES. Figure 4 shows the predicted percentage changes in food basket costs associated with a 5% reduction in HFSS energy share across the three SES groups, which depict a similar pattern as our main finding in Figure 3. The minimal difference in predicted percentage cost changes suggests that low SES households with unhealthy diets face similar relative cost implications as those with higher SES when trying to improve their diet. Since low SES households are often at financial disadvantage, such cost barrier could be more detrimental to their diet quality.

Childhood obesity has been one of the key public health concerns in the UK. There has been reports that many children are too poor to meet the healthy food guidelines of the government (Scott et al., 2018). In the light of this concern, we examine if the linkage between the healthiness and cost of food baskets is similar among households with and without children. Figure 5 shows minimal differences of the predicted basket cost changes with reduced HFSS energy share between households with and without children.

[7] We report the sample mean of per person monthly food basket costs and HFSS energy share across subsamples in Appendix table A3 and the estimation results in Appendix tables A4 and A5.

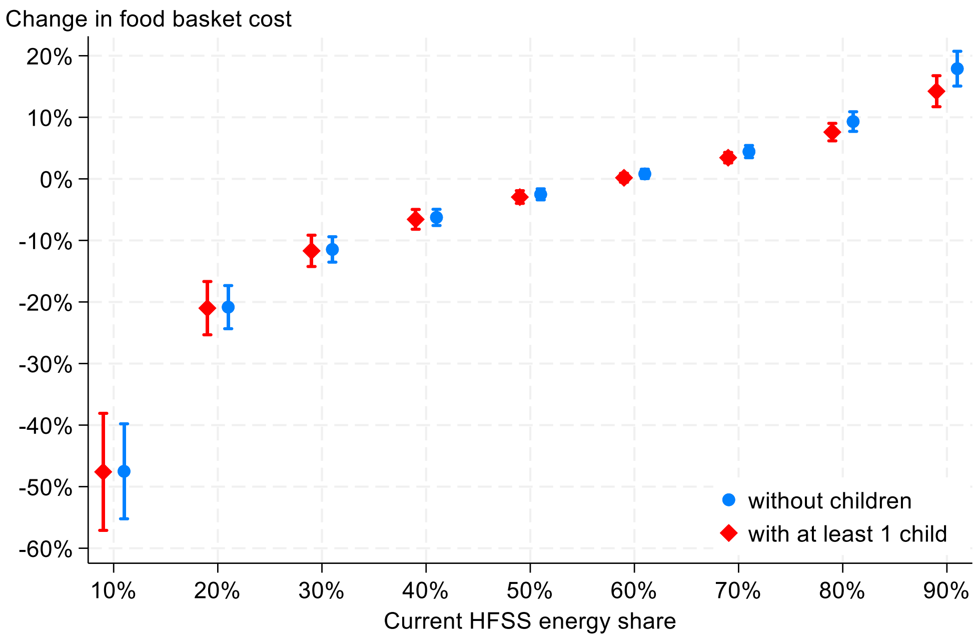
Lastly, we explore if the linkage is heterogenous to the health status of the main shopper in the households. To do so, we classify the households into two groups based on whether the main shopper in the household is overweight (i.e. with body mass index (BMI) equal or more than 25) or otherwise. BMI is defined as weight in kilograms divided by height in meters squared. Around 81% of our household samples reported their BMI (N=25,813). Figure 6 shows the predicted changes in food basket costs under a 5% reduction in HFSS energy share are similar regardless of the BMI of the household's main shopper.

