DISSERTATION

THREE ESSAYS ON NUTRITIONAL DISPARITIES AND POLICIES IN U.S. FOOD PURCHASES

Submitted by

Duoyu Wang

Department of Agricultural and Resource Economics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2025

Doctoral Committee:

Advisor: Rebecca Cleary

Alessandro Bonanno
Joshua Berning
Megan Mueller

ABSTRACT

THREE ESSAYS ON NUTRITIONAL DISPARITIES AND POLICIES IN U.S. FOOD PURCHASES

This dissertation contains three essays examining the nutritional disparities and policy interventions aimed at improving dietary quality in the U.S. The second chapter investigates the nutritional quality gap between food-secure and food-insecure households in the United States using the public access USDA's National Household Food Acquisition and Purchase Survey (FoodAPS). Employing an unconditional quantile decomposition approach, the study reveals that disparities in nutritional quality persist across the distribution of the Healthy Eating Index (HEI)-2010 scores. Differences in household characteristics account for the majority of the diet quality gap among households with lower nutritional quality, whereas for households with higher nutritional quality, the diet quality gap is influenced more by differences in how these characteristics translate into dietary outcomes. The third chapter examines the impact of the Supplemental Nutrition Assistance Program (SNAP) on the dietary quality of Black households using an instrumental variable unconditional quantile regression (IVUQR) approach, again utilizing the public access FoodAPS dataset. The analysis finds that SNAP participation is associated with lower dietary quality for Black households, particularly at lower-to-middle quantiles of the HEI-2010 distribution, primarily due to an increased acquisition of foods high in empty calories. Compared to White SNAP households, as represented by their primary food purchaser, Black SNAP households show more stable nutritional outcomes across food sources, meaning their diet quality remains relatively consistent between food-at-home and food-away-from-home purchases. The fourth chapter evaluates the potential impact of an industry-wide sugar content ceiling on the U.S. sugar-sweetened beverage (SSB) market. Utilizing demand estimates from a random coefficients logit model and Circana point-of-sale (Pos) scanner data, the study simulates an how industry-wide product reformulation would affect equilibrium shares and per capita sugar amounts from purchasing SSBs.

ACKNOWLEDGMENTS

This dissertation represents the end of a long and challenging journey, and I couldn't have made it this far without the help, guidance, and encouragement of so many people. I want to take a moment to thank everyone who has supported me along the way.

First, I want to express my deepest gratitude to my advisor, Dr. Rebecca Cleary, for her incredible patience, guidance, and support. She has been there for me every step of the way, offering thoughtful advice and helping me grow as a researcher. Without her efforts, this dissertation wouldn't have been possible. I am also deeply thankful to Dr. Alessandro Bonanno for his generosity and support. He played a key role in helping me access the data that was crucial to my research, and his patience, detailed feedback, and encouragement have been invaluable. I greatly appreciate the time and effort he has invested in helping me throughout this process. I would also like to thank Dr. Joshua Berning and Dr. Megan Mueller for their thoughtful feedback and constructive suggestions, which have strengthened my work and deepened my understanding throughout this process. I also want to thank Dr. Hayley Chouinard, our department head, Dr. Jordan Suter, the chair of our graduate program, and Kathy Bruce, our program coordinator, for their help and for creating such a supportive environment. Their dedication to the department has made a big difference in my graduate experience.

Most of all, I want to thank my husband and my parents for their unwavering love and support. They have been my foundation through every challenge, always believing in me even when I doubted myself. Their constant encouragement has carried me through the most difficult times, and I am endlessly grateful for their presence in my life. To my friends and fellow graduate students, thank you for the camaraderie, laughter, and shared moments of support. Your friendship has made this journey more meaningful, and I am grateful for the connections we've built along the way. Finally, thank you to everyone who has been part of this journey. Your kindness and support have meant more to me than words can fully express, and I am truly grateful for everything you have done to help me reach this milestone. Thank you all, from the bottom of my heart.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

DISCLAIMERS

The findings and conclusions in this dissertation are those of the author and should not be construed to represent any official USDA or U.S. Government determination or policy. The analysis, findings, and conclusions expressed in Chapter 4 of this dissertation should also not be attributed to Circana (formerly IRI).

