# Health News Environment and the Distribution of Diet Quality

**Yizao Liu, Rebecca Cleary, and Andrea Carlson**

This paper investigates the relationship between the health news environment and the healthfulness of consumers' food purchases. We estimate an unconditional quantile regression model with the 2015-2018 household-level purchase data and find that more health information is associated with a better diet. The relationship is slightly stronger at the lowest quantile of the HEI distribution, although we cannot reject equivalence with the estimate at the mean for most quantiles. Further, the association between health information and diet quality is weaker in households with higher education, implying that high education households are less likely to be affected by media information.

*Key words:* Health Information, Diet Quality, Healthy Eating Index (HEI), Unconditional Quantile Regression, Purchase to Plate Crosswalk (PPC), IRI Consumer Network Data

## Introduction

Improving diet quality or healthfulness has long been a target of public health policy because of its direct impact on human health. Various factors, policies and interventions to improve diet quality have been discussed, including providing better access to a healthy food environment (Volpe, Okrent, and Leibtag, 2013; Courtemanche, et al., 2018; Allcott et al. 2019), nutrition and income assistance programs such as SNAP (Hastings, Kessler, and Shapiro, 2019; Katare, Binkley, and Chen, 2021), school food programs (Smith, 2017; Cleary, et al. 2021), nutrition labels (Christoph and An, 2018; Buyuktuncer et al., 2018), food reformulations (Ale-Chilet and Moshary, 2022), etc. One possible factor that might affect diet quality is media exposure to health and nutrition information, including media stories. In 2020, US adults spent an average of 13 hours per day with a combination of various media, such as internet, television (TV), radio, magazines and others (eMarketer, 2021). Media stories on television, radio, newspapers, especially online health information, has further resulted in more consumers actively searching for and acquiring health knowledge from online sources (Diviani, et al. 2015). Exposure to mass media may have a considerable impact on consumers' awareness of the importance of a healthy diet and food choices, by providing a very low-cost way for consumers to receive health information.

Lower information costs tend to increase the probability that heterogeneous consumers will choose healthier food products (Zhu, Lopez and Liu, 2016). Consumers may learn about the importance of a healthy diet and how to eat healthier over time from media health information (Smed, 2012), and make healthier food choices. However, today's consumers are also surrounded by an overwhelming amount of information about health and diet from different sources including healthcare providers, news coverage on new health findings, food advertisements, and word-of-mouth among consumers. Sometimes, the health information is confusing and conflicting, sending contradictory messages about health benefits of certain foods, eating behaviors or nutrition information (Im and Huh, 2017). Consequently, confused consumers may be less likely to comply with expert nutrition and health advice, even when there is no ambiguity that compliance will lower health risks (Nagler, 2014). Overall, the net impact of health information on consumers' diet quality is still uncertain and needs to be empirically examined.

Health information plays an important role in public health. Many healthcare professionals, organizations, and governments use health campaigns to raise awareness, promote healthy diets, encourage healthy behaviors, and curb harmful behaviors. For example, former First Lady Michelle Obama started a health promotion campaign, "*Let's Move*", in 2010 to provide families access to health education and fostering healthier environments to reduce the rates of childhood obesity. Both traditional and online media outlets provided extensive coverage on the First Lady's engagements during the campaign, reaching a much wider audience (Andersen, Wylie and Brank, 2017). In 2017, MTV ran a campaign on World AIDS Day to encourage young people to get tested for HIV, which received 1.2 million hits in five hours and became the number one trending topic in nine minutes (Butteriss and Bradley, 2019). While a successful campaign typically aims to expose a high proportion of a large population to health messages through media, the knowledge of the direction and magnitude of the impact of health information is of critical importance for policymakers when evaluating the true effectiveness of a health campaign.

This paper investigates the relationship between health information from the media and the healthfulness of consumers' retail food purchases. We measure the amount of health and diet related information available to consumers using the number of media stories on newspapers, magazines, TV, radio, and online sources covering healthy diets collected from the NexisUni database. The healthfulness of food purchases is measured by the 2015 Healthy Eating Index (HEI-2015), which is a validated index of dietary quality measuring adherence to key recommendations in the *Dietary Guidelines for Americans 2015-2020* (DGA-2015). We use 2015-2018 household-level retail food purchase data matched to USDA nutrition data (Carlson et al. 2019) to estimate the HEI-2015 based on food-at-home purchases for each household. We estimate a panel unconditional quantile regression model of household diet quality to investigate across the distribution of diet quality, as the tails of diet quality are clinically important. Those with the poorest diet quality are more at risk for diet-related chronic diseases such as hypertension, diabetes, and cardio-vascular disease, as well as more limited mobility than those with higher quality diets.

Recent literatures have shown that many consumers use media information as an important primary source of information on health-related decisions, including dietary choices (Greiner, Smith, and Guallar, 2010; Pieniak, Vanhonacker, and Verbeke, 2008) and health care choices (Beck et al., 2014; Suenaga and Vicente, 2022). A number of previous studies examine the effect of media information by focusing on public information campaigns, published scientific articles, or mass media stories on consumer demand in specific food categories, such as shell egg (Brown and Schrader, 1990), meat and fish (Tonsor, Mintert, and Schroeder, 2010), fruit and vegetables (Smed, 2012), and bottled water (Lu and Liu, 2017). Most of these studies find that information about the long-term health effects of dietary choices may have small effects on food demand. Information about food safety issues including high-risk, low-probability events such as food scares are in some cases found to have larger effects on consumption (e.g., Liu, Lien, and Asche 2016; Rieger, Kuhlgatz, and Anders 2016). Different from the previous literature, this paper focuses on the overall healthfulness of all retail food purchases and calculate the household's

monthly HEI score over time. In contrast to isolated food groups, the collective assessment of households' overall food purchases yields a reasonably accurate estimate of their general diet quality. This approach considers the role of health information more comprehensively, accounting for possible consumer substitution across various food groups.