Figure 4. Predicted percentage change in food basket costs with a 5% reduction in HFSS energy share by SES



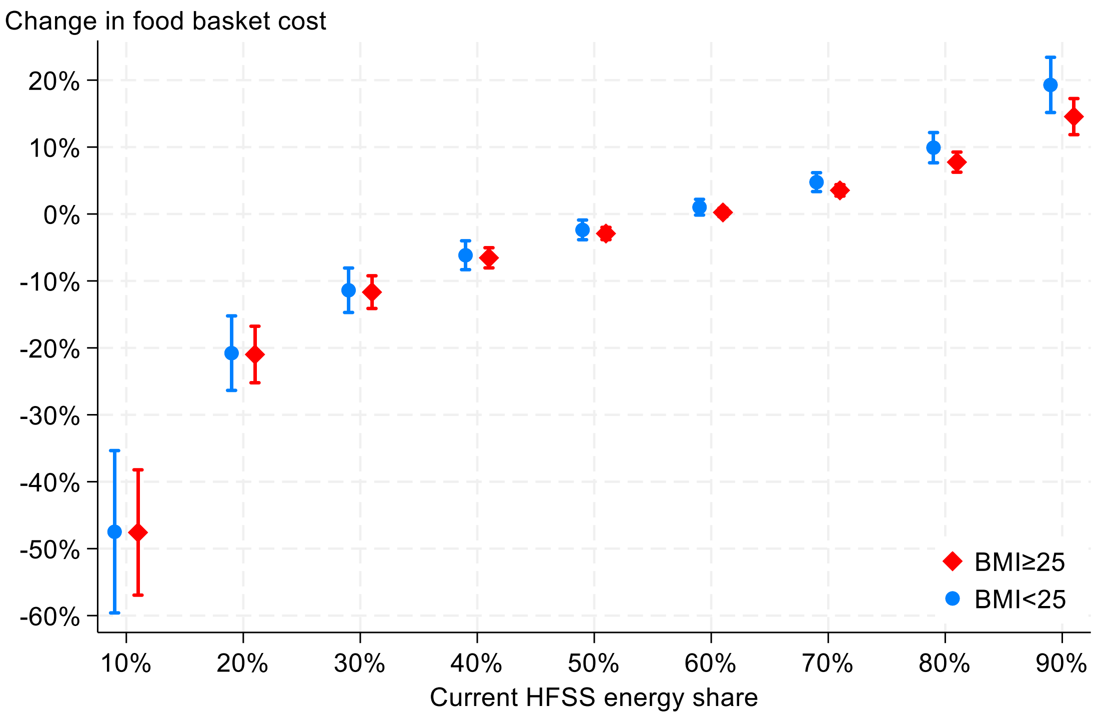
Note: The changes are estimated using the IV estimates in Appendix Table A4.

Figure 5. Predicted percentage change in food basket costs with a 5% reduction in HFSS energy share among households with and without children



Note: The changes are estimated using the IV estimates in Appendix Table A5.

Figure 6. Predicted percentage change in food basket costs with a 5% reduction in HFSS energy share by the BMI of main food shoppers



Note: The changes are estimated using the IV estimates in Appendix Table A5.

6. Discussion and policy implications

This paper uses a large representative British household purchase panel of foods and beverages bought to consume at home in 2017 to explore the association between the healthiness and cost of diet. Its objective is to understand how within-household variations in the energy share from HFSS foods drive changes in food basket costs. We first document the large variations in the healthiness of food baskets of similar costs, challenging the commonly held view that healthy diets are more expensive than unhealthy diets. We then demonstrate that the relationship in fact is non-linear, and the healthiest diets can be as affordable as the least healthy diets. This linkage is consistently found among households across SES, with and without children as well as those whose main shopper is overweight or otherwise. It implies that buying a basket consisting predominantly of either non-HFSS energy or HFSS energy is likely to be less expensive than a mixed basket. Based on our IV results, substitution of HFSS energy with non-HFSS-energy initially increases the cost of the basket but once the share of energy from HFSS foods decreases to 56.3% then further substitution of HFSS energy with non-HFSS energy will start lowering the cost of the basket. In other words, if the current food basket is rather unhealthy, small substitutions towards non-HFSS energy are likely to increase the cost and only a substantial change would lead to an equivalent costing healthy basket. Making such big changes are challenging (Hut and Oster (2022)). Hence, while food cost does not fully explain unhealthy diet as there are cost-equivalent baskets with much lower HFSS energy share, it is likely to be a detriment to improving diets in small steps.