Chapter 1 Introduction

Nutritional disparities persist as a major challenge in U.S. food policy, particularly affecting low-income and racially marginalized communities. While traditional efforts have focused on achieving food security, the consistent availability of enough food, recent USDA frameworks emphasize nutrition security, defined as consistent access to safe, healthy, affordable, and culturally appropriate foods essential for optimal health (U.S. Department of Agriculture, Food and Nutrition Service., 2022). This shift reflects growing recognition that caloric sufficiency alone is insufficient to promote long-term well-being. This dissertation addresses critical gaps in our understanding of nutrition disparities and evaluates the role of federal policy interventions and market regulations in shaping dietary quality. Across three essays, I investigate how nutritional inequalities arise across food security and racial lines, and assess the effects of public policies, including the Supplemental Nutrition Assistance Program (SNAP) and industry-wide sugar content reformulation, on the U.S. food purchase environment.

Chapter 2 examines whether nutrition security provides information beyond conventional food security measures by exploring differences in dietary quality between food-secure and food-insecure households. Using public access data from the USDA's National Household Food Acquisition and Purchase Survey (FoodAPS), I calculate household-level Healthy Eating Index (HEI-2010) scores based on both food-at-home and food-away-from-home acquisitions. Employing an unconditional quantile regression framework with Recentered Influence Functions, I find that disparities in nutritional quality are not uniform across the HEI-2010 distribution. On average, food-secure households have HEI-2010 scores that are 4.5 points higher than those of food-insecure households, and this nutritional quality gap widens considerably among households with higher nutritional quality. Among households with lower nutritional quality, differences in observable characteristics, such as education, marital status, and food expenditures, relatively contribute nearly all of the gap. In contrast, for households with higher nutritional quality, these characteristics account for only about 58% of the difference. As nutritional quality improves, the ability of households to translate available resources into healthier diets becomes increasingly important. For these households, both differences in characteristics and differences in how effectively those characteristics are used contribute equally to the nutritional quality gap.

Chapter 3 evaluates the causal effect of SNAP participation on the dietary quality of Black and White households, using the public access FoodAPS dataset. To address the endogeneity of SNAP participation, I employ an instrumental variable unconditional quantile regression (IVUQR) approach. I use exogenous variation in state-level welfare programs and SNAP administrative policies as instruments to control for selection into the program. The results show that SNAP participation significantly reduces dietary quality among Black households, particularly among those with lower to middle nutritional quality. This decline is primarily driven by increased consumption of foods high in empty calories, such as sugary beverages and desserts. While Black households exhibit more stable food acquisition patterns compared to White households, they do not experience corresponding improvements in dietary quality through SNAP. These findings highlight the need to reform SNAP beyond its role as a cash-transfer program and strengthen its function as a targeted nutrition intervention, especially for Black communities.

Chapter 4 shifts focus to the industry side of the nutrition environment, simulating how an industry-wide sugar content ceiling could influence the U.S. sugar-sweetened beverage (SSB) market. Using Circana (formerly IRI) Point of Sale scanner data, I estimate demand using a random coefficients logit model and simulate equilibrium outcomes under a hypothetical sugar reformulation policy. The results indicate that consumer demand for SSBs is downward-sloping with respect to price, and that product attributes such as sugar and sodium content reduce utility, while caffeine slightly enhances it. Simulation results show that sugar reformulation policies can lead to meaningful reductions in average sugar purchased per 12-ounce serving, with estimated decreases of 13.5%, 21.5%, and 29.9% under 10%, 20%, and 30% sugar reduction scenarios, respectively. Changes in market shares are modest, suggesting that such policies could improve population health without substantially altering consumer preferences or disrupting existing market dynamics.

Chapter 2 The Distribution of the Gap in Nutrition for Food Insecure U.S. Households