A number of empirical studies found heterogeneity in consumer behavior in the search and use of health information, depending on consumer demographics and health status (Campos, Doxey, and Hammond, 2011), risk perceptions and risk attitudes (Yang and Goddard, 2011), or sources where they receive the information (Nam Verrill, and Kim, 2017). Other studies investigate the heterogeneous demand reaction to health news (Smed 2012; Rieger, Weible, and Anders 2016). Browning, Hansen and Smed (2018) examine the dynamic consumers responses and find that the short- and long-run information impacts vary across individual households. In line with the previous literature, this paper further examines the heterogenous impact of health news across households with different education levels.

We find that more health information in the media is associated with healthier food purchases. There is a slightly stronger link between health information and diet quality at the lowest quantiles of the HEI distribution, although across most quantiles we cannot reject equivalence with the estimate at the mean. For households with very unhealthy food purchases, providing more health information is associated with a larger increase in diet quality compared to households in the middle and upper quantiles of the HEI distribution. We also find that the association between health information and the healthfulness of food purchases is weaker in households with higher education, suggesting that high education households may be less likely to be affected by media information. These findings suggest that current media campaigns aimed to increase the number of accurate media stories on nutrition and healthy diets might be effective in encouraging the purchases of healthier food and beverages but may be slightly more effective for households with the least healthy retail food purchases. Public campaigns may want to consider improving information accuracy, particularly over health-related news and information where misinformation tends to sprout. Furthermore, there is potential to tailor targeted media campaigns to appeal more effectively to communities with lower levels of education and higher rates of unhealthy food purchases,). However, further research is necessary to determine the specific focus and effectiveness of these initiatives in promoting healthier food choices across diverse demographic segments.

## Data

The 2015-2018 IRI Consumer Network Panel (IRI-CNP) and the Purchase-to-Plate Crosswalk (PPC) released by USDA Economics Research Service are the core data for our analysis. The IRI-CNP data contain demographic information for a sample of U.S. households and records households' all retail purchases of food for at-home consumption at the barcode level, including quantities, prices, discounts, and coupons.[1] Households are incentivized to record all of their barcoded purchases with a handheld in-home scanning device. Our final sample includes over 90,000 households, who belong to the "Static Panel" of households that reliably scan their purchases throughout the year, and which have assigned sample weights to result in nationally representative consumer purchases.[2]

While the IRI data include some nutrition data, it is not sufficient. We import data from USDA's Food and Nutrient Database for Dietary Studies (FNDDS) (Martin et al, 2014) and Food Pattern Equivalent Database (FPED) (Bowman et al. 2014) via the Purchase-to-Plate Crosswalk (PPC) (Carlson et al 2019). The crosswalk includes a linking database that links over 95 percent of individual items purchased by IRI-CNP participants to the USDA databases, and conversion

---

[1] Foods purchased from restaurants, fast-food, delis and other establishments where the primary food sold is prepared food are not included in the IRI data.

[2] The number of households in the "Static Panel" varies by month.

factors which convert the purchase weight to the edible weight. The edible weight is used in the USDA nutrition data and represents the weight of the food after the inedible parts (such as skins, bones, seeds and shells) are removed.

Similar to other studies employing household scanner data to assess nutritional quality (Hastings et al., 2021; Hut and Oster, 2022), the IRI-CNP data confront certain limitations. They do not include food intended for consumption away from home, such as in restaurants or schools. Additionally, the data primarily capture purchases of packaged items with barcodes, omitting random-weight products like bulk nuts or granola. As household purchase data, we cannot identify the dietary intake of individual household members or ensure equitable product allocation within households. Recent studies, however, indicate that household food purchases offer a reasonably accurate estimate of overall diet quality (Vepsäläinen et al., 2022).

*Healthy Eating Index-2015*

The HEI-2015 is an index of adherence to key recommendations in the DGA 2015, whose scores range from 0 to 100, with 100 representing a diet that perfectly adheres to the recommendations. Since we use the scores in regression analysis, we calculate household level HEI-2015 scores for each household in a month. While the HEI was originally designed to use with dietary intake data, it has been used with food availability (Miller et al., 2015), food pantry shelves (Grannon et al., 2016), household food acquisition (Mancino et al. 2018) and retail scanner data (Carlson et al, 2019). Our HEI scores are based on household retail food purchases similar to Chrisinger et al. (2018) except that our HEI scores are calculated on the monthly level rather than for each purchase occasion. In general, household food purchase data has been shown to be consistent with overall diet quality measures from recalls but may differ on specific nutrient intake (Appelhans et al 2017). For this reason, we limit our investigation to the total HEI score and do not analyze the individual HEI components separately.

Table 1 presents the weighted descriptive statistics. In our sample, the mean monthly HEI-2015 is 49.95, which is lower than estimates based on a simple average of individual scores using dietary recall data (54) or FoodAPS data for large grocery stores (52) (Mancino, Todd, and Scharadin, 2018). Sweitzer et al. (2017) find that expenditures in the IRI-CNP are lower than in the Consumer Expenditure survey and the National Food Acquisition and Purchase Survey (FoodAPS). Higher income and larger households were more likely to under report than other households. Under reporting is particularly apparent for produce, eggs, seafood, and processed vegetables. More importantly, the PPC only covers about half of produce purchases made by IRI-CNP participants, because participants do not record quantities for random weight items.[3] Since fruits and vegetables comprise 20 percent of the HEI-2015 score, under reporting of produce and processed fruits and vegetables, as well as not being able to include almost half of produce sales in the HEI estimate are likely the main reason for differences between our estimates and other estimates.

*A measure of health Information in the media*

Consumers' knowledge of the healthy diet and nutritional information might be affected by information provided by mass media. To measure the amount of health and diet related information available to consumers, we use the NexisUni database to search for media stories covering healthy diets. The database provides access to over 15,000 news, business, and legal

---

[3] Random weight items are foods that consumers or stores package themselves, such as loose fruits and vegetables or bulk coffee beans.