There are a few caveats in this paper. Despite the richness of our data in capturing take home purchase, it does not include food and drink purchases made in out-of-home settings, such as restaurants and fast-food chains. In our analysis, eating-out patterns are controlled via households and month-postcode fixed effects. The former captures household-specific preference in consuming foods outside of home and the latter account for the seasonal changes in eating out patterns. Our IV instrument further ensure that our results are not biased by the unobserved time-varying element of eating out patterns. Second, it is important to note that the HFSS energy share, our main variable of interest in this study, is only one of the dimensions in which diet quality can be measured. Some nutrition research has argued that people get full by the amount of foods rather than the number of calories contained in foods and hence diets with lower energy density (i.e. calorie content per gram of food) are more preferable. Since our data does not capture precise information on the weight of food items, we are unable to repeat our analysis using alternative measures such as energy density of the basket, which would be a fruitful avenue for future research when appropriate data are available.

Our results provide several insights on the effectiveness of fiscal measures in encouraging people to make healthier food choices. As discussed above, if the current baskets contain predominantly calories from HFSS food, purchasing relatively more these foods will reduce the basket cost. In other word, there is a cost saving incentive for households with poor diets to consume relatively more HFSS food, thus exacerbating their poor diet. This provides evidence in support of implementing taxes targeted at HFSS food to adjust the price signal and hence reduce the cost-saving incentive from increasing calorie share from HFSS food. In recent years, there has been a rise in adoption of sugar-sweetened beverage (SSB) taxation across the world. Current empirical research generally finds these taxes to be effective in reducing SSB consumption.[8] In the UK, the Soft Drinks Industry Levy, a tiered tax on sugar-sweetened beverages, implemented in 2018 is found to have led to reduction in overall calorie as well as sugar purchases from SSBs (Dickson et al., 2023; Rogers et al., 2023). However, non-alcoholic drinks contribute to less than 10% of the total calories in the food baskets of British households, as reflected in Figure 2. It is thus important to consider extending health-related food taxes to other HFSS foods, such as sweet and savoury snacks which take up at least 20% of total calories in the food baskets with HFSS energy share over 60%. On the other hand, subsiding healthier food like fruit and vegetables can lower the basket cost increment and hence incentivise households with poor diets to make

[8] There is a large literature summarising the current empirical evidence on the effects of SSB taxes, see Cawley and Frisvold (2023) and (Kiesel et al., (2023) and for example.

small steps in their substitution away from HFSS food. The heterogeneity analysis indicates that households across socio-economics groups face a similar concave linkage between food basket cost and its healthiness. Financial aid is particularly helpful to improve diets among low SES households as they are often at financial disadvantage and hence less capable to afford more expensive but healthier food baskets.

While this paper provides support for the use of fiscal measures in improving diets, it also highlights their limitations in preventing obesity and improving population health. Using the IV estimates from our full model, we find that if the current food basket contains 80% of calories from HFSS food a 5% reduction in HFSS energy share will increase the basket cost by over 10%. This implies that sufficiently high taxes or subsidies are needed to achieve meaningful changes in the healthiness of household food purchases. This is consistent with the price inelastic demand for food among British households. Cornelsen, et al. (2019) study the food purchases made by British households in 2012-2013 and find that the own price elasticities for various food groups are generally less than one. However, such high taxes on unhealthy food might backfire as the current food basket of people with poor diets would become much more expensive. If they were not prepared or able to make substantial dietary changes, they might look to purchase food baskets with even higher HFSS energy share as long as the linkage between cost and healthiness of food baskets continues to be concave.

Our analysis shows that if households are willing and able to make dramatic changes in their diets it is possible to have a much healthier basket with the same cost. Hence, while food costs are a key factor driving food choices, they do not fully explain the healthiness of food baskets purchased by households. This is in line with Dubois et al. (2014) which highlights the importance of preferences and eating habits, in addition to price and attributes of food products, in explaining observed differences in nutritional composition across countries. Empirical evidence from the US also suggests that difference in neighbourhood food environment (i.e. availability and price) is unlikely to contribute to meaningful differences in nutritional quality between high- and low-income households (Allcott et al 2019). Hastings et al (2021) argue that closing the gap in food expenditure between high- and low- SES households in the US would not fully address their differences in summary measures of food healthfulness. Our findings, along with the conclusions from other studies, point to the need of complementary policies to address the non-financial barriers to healthier food choices. Fiscal measures would be more effective in shifting food purchases away from HFSS food if households also developed higher preferences for healthier foods and were equipped with the knowledge, facilities, and skills to improve their diet.