2.1 Introduction

Food security is defined as all people having consistent access at all times to sufficient food for an active and healthy life (U.S. Department of Agriculture, Economic Research Service., 2025). While this concept has been instrumental in addressing issues related to caloric intake, it does not always guarantee a nutritious diet. People who are considered food secure may still eat diets that lack important nutrients. These diets are often high in calories but low in nutritional value, making individuals more likely to experience obesity and other diet-related diseases (Nicklas et al., 2013; Nguyen et al., 2015; Leung et al., 2017). This concern is evident in developed countries, where people may meet calorie needs but still have poor diets. For example, despite approximately 8.9% of U.S. households with children experiencing food insecurity in 2023 (U.S. Department of Agriculture, Economic Research Service., 2024b), about 56% of U.S. children still consume diets characterized by poor nutritional quality (Liu et al., 2020). These statistics underscore a critical limitation of food security when assessed solely in terms of caloric sufficiency (Nguyen et al., 2014; Niles et al., 2020; Hanson and Connor, 2014). Despite the inclusion of the word "healthy" in the official definition of food security, its practical application in policy often overlooks the nutritional dimension. For instance, health screening tools commonly used to assess food insecurity do not evaluate nutritional quality (Mozaffarian et al., 2021). The standard 18-question USDA Household Food Security assessment from 2000 lacks nutritional content evaluation, with only two brief references to "balanced meal". Despite being the primary federal nutrition assistance program, the Supplemental Nutrition Assistance Program (SNAP) allocates minimal resources toward nutritional improvement. SNAP Education, the nutrition education and obesity prevention component of SNAP, aims to improve the likelihood that SNAP-eligible individuals make nutritious food choices within a limited budget and adopt physically active lifestyles aligned with the current Dietary Guidelines for Americans and USDA food guidance. However, it reaches only about 15% of SNAP participants (Gleason et al., 2018), leaving most recipients without direct support for improving their diet. These challenges highlight the need for a broader framework that addresses not just whether people have enough to eat, but whether they are eating well.

In response to the recognition that food access alone does not ensure good health, the concept of nutrition security has emerged as a critical complement to food security. While food security efforts focus on the availability and quantity of food, nutrition security expands this scope to include the quality and nutrient adequacy of the diet (Pinstrup-Andersen, 2009; Ingram, 2020). It emphasizes consistent access to the essential vitamins, minerals, and other nutrients required for healthy growth and disease prevention. Given the rise in diet-related illnesses and persistent racial disparities in access to nutritious food, the concept of nutrition security has become an increasingly important focus in public health (Mozaffarian et al., 2021). Although closely related, nutrition security is conceptually distinct from food security (Seligman et al., 2023). Some researchers argue that food and nutrition security are inherently interdependent, suggesting that addressing one without the other is ultimately unsustainable (Hwalla et al., 2016). Nutrition security expands upon the concept of food security by highlighting the coexistence of food insecurity with diet-related diseases and nutritional disparities (U.S. Department of Agriculture, Food and Nutrition Service., 2022). The failure to address the nutrition inequalities between food-secure and food-insecure households may exacerbate existing health inequities, particularly among lower socioeconomic status populations (Hanson and Connor, 2014; Crews et al., 2015).

Previous studies provide insights into the factors associated with the nutritional inequalities between food-secure and food-insecure households. Hanson and Connor (2014) highlights that lower education levels and income, key indicators of food insecurity, are linked to poorer diet quality among both adults and children, suggesting that educational and economic status influence dietary choices across different household types. This is further compounded by employment status. Unemployed adults are more likely to experience food insecurity and poorer dietary quality compared to those who are employed (Hanson and Connor, 2014). Leung and Tester (2019) expands this demographic analysis by examining differences within racial and ethnic groups. The study finds that among groups such as Non-Hispanic Whites and Asians, food-secure individuals exhibit better nutritional quality than their food-insecure counterparts. Health behaviors are also related to these disparities. Smoking, which is more common among low-income, food-insecure households (Armour et al., 2007), is associated with poorer diet quality regardless of other socioeconomic, lifestyle, or biological factors (Alkerwi et al., 2017). Finally, Silva et al. (2016) emphasize the role of food expenditures. Food-secure households not only spend more on food overall but also allocate resources toward a wider variety of nutritious options, resulting in better dietary outcomes compared to food-insecure households.

My study aims to empirically examine whether and how nutrition security provides information distinct from food security by analyzing nutritional inequalities between food-secure and food-insecure households. Unlike previous research focused on mean differences (Hastings et al., 2021; Nguyen et al., 2014; Niles et al., 2020; Hanson and Connor, 2014; Mancino and Gregory, 2020), I investigate how the gap in nutritional quality across food security status varies across the distribution of nutritional quality. This distributional approach is particularly important for policy design as mean differences can mask two critical issues. First, the presence of a few households with very high nutritional quality can raise the average in a way that does not reflect the typical experience of most households. Second, households most vulnerable to diet-related disease are often those with the lowest nutritional quality, and mean values may fail to reflect their conditions accurately. By examining disparities in the nutritional quality of food acquired between food-secure and food-insecure households across multiple quantiles of the diet quality distribution, we can identify which segments of the population experience the widest nutrition gap related to food insecurity and determine which households are most disadvantaged in terms of nutritional outcomes. This approach also directly engages with the concept of nutrition security. If food security were sufficient to ensure good nutrition, we would expect minimal differences in diet quality among households with similar characteristics, such as income, eduction and household composition. Conversely, finding large disparities in diet quality, especially after accounting for these factors, would indicate that nutrition security provides distinct information beyond food security status alone.