**Table 1: Summary Statistics, Estimated Using Survey Weights**

|  | All Sample | |
| --- | --- | --- |
| **Variable** | **Mean** | **S.D.** |
| Household HEI-2015 of retail food purchases | 49.9515 | 12.3185 |
| Health News | 2,596.3800 | 388.2719 |
| *Education* | | |
| Less than High School | 0.0031 | 0.0558 |
| High School | 0.1759 | 0.3807 |
| College | 0.6195 | 0.4855 |
| Postgraduate | 0.2015 | 0.4011 |
| *Race and Ethnicity* | | |
| Hispanic | 0.0769 | 0.2664 |
| White | 0.7613 | 0.4263 |
| Black | 0.1214 | 0.3266 |
| Asian | 0.0433 | 0.2036 |
| Other Races | 0.0739 | 0.2617 |
| Income (in $1,000) | 74.1706 | 55.0340 |
| Employed part-time | 0.2210 | 0.4149 |
| Employed full-time | 0.6251 | 0.4841 |
| Presence of Children | 0.3237 | 0.4679 |
| Household Size | 2.5544 | 1.4533 |
| Q1 – Jan, Feb, March | 0.2489 | 0.4324 |
| Q2 – April, May, June | 0.2520 | 0.4342 |
| Q3 – July, Aug, Sept | 0.2510 | 0.4336 |
| Q4 – Oct, Nov, Dec | 0.2480 | 0.4319 |
| No. of Observations | 2,906,400 | |

*Source*: Author estimates using data from NexisUni, the IRI Consumer Network, 2015-2018, USDA's Food and Nutrient Database for Dietary Studies, Food Pattern Equivalent Database, and the Purchase to Plate Crosswalk. State summary statistics are not available to maintain confidentiality of the data. The households are weighted using the projection factors.
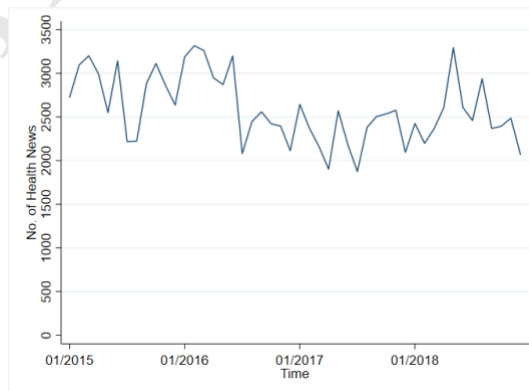


**Figure 1: Average Number of Health News Items Over Time**
*Source*: Author estimates using NexisUni Data.

sources, which is the world's largest electronic database for legal and public-records–related information. We consider newspapers, magazines, TV, radio, and online sources in our study. Specifically, we searched for all media stories that contain the keywords related to the topic of health and diet and include three groups of keywords: 1) healthy diets, healthy food, healthy dinner, healthy lunch, healthy breakfast, healthy eating healthful diet, healthy snack; 2) MyPlate, food pyramid, food group, my pyramid; 3) weight loss, diet plan, diet recommendation. We use the total number of media stories in each month in a state as proxies for the amount of information.

    The summary statistics for media coverage data are also shown in Table 1. On average, there are 2,596 media stories covering health information related to diet a month. Figure 1 presents how the media coverage changed over time during our sample period. Overall, the amount of health information is relatively stable, with the numbers being slightly higher in 2015 and 2016.

*Other covariates*

Demographic information of households is included into our analysis to evaluate the heterogeneous effects of health information across education level. The weighted summary statistics are also shown in Table 1.[4] Education has been found to be positively associated with dietary quality and we include indicators for households' education levels. To control for differences in income, we include the midpoints of the eight income brackets included in the IRI-CNP data and the weighted median income in our sample is $74,170. From our base data, we also include household size and the average household size in our sample is 2.55. The composition of the household also has a link with dietary quality, and we include a dummy variable of children, which equals to 1 if there is at least one child under age 18 in the household. On average, 32% of households report having children in the sample. Employment type may also impact the time that households have available to purchase and prepare healthful and we include indicators for full and part-time employment of the household head(s) (the excluded group is not employed). We include indicators for the race (White, Black, Asian, other or mixed) and ethnicity (Hispanic) principal survey respondent because dietary quality can differ by race and ethnicity. We also include market and seasonal (calendar quarters) indicators to control for purchase differences across space and that vary within the year. These social-demographic indicators also control for different levels of under reporting by IRI-CNP households.

## Model

*Econometric approach and specification*

To determine the impact of media stories on different quantiles of diet quality, we utilize the unconditional quantile regression (UQR) approach proposed by Firpo et al. (2009). UQR offers several distinct advantages over the conditional quantile (CQR) approach, namely that the UQR measures the unconditional association of health information in the media on household diet quality and does not depend on the covariates available or their specific values. The primary advantage of the UQR approach is that the estimated coefficient of the UQR measures the association of health news stories on household diet quality *unconditional* on the presence or specific values of the covariates. In contrast, the estimated coefficients of the CQR measure the association of the number of health news stories on a quantile of household diet quality *conditional* on the specific values of the other covariates, and the interpretation of the estimated association is limited, particularly in the presence of multiple covariates (Borah and Basu 2013). In policy analysis, the unconditional relationship between the covariates and household diet quality is often relevant. UQR is therefore recommended when addressing questions of policy

---

[4] Demographic information is collected on a yearly basis by IRI. Demographics that change within the year (births, deaths, moves) will only be reflected when IRI distributes their annual survey.

relevance, for example, the association of the health news story environment and household diet quality (Borah and Basu 2013). [5] This UQR approach has been used to provide policy recommendations in a food and/or health context to measure medication adherence (Borah and Basu 2013), the relationship between income and health biomarkers (Carrieri and Jones 2017), the relationship between income and body mass index (BMI) (Rodriguez-Caro et al 2016), and the relationship between BMI and food label use (Bonanno et al 2018).

We assume a simple linear relationship between the healthfulness of household $h's$ retail food purchases in market $m$ and time $t$, $H_{hmt}$, and the amount of health information in the media in market $m$ where the household lives at time $t$, $INFO_{mt}$. The one-period lag of the health information, $INFO_{m,t-1}$, is also included in the model to capture the potential carry-over effect of health information that may last more than one period. Most households stay multiple years in the IRI-CNP. To incorporate the panel feature of our data, we conduct a Panel UQR regression with the household fixed effects. The relationship is given in the following expression:

$$(1) \qquad H_{hmt} = \beta_0 + \beta_1 INFO_{mt} + \beta_2 INFO_{m,t-1} + Season_t + \ Market_h + HH_h + e_h$$

where $\beta$ is a conformable vector of parameters to be estimated and $e_h$ is an error term. *Season* and *Market* are vectors of dummy variables for seasonality and market. $HH_h$ denotes the household fixed effects.