Reference

Allcott, H., Diamond, R., Dubé, J. P., Handbury, J., Rahkovsky, I., & Schnell, M. (2019). Food Deserts and the Causes of Nutritional Inequality. *The Quarterly Journal of Economics*, *134*(4), 1793–1844. https://doi.org/10.1093/QJE/QJZ015

Batis, C., Marrón-Ponce, J. A., Stern, D., Vandevijvere, S., Barquera, S., & Rivera, J. A. (2021). Adoption of healthy and sustainable diets in Mexico does not imply higher expenditure on food. *Nature Food 2021 2:10*, *2*(10), 792–801. https://doi.org/10.1038/s43016-021-00359-w

Bechthold, A., Boeing, H., Schwedhelm, C., Hoffmann, G., Knüppel, S., Iqbal, K., De Henauw, S., Michels, N., Devleesschauwer, B., Schlesinger, S., & Schwingshackl, L. (2017). Food groups and risk of coronary heart disease, stroke and heart failure: A systematic review and dose-response meta-analysis of prospective studies. *Https://Doi.Org/10.1080/10408398.2017.1392288*, *59*(7), 1071–1090. https://doi.org/10.1080/10408398.2017.1392288

Bell, M., Woolley, N., Toms, H., & Lebre de Freitas, G. (2023). *The rising cost of obesity in the UK*. https://www.frontier-economics.com/uk/en/news-and-insights/news/news-article/?nodeId=20358

Berger, N., Cummins, S., Smith, R. D., & Cornelsen, L. (2019). Recent trends in energy and nutrient content of take-home food and beverage purchases in Great Britain: an analysis of 225 million food and beverage purchases over 6 years. *BMJ Nutrition, Prevention & Health*, *2*(2), 63. https://doi.org/10.1136/BMJNPH-2019-000036

Carlson, A., & Frazão, E. (2012). Are Healthy Foods Really More Expensive? It Depends on How You Measure the Price. *USDA-ERS Economic Information Bulletin No.96*, *96*, 1–58. https://doi.org/10.2139/SSRN.2199553

Cawley, J., & Frisvold, D. (2023). *Review: Taxes on sugar-sweetened beverages: Political economy, and effects on prices, purchases, and consumption*. https://doi.org/10.1016/j.foodpol.2023.102441

Cherchye, L., De Rock, B., Griffith, R., O’connell, M., Smith, K., & Vermeulen, F. (2020). A new year, a new you? Within-individual variation in food purchases. *European Economic Review*, *127*, 103478. https://doi.org/10.1016/j.euroecorev.2020.103478

Cornelsen, L., Berger, N., Cummins, S., & Smith, R. D. (2019). Socio-economic patterning of expenditures on ‘out-of-home’ food and non-alcoholic beverages by product and place of purchase in Britain. *Social Science and Medicine*, *235*, 112361. https://doi.org/10.1016/j.socscimed.2019.112361

Cornelsen, L., Mazzocchi, M., & Smith, R. D. (2019). Fat tax or thin subsidy? How price increases and decreases affect the energy and nutrient content of food and beverage purchases in Great Britain. *Social Science and Medicine*, *230*, 318–327. https://doi.org/10.1016/j.socscimed.2019.04.003

Department of Health. (2011). *Nutrient Profiling Technical Guidance*.

Dickson, A., Gehrsitz, M., & Kemp, J. (2023). Does a Spoonful of Sugar Levy Help the Calories Go Down? An Analysis of the UK Soft Drinks Industry Levy. *The Review of Economics and Statistics*, 1–29. https://doi.org/10.1162/REST_A_01345

Dimbleby, H. (2021). *National Food Strategy - The Report*.