To investigate these issues, I use data from the public access USDA's National Household Food Acquisition and Purchase Survey (FoodAPS). FoodAPS is suited for this analysis because it provides detailed information on household food acquisitions over a one-week period, including both food at home (FAH) and food away from home (FAFH). Unlike earlier studies that rely on 24-hour dietary recall data to assess individual consumption (Leung and Tester, 2019), FoodAPS captures household food acquisition behavior at the household level over an entire week. This provides a more accurate reflection of access in everyday circumstances. Because food security is fundamentally about having reliable access to adequate food, acquisition data offer a more direct measure of household food security than short-term dietary intake alone. To nutrition security, I calculate the Healthy Eating Index (HEI)-2010 score using the FAH and FAFH acquisition data, a widely accepted and validated measure of dietary quality (Fulgoni et al., 2009; Guenther et al., 2014). In my analysis, food-secure households have an average HEI-2010 score about 4.5 points higher than food-insecure households, aligning with findings from previous research. For instance, when focusing on FAH acquisitions, low-income food-secure households scored about 4 points higher on the HEI-2010 than comparable food-insecure households (Mancino and Gregory, 2020). Similarly, Leung and Tester (2019) observed that food-insecure adults scored 2.2 points lower on HEI-2015 compared to food-secure adults, on average. Such mean differences highlight that food-insecure individuals generally eat less healthy diets. More importantly, my findings show that the gap in nutritional quality between food-secure and food-insecure households is not uniform across the distribution. Among households with lower nutritional quality, the difference is relatively small, but it becomes larger at higher levels of nutritional quality. This study quantifies that variation and explores the underlying factors associated with it.

My analysis uses an unconditional quantile decomposition method to understand what drives the nutrition gap at different points in the nutritional quality distribution. This approach, based on Recentered Influence Function (RIF) regression (Firpo et al., 2009), lets us break down the diet quality gap at each quantile into two parts. The first part is the composition effect, which reflects the portion of the nutritional quality gap attributable to differences in household characteristics, such as income, education, and food spending, between food-secure and food-insecure groups. The second part is the coefficient effect, which captures the portion of the nutritional quality gap due to differences in the returns or effects of those characteristics on nutritional quality. The results show that differences in household characteristics explain most of the disparity for households with lower nutritional quality, whereas both differences in household characteristics and differences in how they are associated with diet quality contribute more evenly to the nutritional quality gap among households with higher nutritional quality. Specifically, differences in characteristics account for 98% of the gap among households with lower nutritional quality, 83% for households at median levels of nutritional quality, and 58% for households with higher nutritional quality. For those making less healthy purchases, demographic differences account for about 52% of the disparity, while differences in health behaviors and food expenditures explain 27% and 18%, respectively. Among demographic factors, education level stands out as a particularly strong contributor to the disparity among nutritionally vulnerable households. This suggests that efforts to improve diet quality may need to address underlying disparities in education as part of a long-term strategy to reduce nutritional inequality. This analysis is not only useful for understanding disparities in dietary outcomes, but also provides guidance for designing targeted interventions. By identifying which factors matter most at different points in the distribution, it supports more tailored strategies to address both food insecurity and nutrition insecurity more effectively.

The rest of this paper is organized as follows: Section 2.2 describes the public access FoodAPS dataset used in the analysis. Section 2.3 introduces the empirical model and methodology applied to analyze the factors contributing to nutritional disparities. Section 2.4 discusses the results, and Section 2.5 concludes with policy implications.