To estimate the Panel UQR specified in equation (1), we employ the re-centered influence function (RIF) method of Firpo et al. (2009). The RIF summarizes the impact of an individual observation on a given quantile of household diet quality. Firpo et al (2009)'s method uses RIFs to estimate unconditional partial effects of infinitesimal changes in the distribution of the covariates on a given quantile of household diet quality. We follow Firpo et al (2009) and assume a linear relationship between the RIF of a given quantile of household diet quality and the covariates. Under the assumption of linearity, we can use OLS to capture how marginal changes in the distribution of the covariates affects a given quantile of household diet quality. However, instead of using the $\tau$th quantile of household diet quality, $q_\tau$, as the dependent variable, we use its RIF, and the expectation of the RIF regression function for the household can be specified as:

$$(3) \qquad E[RIF(H_{hmt}; q_\tau|X)] = \beta_0 + \beta_1 INFO_{mt} + \beta_2 INFO_{m,t-1}$$
$$+ Season_t + Market_h + HH_h + e_h$$

where X include $INFO_{mt}, INFO_{m,t-1}, Season_t, Market_h$, and $HH_h$. The parameter estimates can be interpreted as the unconditional marginal effects at each quantile. We employ the RIF algorithm developed by Firpo et al (2009) within the Stata 18 framework.

*Identification*

There are potentially two types of endogeneity related to health news stories. It is worth noting that our health information is collected at the aggregate state level. That is, we only know how many pieces of health news stories are available to all consumers in a state in a month, but not the individual reception (consumption) of health news stories. In addition, our measure of the information volume may not include all the information that a consumer could receive and not all consumers are exposed to the health information. The individual reception of the health news stories might depend on individual subscriptions to media sources, which may vary across household demographics. Although the arrival of news is exogenous to individual consumers when making food-related decisions, with the use of the state level health news environment, there might be a potential concern of endogeneity in the sense that there might be more health news in

---

[5] For a complete discussion on the differences between conditional mean, conditional quantile, and unconditional quantile estimation see Borah and Basu (2013).
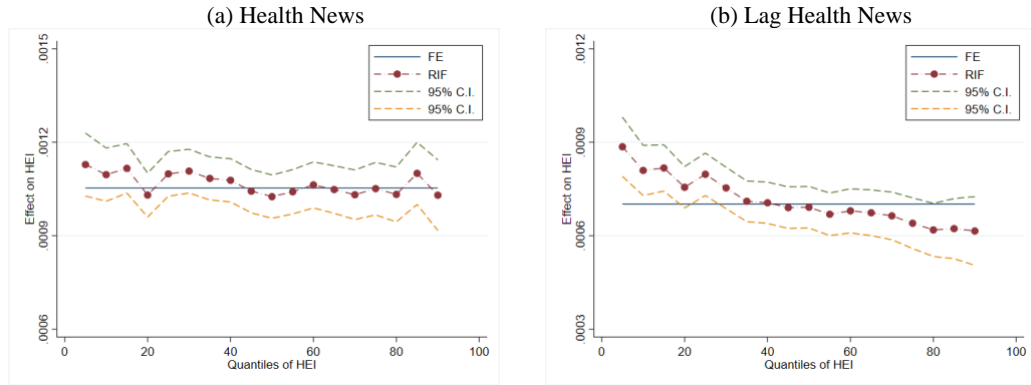
**Figure 2. Estimated Relationship between HEI and Health News**
*Source*: Author estimates using data from NexisUni, the IRI Consumer Network, 2015-2018, USDA's Food and Nutrient Database for Dietary Studies, Food Pattern Equivalent Database, and the Purchase to Plate Crosswalk. The figures are drawn using estimates from Table 2. The household, season, and state fixed effects were included in the regressions. HEI = Healthy Eating Index.

states that have higher levels of health and income. To alleviate this concern, we use the average obesity rate and average income at the state and annual level as instruments for the number of health news pieces. It is less likely that a household's HEI will be affected directly by the state average. We use a control function approach to estimate the unconditional quantile regression models with instruments. Our conducted endogeneity test affirms its presence, and the instrument tests support the statistical validity of our chosen instruments.

However, it's crucial to note that this analysis faces additional sources of endogeneity. The way individuals receive and respond to information can vary significantly. Some households may be more health-conscious, actively engaging with health news and making healthier food choices, while others may not be paying attention to health news. Unfortunately, data on the specific reception of information at the household level is unavailable to us. Consequently, we are unable to address the endogeneity arising from the unobserved reception of health information and its impact on the Healthy Eating Index (HEI).

Additionally, our analysis does not incorporate a causal mechanism to explain how health news stories might influence individual food consumption and, consequently, their HEI. Developing such causal mechanisms would require a theoretical framework, which is beyond the scope of our current study. Thus, while we can partially mitigate the endogeneity related to the health news environment, we refrain from making any claims about causal effects.

## Results

### Panel UQR Regressions

Table 2 shows the results of Panel UQR described in equation (3). Column 1 includes the results from the fixed effect (FE) regression at the mean and columns 2-8 presents the results of the Panel UQR regression at the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentile of households HEI. Further, to make the interpretation of our coefficients of interest easier, the estimated coefficients for Panel UQR regressions at every 5 percentiles of the HEI distribution are plotted in Figure 2, along with their 95% confidence intervals. As a comparison, the FE regression results are also plotted as a straight line.