Dubois, P., Griffith, R., & Nevo, A. (2014). Do Prices and Attributes Explain International Differences in Food Purchases? *American Economic Review*, *104*(3), 832–867. https://doi.org/10.1257/AER.104.3.832

Grosso, G., Bella, F., Godos, J., Sciacca, S., Del Rio, D., Ray, S., Galvano, F., & Giovannucci, E. L. (2017). Possible role of diet in cancer: systematic review and multiple meta-analyses of dietary patterns, lifestyle factors, and cancer risk. *Nutrition Reviews*, *75*(6), 405–419. https://doi.org/10.1093/NUTRIT/NUX012

Hastings, J., Kessler, R., & Shapiro, J. M. (2021). The Effect of SNAP on the Composition of Purchased Foods: Evidence and Implications. *American Economic Journal: Economic Policy*, *13*(3), 277–315. https://doi.org/10.1257/POL.20190350

Hut, S., & Oster, E. (2022). Changes in household diet: Determinants and predictability. *Journal of Public Economics*, *208*, 104620. https://doi.org/10.1016/J.JPUBECO.2022.104620

Jones, A., Dunford, E., Crossley, R., Thout, S., Rayner, M., Neal, B., Jones, A., Dunford, E., Crossley, R., Thout, S. R., Rayner, M., & Neal, B. (2017). An Evaluation of the Healthiness of the Indian Packaged Food and Beverage Supply. *Nutrients*, *9*(10), 1103. https://doi.org/10.3390/nu9101103

Kiesel, K., Lang, H., & Sexton, R. J. (2023). A New Wave of Sugar-Sweetened Beverage Taxes: Are They Meeting Policy Goals and Can We Do Better? *Annual Review of Resource Economics*, *15*, 407–432. https://doi.org/10.1146/ANNUREV-RESOURCE-111522-111325

Lee, A. J., Kane, S., Herron, L. M., Matsuyama, M., & Lewis, M. (2020). A tale of two cities: The cost, price-differential and affordability of current and healthy diets in Sydney and Canberra, Australia. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 1–13. https://doi.org/10.1186/S12966-020-00981-0/TABLES/4

Lee, D. S., McCrary, J., Moreira, M. J., & Porter, J. (2022). Valid T-Ratio Inference for IV. *American Economic Review*, *112*(10), 3260–3290. https://doi.org/10.1257/AER.20211063

Malik, V. S., Willett, W. C., & Hu, F. B. (2012). Global obesity: trends, risk factors and policy implications. *Nature Reviews Endocrinology 2012 9:1*, *9*(1), 13–27. https://doi.org/10.1038/nrendo.2012.199

Office for National Statistics. (2020, July). *Family spending in the UK: April 2018 to March 2019*. https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/coronavirusandhomeworkingintheuk/april2020

Rao, M., Afshin, A., Singh, G., & Mozaffarian, D. (2013). Do healthier foods and diet patterns cost more than less healthy options? A systematic review and meta-analysis. *BMJ Open*, *3*(12), e004277. https://doi.org/10.1136/BMJOPEN-2013-004277

Rogers, N. T., Pell, D., Mytton, O. T., Penney, T. L., Briggs, A., Cummins, S., Jones, C., Rayner, M., Rutter, H., Scarborough, P., Sharp, S., Smith, R., White, M., & Adams, J. (2023). Changes in soft drinks purchased by British households associated with the UK soft drinks industry levy: a controlled interrupted time series analysis. *BMJ Open*, *13*, 77059. https://doi.org/10.1136/bmjopen-2023-077059

Rosenheck, R. (2008). Fast food consumption and increased caloric intake: a systematic review of a trajectory towards weight gain and obesity risk. *Obesity Reviews*, *9*(6), 535–547. https://doi.org/10.1111/J.1467-789X.2008.00477.X

Scott, C., Sutherland, J., & Taylor, A. (2018). *Affordability of the UK's Eatwell Guide*.

Stock, J., & Yogo, M. (2005). Asymptotic distributions of instrumental variables statistics with many instruments. *Stock, J., & Yogo, M. (2005). Asymptotic DistributioIdentification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, *6*, 109–120.

Vandevijvere, S., Seck, M., Pedroni, C., De Ridder, K., & Castetbon, K. (2020). Food cost and adherence to guidelines for healthy diets: evidence from Belgium. *European Journal of Clinical Nutrition 2020 75:7*, *75*(7), 1142–1151. https://doi.org/10.1038/s41430-020-00815-z

Appendix

Table A1 Summary statistics

	Mean	SD
Household demographics		
Presence of children (Yes/ No)	0.620	0.967
Household Size	2.713	1.329
Age of main shopper	50.657	15.047
Regions		
London	0.167	0.373
Midlands	0.150	0.357
North East	0.048	0.214
Yorkshire	0.123	0.328
Lancashire	0.106	0.308
South	0.105	0.307
Scotland	0.093	0.290
E. England	0.087	0.282
Wales + West	0.087	0.281
South West	0.034	0.182
Occupational social grades		
Low SES	0.213	0.409
Middle-SES	0.569	0.495
High- SES	0.218	0.413
Monthly Food Purchases		
Per person monthly food basket cost	82.972	57.580
HFSS energy share (HES)	0.501	0.130
Food Availability (IV)	0.496	0.041
Number of household	31723	

Note: The occupational social grades are based on the occupation of the main food shopper of the household. Respondents with higher and intermediate managerial, administrative or professional occupations are classified as high-SES. Mid-SES are skilled manual workers and those with supervisory or clerical and junior managerial, administrative or professional occupations. Low-SES consists of respondents who are semi-or unskilled manual workers, state pensioners, causal or lowest grade workers and those unemployed with state benefits (Cornelsen, Berger, et al., 2019).

Table A2. List of food products within the 19 food groups

Food group	**Food products**
Alcohol	Alcohol
Bread and morning goods	Bread products, other morning goods
Breakfast cereals	Breakfast cereals
Fat and oils	Margarine and vegetable oils, butter and animal fats
Fruits	Fresh fruits, tinned fruits
Milk and Yogurt	Whole milk, yogurts, low fat milk
Other	Water, table salt, table sauces and condiments, meat substitutes, sugar, honey, syrup, semiproducts, slimming products
Other dairy products	Cheese, cream, dairy-based drinks, ice creams
Other non-alcoholic drinks	Diet drinks, pure fruit juice and smoothies, juice drinks, hot and powdered drinks
Processed meat/ fish	Processed meat, processed fish
Pulses, seeds and nuts	Pulses, seeds and nuts
Ready meals/ convenience food	Soup, ready meals, convenience food
Red meat	Red meat
Regular drinks	Regular drinks
Rice/Pasta/Flour	Pasta and rice, other grains and flour
Savoury snacks	Crisps and savoury snacks
Sweet snacks	Desserts and puddings, chocolate and confectionary, biscuits
Vegetables	Fresh vegetables, canned vegetables, fresh salad, potatoes
White meat/fish/egg	Eggs, fish, white meat

Table A3 Summary statistics by household subsamples

Household subsamples	Per person monthly food basket cost		HFSS energy share		No of
	Mean	SD	Mean	SD	households
	Panel A social economic status				
Low SES	78.840	48.644	0.483	0.098	6,754
Middle SES	80.003	50.612	0.499	0.097	18,038
High SES	77.275	49.134	0.521	0.102	6,928
	Panel B Presence of children				
With children	50.500	27.261	0.498	0.102	20,450
Without children	94.954	52.429	0.504	0.093	11,270
	Panel C Body Mass Index (BMI) of the main shopper in the household				
BMI >= 25	88.522	60.339	0.507	0.129	15,560
BMI <25	78.614	54.055	0.488	0.131	10,256