Panel (a) of Figure 2 shows the relationship between health information and HEI. The FE analysis at the mean suggests that one piece of health news is associated with an increase in HEI by 0.0011 points for an average household in a month. Since the average number of 2,596 pieces

**Table 2: Relationship Between HEI and Health News: Panel UQR**

| Variable | FE (1) | Q5 (2) | Q10 (3) | Q25 (4) | Q50 (5) | Q75 (6) | Q90 (7) | Q95 (8) |
|---|---|---|---|---|---|---|---|---|
| | | | | Panel UQR Regressions | | | | |
| Health News | 0.0011*** | 0.0011*** | 0.0011*** | 0.0011*** | 0.0010*** | 0.0011*** | 0.0010*** | 0.0010*** |
| | (0.0000) | (0.0001) | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0001) | (0.0001) |
| Lag Health News | 0.0007*** | 0.0009*** | 0.0008*** | 0.0008*** | 0.0007*** | 0.0006*** | 0.0006*** | 0.0006*** |
| | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0000) | (0.0001) | (0.0001) |
| State FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Household FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Season FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| No. of Obs. | 2,906,400 | 2,906,400 | 2,906,400 | 2,906,400 | 2,906,400 | 2,906,400 | 2,906,400 | 2,906,400 |

*Note*: Robust standard errors below coeffecent estimate. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The demographic variables are omitted from the regression because there are minimal variation in demogrphics within a household over time. Household fixed effects, season fixed effects, and market fixed effects were included in the regressions, however omitted here for brevity. Both the FE and Panel UQR are estimated with instruments described in the Model section. HEI = Healthy Eating Index; UQR = unconditional quantile regression.

*Source*: Author estimates using data from NexisUni, the IRI Consumer Network, 2015-2018, USDA's Food and Nutrient Database for Dietary Studies, Food Pattern Equivalent Database, and the Purchase to Plate Crosswalk.

**Table 3: Relationship Between HEI and Health News Across Education Levels**

| | | | | RIF Regressions | | | | |
|---|---|---|---|---|---|---|---|---|
| | **OLS** | **Q05** | **Q10** | **Q25** | **Q50** | **Q75** | **Q90** | **Q95** |
| **Variable** | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** | **(6)** | **(7)** | **(8)** |
| Health News | 0.0020*** | 0.0024*** | 0.0022*** | 0.0018*** | 0.0020*** | 0.0019*** | 0.0024*** | 0.0021*** |
| | (0.0004) | (0.0009) | (0.0007) | (0.0005) | (0.0005) | (0.0005) | (0.0007) | (0.0008) |
| Health News * | | | | | | | | |
| High School | -0.0010*** | -0.0012 | -0.0010 | -0.0007 | -0.0010** | -0.0010* | -0.0016*** | -0.0015* |
| | (0.0004) | (0.0009) | (0.0007) | (0.0005) | (0.0005) | (0.0005) | (0.0007) | (0.0008) |
| Health News * | | | | | | | | |
| College | -0.0008** | -0.0014 | -0.0012* | -0.0007 | -0.0009* | -0.0006 | -0.0011 | -0.0008 |
| | (0.0004) | (0.0009) | (0.0007) | (0.0005) | (0.0005) | (0.0005) | (0.0007) | (0.0008) |
| High School | 2.7886*** | 4.6610** | 3.7954** | 2.5036* | 2.8373*** | 2.0968 | 3.4964** | 3.1816 |
| | (0.9484) | (2.2268) | (1.6939) | (1.3094) | (1.1722) | (1.3053) | (1.7043) | (2.0814) |
| College | 4.7239*** | 6.1607*** | 5.5480*** | 4.4698*** | 5.0016*** | 4.1049*** | 5.2245*** | 4.6546** |
| | (0.9435) | (2.2183) | (1.6861) | (1.3026) | (1.1661) | (1.2995) | (1.6987) | (2.0756) |
| Demographic Variables | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| State FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Season FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| No. of Obs. | 2,906,400 | 2,906,400 | 2,906,400 | 2,906,400 | 2,906,400 | 2,906,400 | 2,906,400 | 2,906,400 |

*Note*: Robust standard errors below coeffecient estimate. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The demographic variables, season fixed effects, and state fixed effects were included in the regressions, however omitted here for brevity. HEI = Healthy Eating Index; 2SLS=Two Stage Least Squares; RIF = Recentered Influence Function.
*Source*: Author estimates using data from NexisUni, the IRI Consumer Network, 2015-2018, USDA's Food and Nutrient Database for Dietary Studies, Food Pattern Equivalent Database, and the Purchase to Plate Crosswalk.

of news a month, health information is associated with a contribution to a household's HEI by 2.86 points a month.[1] The results imply the importance of health information: with the constant flow of health news from various media sources, it is possible for health information to have a continued contribution to a household's purchases of healthy foods.

The Panel UQR regressions indicate that the relationship between health information and HEI is positive and statistically significant across all quantiles of the HEI distribution but varies in magnitude along the distribution. Specifically, the association between health information and HEI is slightly higher at the lowest quantiles of the HEI distribution. However, the differences across quantiles are very small and the FE estimate lies within the confidence interval for most quantile estimates. Therefore, we are not able to reject equivalence with the estimates at the mean for most quantiles. For households with very unhealthy food purchase baskets, providing more health information is slightly associated with larger increases compared to very healthy households. One possible explanation could be that households with very unhealthy food choices might have a lower baseline awareness of nutritional information or the health implications of their food selections. Therefore, the introduction of health information could lead to more significant changes in behavior in these households. Further, households with unhealthy food habits might be more slightly motivated to make positive changes when provided with health information. In addition, lagged health news also have a positive and significant impact on HEI, but the magnitude is smaller compared to the current period health news, which suggests that the health media stories viewed one month ago could be forgotten by the next month.

As a robustness analysis, we further conduct a UQR analysis with household demographic variables, without including the panel features. The results are presented in the Table S1, are consistent with the Panel RIF results: one piece of health news is associated with an increase in HEI by 0.0011 points for an average household in a month at the mean. Similar to the Panel RIF regression results, we find that the differences across quantiles are very small and the 2SLS Panel Regression estimate at the mean lies within the confidence interval for most quantile estimates, and we are not able to reject equivalence with the estimate at the mean for most quantiles.

*Health information and education*

Personal health literacy refers to the degree to which individuals have the ability to find, understand, and use information and services to inform health-related decisions and actions for themselves and others.[2] From newspaper to online blogs, the constant and large stream of health news can make it difficult for consumers to distinguish reliable information. They could be confused by the vast amount of, and sometimes conflicting, health related news they received or have difficulty understanding the benefit or problems of certain diets or nutrition. Therefore, the relationship between health news and diet quality could vary depending on the levels of health literacy of a consumer. In this section, we use the highest education level in a household as a proxy of the households' level of health literary and evaluate the heterogeneous association between health news and healthfulness of retail food purchases across households with varying education levels. Health literacy have been shown to have a relationship with level of education: people with lower education were found to demonstrate lower health literacy skills in comparison with people with higher education (Howard et al., 2006; Lee et al., 2010).