Table A4 IV estimates by socioeconomic status

	(1)	(2)	(3)
Household groups	High SES	Middle SES	Low SES
Panel A. First stage results (Dependent variable: HES_{it})			
$\omega_{h(i)t}$	-1.151***	-1.114***	-1.023***
	(0.260)	(0.159)	(0.247)
$\omega_{h(i)t} \times \omega_{h(i)t}$	1.353***	1.275***	1.194***
	(0.251)	(0.158)	(0.251)
F-statistics	76.65	160.16	35.46
Panel BC: First stage results (Dependent variable: $HES_{it} \times HES_{it}$)			
$\omega_{h(i)t}$	-2.138***	-1.857***	-1.851***
	(0.296)	(0.194)	(0.290)
$\omega_{h(i)t} \times \omega_{h(i)t}$	2.325***	2.009***	2.037***
	(0.290)	(0.193)	(0.299)
F-statistics	72.22	137.32	167.64
Panel C. Second stage results (Dependent variable: Cost of food baskets)			
HES_{it}	897.885***	885.715***	908.039***
	(112.332)	(80.371)	(165.386)
$HES_{it} \times HES_{it}$	-770.552***	-809.360***	-789.894***
	(92.150)	(69.088)	(129.408)
Cragg-Donald Wald F statistic	97.33	145.178	33.51
Kleibergen-Paap Wald rk F statistic	32.50	84.190	13.81
Kleibergen-Paap rk LM F statistic	<0.001	<0.001	<0.001
Household fixed effect	Yes	Yes	Yes
Month fixed effect	No	No	No
Month-postcode area fixed effect	Yes	Yes	Yes
Observations	69,267	187,901	328,249
HES^{*}	0.583	0.547	0.575
	(0.014)	(0.011)	(0.023)

Note: HES_{it}: HFSS energy share of monthly food baskets. Standard error clustered at postcode areas in parentheses *** p<0.01, ** p<0.05, * p<0.1 HES^{*}: the level of HFSS energy share associated with the highest basket cost.

Table A5 IV estimates by presence of children and BMI of the main food shopper

	(1)	(2)	(4)	(5)
Household groups	With children	Without children	BMI≤25	BMI>25
Panel A. First stage results (Dependent variable: HES_{it})				
$\omega_{h(i)t}$	-1.320***	-0.929***	-1.174***	-1.043***
	(0.165)	(0.162)	(0.204)	(0.178)
$\omega_{h(i)t} \times \omega_{h(i)t}$	1.514***	1.093***	1.351***	1.224***
	(0.163)	(0.159)	(0.204)	(0.175)
F-statistics	123.71	295.51	85.69	295.51
Panel B: First stage results (Dependent variable: $HES_{it} \times HES_{it}$)				
$\omega_{h(i)t}$	-2.233***	-1.693***	-2.078***	-1.859***
	(0.200)	(0.185)	(0.237)	(0.213)
$\omega_{h(i)t} \times \omega_{h(i)t}$	2.427***	1.849***	2.246***	2.034***
	(0.200)	(0.194)	(0.238)	(0.211)
F-statistics	132.50	114.51	90.55	126.18
Panel C. Second stage results (Dependent variable: Cost of food baskets)				
HES_{it}	621.327***	1,006.269***	696.595***	1,027.249***
	(62.382)	(82.051)	(87.422)	(102.076)
$HES_{it} \times HES_{it}$	-545.697***	-913.687***	-629.860***	-905.203***
	(50.510)	(69.669)	(71.013)	(85.446)
Kleibergen-Paap rk LM statistic	<0.001	<0.001	<0.001	<0.001
Cragg-Donald Wald F statistic	106.012	167.189	95.712	153.63
Kleibergen-Paap Wald rk F statistic	49.94	61.100	51.886	59.56
Household fixed effect	Yes	Yes	Yes	Yes
Month fixed effect	No	No	No	No
Month-postcode area fixed effect	Yes	Yes	Yes	Yes
Observations	110,398	217,847	107,202	163,518
HES^*	0.569	0.557	0.545	0.567
	(0.012)	(0.008)	(0.018)	(0.011)

Note: HES_{it}: HFSS energy share of monthly food baskets. Standard error clustered at postcode areas in parentheses *** p<0.01, ** p<0.05, * p<0.1 HES^*: the level of HFSS energy share associated with the highest basket cost.