We classify education levels obtained by the household heads into three categories and use two dummy variables: *High School* and *College and above,* using less than high school as the reference. Specifically, we include the interactions of health news and these education-level

---

[1] One of the ways that Guenther et al (2014) validate the HEI as a measure of diet quality is by finding a difference in scores between men and women of about 2 points.

[2] The definition of health literacy was updated in August 2020 with the release of the U.S. government's Healthy People 2030 initiative. https://health.gov/healthypeople/priority-areas/health-literacy-healthy-people-2030
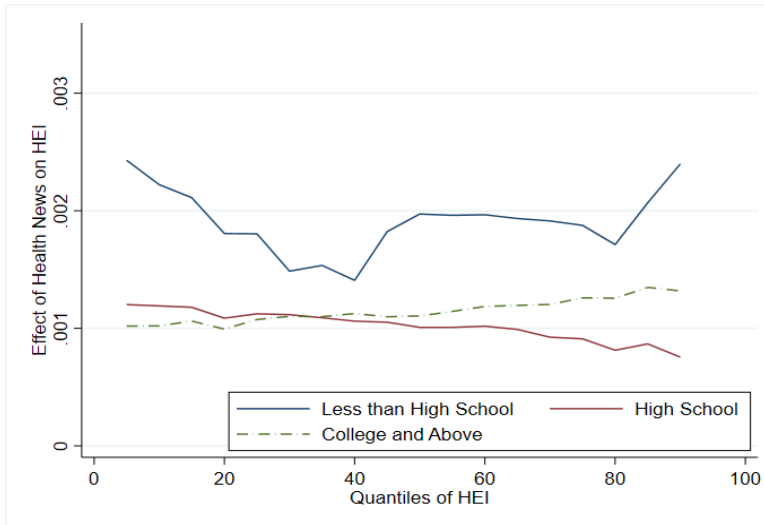
**Figure 3: Estimated Relationship between Health News and Healthfulness of Retail Food Purchases by Education Levels**
*Source*: Author estimates using data from NexisUni, the IRI Consumer Network, 2015-2018, USDA's Food and Nutrient Database for Dietary Studies, Food Pattern Equivalent Database, and the Purchase to Plate Crosswalk

dummy variables. The results are presented in Table 3. From the 2SLS results, the coefficients of the interaction terms are all negative and significant, suggesting the significant additional effect of health information among higher-educated households relative to lower-educated households. We further calculate the net effects of the health information across education levels. On average, one additional piece of health news is associated with an increase in the households' HEI by 0.0020 for households with heads not finishing high school, 0.0010 for high school education, 0.0012 for college education and above. Figure 3 plots the net relationship between health information and HEI across education levels. The relationship between health information and households whose heads have less than a high school education is significantly stronger compared to households with higher education levels. Further, the effect is higher at the highest and the lowest quantiles of the HEI distribution for households with low education levels. Between households with high school and college and above, there is no statistically significant difference in the effects of health information across almost all quantiles of diet qualities, except for the highest quantiles.

At the first glance, these results are contradictory to the concept of health literacy, where we might expect that the high educated groups would be more willing to adopt the health information and change their diet decisions. However, considering the increase in exposure to media in all forms and the vast amount of information consumers may receive, it is possible that some information might not be reliable. High education groups usually have a higher diet quality already and are more likely to think critically about the information they receive (Huber and Kuncet, 2016). Therefore, they are less likely to be affected by media information. On the contrary, the consumers with lower education may be more easily influenced by information from media. In addition, the households with higher education levels are more likely to have more existing knowledge of a healthy diet and a greater understanding of dietary quality (Johnston et al., 2015), and therefore, they have less to gain from the media stories than lower educated households and the association is not as strong as those on households with lower education levels.

## Conclusions and Policy Implications

We use a panel unconditional quantile regression model of the healthfulness of household retail food purchases to investigate the relationship between health information and household diet quality beyond the mean. We quantify the volume of health and diet-related information available to consumers using the number of media stories on newspapers, magazines, TV, radio, and online sources covering healthy diets, which is collected from the NexisUni database. Using 2015-2018 household-level retail food purchase data, we estimate the 2015 HEI for each household by matching purchases to nutrient and food composition data. Results from this paper shed light on related government policies aiming to improve the healthfulness of diet.

First, we find that the association between health information and healthfulness of retail food purchases is higher at the lowest quantiles of the HEI distribution, although we cannot reject equivalence with the FE estimate for most quantiles. For households with very unhealthy retail food purchases, providing more health information is associated with a slightly larger effect compared to households whose retail food choices are in the middle and high quantiles of healthfulness. One possible explanation could be that households with very unhealthy food choices might have a lower baseline awareness of nutritional information or the health implications of their food selections. Therefore, the introduction of health information could lead to more significant changes in behavior in these households. Further, households with unhealthy food habits might be more slightly motivated to make positive changes when provided with health information. This could imply that media stories, educational efforts, or initiatives focused on health information may have a more noticeable effect on households with less healthy food choices, potentially contributing to improved healthfulness of food purchases. To provide a more effective campaign, many healthcare professionals, organizations, and governments could use this information to allocate their resources more efficiently.

Regarding the interpretation of our results, there are also a couple of limitations that need to be acknowledged. First, our calculation of HEI scores only covers households' retail food purchases, but not food away from home or other sources. Although household food purchase data has been shown to be consistent with overall diet quality it may differ on specific nutrient intake. Second, our HEI score estimates are lower compared to other studies based on a simple average of individual scores using a week of food acquisitions, most likely due to under reporting of some food categories such as fruits and vegetables, and the fact that random weight items, such as loose fruits and vegetables are not included in the data. Third, due to data limitations, we focus on the general health news environment at the aggregate state level, which is the volume of available health news stories available to all consumers in a state in a month. Our measure of the information volume may not include all the information that a consumer could receive and not all consumers are exposed to all health information. The individual exposure to available health news stories might depend on subscriptions to media sources, which may vary across household demographics.

Second, although the impact of one piece of health-related news on a household's monthly HEI score is relatively modest (0.0011 points), the average monthly impact of a constant flow of health news is meaningful. Health-related news is associated with an average 2.86-point increase in HEI in a month. This contribution aligns with the effects of other influential demographic variables; for instance, college education of the household head(s) is associated with 2.59 points increase in HEI, compared to households headed by individual(s) without a high school degree (see Appendix Table 1). Comparing this effect to other interventions influencing diet quality reveals notable findings. Hasting et al. (2021) finds that the effect of Supplemental Nutrition Assistance Program (SNAP) participation on HEI is small, ranging from $-0.101$ to $0.448$ points. Feng, Fan, and Jaenicke (2023) find that although SNAP has no significant impact on households' dietary quality on average, for households with initially low-to-intermediate dietary quality, SNAP participation reduces their HEI scores by more than 7 points. Allcot et al. (2019) finds that entry of supermarkets will increase the health index for households living in food desert by 0.014

points. Scharadin and Jaenicke (2020) estimated that a one-hour reduction in secondary childcare per day leads to a 2.35-point increase in household HEI, and 30 min of additional food-at-home time per day would increase HEI by 2.01 points. From a policy perspective, our results imply that a constant flow of accurate health-related news positively contributes to a household's diet quality. In contrast, the demographic variables are impossible, or difficult to change in the short run. However, a successful media campaign typically stimulates extensive media coverage across various platforms, including TV, radio, and online sources. Additionally, it may foster broader discussions related to healthy diets and nutrition in the media landscape. Our results suggests that the improve health news environment could be effective in encouraging the purchases of healthier food and beverages thus improving diet quality.

Third, considering the increase in exposure to media in all forms and the vast amount of information consumers may receive, it is possible that some information might not be reliable. Nutrition research is complex and is often oversimplified by the media. It is difficult for readers to distinguish reliable research from weak studies and sensational headlines. Our results finds that consumers with lower education may be more easily influenced by information from media, compared to higher educated groups. As a result, effort should continue to improve information accuracy, particularly over health-related news and information where misinformation tends to sprout.

Overall, our study sheds light on the relationship between health news environment and the healthfulness of household retail food purchases. Our findings underscore the potential of media campaigns to positively impact diet quality, highlighting potential benefits of targeted efforts to improve information accuracy and accessibility, particularly for vulnerable populations. By leveraging the influence of media, policymakers and healthcare professionals can play a vital role in promoting healthier food choices and improving public health outcomes.

*[First submitted September 2023; accepted for publication April 2024.]*

## References

Alé-Chilet, J. and Sarah Moshary. 2022. Beyond Consumer Switching: Supply Responses to Food Packaging and Advertising Regulations. *Marketing Science*, 41(2):243-270.

Andersen, J. L. Wylie and E.M. Brank. 2017. Public health framing and attribution: Analysis of the first lady's remarks and news coverage on childhood obesity, *Cogent Social Sciences*, 3:1. DOI:10.1080/23311886.2016.1268748

Beck, F., Richard, J.-B., Nguyen-Thanh, V., Montagni, I., Parizot, I., Renahy, E.2014. Use of the Internet as a health information resource among French young adults: results from a nationally representative survey. *J. Med. Internet Res*. 16(5), e128. DOI: 10.2196/jmir.2934

Bonanno, A., F. Bimbo, R. Cleary, and E. Castellari. 2018. Food Labels and Adult BMI in Italy – An Unconditional Quantile Regression Approach. *Food Policy*, 74: 199-211. DOI:10.1016/j.foodpol.2017.12.008

Borah, B.J. and A. Basu. 2013. Highlighting the Differences Between Conditional and Unconditional Quantile Regression Approaches Through an Application to Assess Medication Adherence. *Health Economics*, 22(9): 1052-1070. DOI: 10.1002/hec.2927

Bowman, S.A., J.C. Clemens, J.E. Friday, R.C. Thoerig, and A.J. Moshfegh. 2014. *Food Patterns Equivalents Database 2011-12: Methodology and User Guide*, U.S. Department of Agriculture, Agricultural Research Service, Beltsville Human Nutrition Research Center, Food Surveys Research Group. DOI:10.13140/RG.2.1.2201.5520

Brown, D.J., and L.F. Schrader. 1990. Cholesterol Information and Shell Egg Consumption. American Journal of Agricultural Economics 72(3):548–55. DOI:10.2307/1243023

Browning, M., L.G. Hansen, and S. Smed. 2019. Heterogeneous Consumer Reactions to Health News. *American Journal of Agricultural Economics*, 101(2): 579-599. DOI:10.1093/ajae/aay055

Butteriss, C. and J. Bradley. 2019. The World's Best Public Health Social Media Campaigns. www. bangthetable.com

Buyuktuncer, Z., Ayaz, A., Dedebayraktar, D., Inan-Eroglu, E., Ellahi, B. and Besler, H.T., (2018). Promoting a healthy diet in young adults: the role of nutrition labelling. *Nutrients*, 10(10), p.1335. DOI: 10.3390/nu10101335

Campos, S., J. Doxey, and D. Hammond.2011. Nutrition labels on Pre-Packaged Foods: A Systematic Review. *Public Health Nutrition*,14 (8): 1496–506. DOI: 10.1017/S136898

Carlson, A., E. Page, T.P. Zimmerman, C.E. Tornow, and S. Hermansen. 2019. *Linking USDA Nutrition Databases to IRI Household-Based and Store-Based Scanner Data*, TB-1952, U.S. Department of Agriculture, Economic Research Service.

Carrieri, V. and A.M. Jones. 2017. The Income-Health Relationship 'Beyond the Mean': New Evidence from Biomarkers. *Health Economics*, 26: 937-956. DOI: 10.1002/hec.3372

Chrisinger, B. W., Kallan, M. J., Whiteman, E. D., & Hillier, A. (2018). Where do US households purchase healthy foods? An analysis of food-at-home purchases across different types of retailers in a nationally representative dataset. *Preventive Medicine*, 112, 15-22. DOI: 10.1016/j.ypmed.2018.03.015

Christoph, M.J. and An, R., (2018). Effect of nutrition labels on dietary quality among college students: a systematic review and meta-analysis. *Nutrition reviews*, 76(3), pp.187-203. DOI: 10.1093/nutrit/nux069

Cleary, Rebecca; Bonanno, Alessandro; Ghazaryan, Armen; Bellows, Laura; McCloskey, Morgan, 2021. School Meals and Quality of Household Food Acquisitions, *Applied Economic Perspectives and Policy*, v. 43, iss. 4, pp. 1385-1442. DOI:10.1002/aepp.13130

Diviani, N. B. van den Putte, S. Giani, and J. van Weert. 2015. Low Health Literacy and Evaluation of Online Health Information: A Systematic Review of the Literature. *Journal of Medical Internet Research.* 17(5). DOI: 10.2196/jmir.4018

eMarketer. 2021. US Time Spent with Media 2021. w.insiderintelligence.com/content/us-time-spent-with-media-2021

Firpo S, Fortin N, Lemieux T. 2009. Unconditional quantile regressions. *Econometrica* 77(3): 953–973.

Grunert, K.G. and J.M. Wills. 2007. A Review of European Research on Consumer Response to Nutrition Information on Food Labels. *Journal of Public Health*, 15(5):385-99.

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

Huang, Lu, and Yizao Liu, 2017. Health information and consumer learning in the bottled water market, *International Journal of Industrial Organization,* Volume 55, Pages 1-24. DOI:10.1016/j.ijindorg.2017.08.002

Huber, C. and N.R. Kuncel. 2016. Does College Teach Critical Thinking? A Meta-Analysis. *Review of Educational Research*. 86(2): 431-468. DOI:10.3102/0034654315605917

Katare, Bhagyashree; Binkley, James K.; Chen, Kaiyan, 2021. Nutrition and Diet Quality of Food at Home by Supplemental Nutrition Assistance Program (SNAP) Status, *Food Policy*, v. 105. DOI:10.1016/j.foodpol.2021.102165

Im, H, and J. Huh. 2017. Does Health Information in Mass Media Help or Hurt Patients? Investigation of Potential Negative Influence of Mass Media Health Information on Patients' Beliefs and Mediacation Regimen Adherence. *Journal of Health Communication,*22:214-222. DOI:10.1080/10810730.2016.1261970

Lee S. Y., Tsai T. I., Tsai Y. W., Kuo K. N. Health literacy, health status, and healthcare utilization of Taiwanese adults: Results from a national survey. *BMC Public Health*. 2010;10:614. DOI: 10.1186/1471-2458-10-614

Liu, P., K. Lien, and F. Asche. 2016. The Impact of Media Coverage and Demographics on the Demand for Norwegian Salmon. *Aquaculture Economics & Management*, 20 (4):342–56. DOI:10.1080/13657305.2016.1212126

Mancino, L., J. Guthrie, M.V. Ploeg, and B.H. Lin. 2018. *Nutritional Quality of Foods Acquired by Americans: Findings from USDA's National Household Food Acquisition and Purchase Survey,* EIB-188, U.S. Department of Agriculture, Economic Research Service.

Mancino, L., J.E. Todd, and B. Scharadin. 2018. *USDA's National Household Food Acquisition and Purchase Survey: Methodology for Imputing Missing Quantities to Calculate Healthy Eating Index-2010 Scores and Sort Foods into ERS Food Groups,* TB-1947, USDA-ERS,

Martin, C.L., J.B. Montville, L.C. Steinfeldt, G. Omolewa-Tomobi, K.Y. Heendeniya, M.E. Adler, and A.J. Moshfegh. 2014. *USDA Food and Nutrient Database for Dietary Studies 2011-2012*, USDA, Agricultural Research Service, Food Surveys Research Group.

Miller, P.E., J. Reedy, S.I. Kirkpatrick, and S.M. Krebs-Smith. 2015. The United States Food Supply Is Not Consistent with Dietary Guidance: Evidence from an Evaluation Using the Healthy Eating Index-2010, *Journal of the Academy of Nutrition and Dietetics* 115 (1):95-100. DOI: 10.1016/j.jand.2014.08.030

Nagler, Rebekah H. 2014. Adverse Outcomes Associated with Media Exposure to Contradictory Nutrition Messages. *Journal of Health Communication*, 2014; 19 (1): 24 DOI: 10.1080/10810730.2013.798384

Rieger, J., C. Kuhlgatz, and S. Anders.2016. Food Scandals, Media Attention and Habit Persistence among Desensitised Meat Consumers. *Food Policy*, 64: 82–92.DOI:10.1016/j.foodpo

Smed, S. 2012. Information and Consumer Perception of the Organic Attribute: The Case of Fresh Fruit and Vegetables. *Agricultural Economics* 43 (S1): 33–48.

Smith, T.A. 2017. "Do School Food Programs Improve Child Dietary Quality*?" American Journal of Agricultural Economics*, 99 (2): 339-356. DOI:10.1093/ajae/aaw091

Suenaga, H., Vicente, M.R. 2022. Online and offline health information seeking and the demand for physician services. *Eur J Health Econ* 23, 337–356. DOI: 10.1007/s10198-021-01352-7

Sweitzer, M., D. Brown, S. Karns, M.K. Muth, P. Siegel, and C. Zhen. 2017. *Food-at-Home Expenditures: Comparing Commercial Household Scanner Data from IRI and Government Survey Data,* TB-1949, U.S. Department of Agriculture, Economic Research Service.

Tonsor, G., J. Mintert. and T. Schroeder. (2010). U.S. Meat Demand: Household Dynamics and Media Information Impacts. *Journal of Agricultural and Resource Economics*. 35(1); 1-17. DOI: 10.22004/ag.econ.61056

Volpe, V. A. Okrent, and E. Leibtag. 2013. The Effect of Supercenter-Format Stores on the Healthfulness of Consumers' Grocery Purchase. *American Journal of Agricultural Economics*. 95(3): 568–589. DOI: 10.1093/ajae/aas132

Zhu, C.; R. A. Lopez and X. Liu. 2016. Information Cost and Consumer Choices of Healthy Foods, *American Journal of Agricultural Economics*, Volume 98, Issue 1, Pages 41–53.