2.2 Data

2.2.1 FoodAPS Data

FoodAPS is a nationally representative dataset that captures comprehensive information on food purchases and acquisitions made by 4,826 households over a 7-day survey period, from April 2012 to January 2013. It provides detailed data on food acquisitions for FAH and FAFH consumption, offering insights into various food sources, including those obtained through food and nutrition assistance programs. The survey's primary respondent is the main food shopper or meal planner of the household, typically responsible for most food acquisitions. Each primary respondent participated in two in-person interviews, an initial and a final one, as well as up to three telephone interviews, which gathered both household- and individual-level characteristics. primary respondents reported food purchases and acquisitions for all household members in food diaries. For packaged FAH items, respondents used barcode scanners to record products, while non-packaged items, like fresh produce, were logged with generic codes, weight, quantity, and cost based on store receipts. For FAFH acquisitions, receipts from restaurants and stores served as the primary data source. FAH includes food acquired from sources such as supermarkets, farmers' markets, home gardens, and food pantries. FAFH encompasses meals and snacks purchased from locations like restaurants, fast-food outlets, and entertainment venues. FoodAPS uniquely captures household food acquisitions from these distinct angles, allowing for an in-depth analysis of eating habits both at home and away. Additionally, each food item in FoodAPS was assigned a USDA nutrient food code, linked to the Nutrient Database and the Food Patterns Equivalents Database (FPED). FPED provides data in either cup or ounce equivalents for specific food groups, such as fruits and whole grains, enabling the calculation of the HEI score for each household based on their FAH and FAFH acquisitions.

FoodAPS also provides detailed household-level food expenditure data for both FAH and FAFH. Solely focusing on dietary intake misses the broader nutrition picture, since food purchases made for the whole household influence overall nutrition (French et al., 2009). Household budget constraints directly determine what foods can be purchased, making expenditure data particularly relevant for understanding food choice behaviors. Furthermore, the survey's week-by-week longitudinal data collection offers a more detailed view of short-term expenditure fluctuations compared to the limited scope of typical 24-hour dietary recalls (Nguyen and Powell, 2014; French et al., 2009).

In my analysis, I restricted the sample to households with primary respondents aged 18 years or older. This decision was made to focus on adult-headed households, as these individuals are typically responsible for major household food acquisition and meal planning decisions. This selection reduced the sample size from 4,826 to 4,721 households.

2.2.2 Healthy Eating Index

In this study, I use the Healthy Eating Index (HEI) as a measure of nutritional quality, a well-established metric frequently used by researchers (Guenther et al., 2014). The HEI assesses diet quality by evaluating how well a set of foods aligns with the recommendations of the Dietary Guidelines for Americans (DGA). I specifically use the Healthy Eating Index-2010 (HEI-2010) (Guenther et al., 2014), as it is suitable for FoodAPS data, which was collected between 2012 and 2013. The HEI-2010 consists of 12 components: 9 adequacy components (total fruit, whole fruit, total vegetables, greens and beans, whole grains, dairy, total protein foods, seafood and plant proteins, and fatty acids) and 3 moderation components (refined grains, sodium, and empty calories) (Guenther et al., 2013). The adequacy components, except fatty acids, are scored based on nutrient density per 1,000 calories, while fatty acids are scored by the ratio of polyunsaturated and monounsaturated fats to saturated fats (Guenther et al., 2013; Berube et al., 2017). The moderation components are also scored based on nutrient density per 1,000 calories, with empty calories being scored by their proportion of total energy intake. Higher scores for adequacy components and lower scores for moderation components reflect better nutritional intake. The overall HEI score, ranging from 1 to 100, aggregates these components to provide a measure of diet quality. A higher HEI score indicates a healthier diet, with scores above 80 denoting good nutritional quality, scores between 51 and 80 suggesting room for improvement, and scores below 50 indicating poor nutritional quality (Beatty et al., 2014; Berube et al., 2017).

The analysis includes both FAH and FAFH acquisitions to present a comprehensive picture of household food acquisition patterns. It is important to note that these HEI-2010 scores are calculated at the household level, reflecting the nutritional quality of acquired food rather than consumed food. Since the HEI-2010 scores are not directly available in FoodAPS dataset, I calculated them for each household using their FAH and FAFH acquisitions over a 7-day period based on the HEI-2010 components and scoring criteria, as detailed in Table A1. Table 2.1 presents the mean and standard deviation values for the HEI-2010 scores of food-secure and food-insecure households. The mean HEI-2010 score for food-secure households is 52.926, while for food-insecure households, the mean score is 48.410. The difference in mean HEI-2010 scores between the two groups is statistically significant, with food-secure households scoring an average of 4.516 points higher than food-insecure households. This finding suggests that, on average, food-secure households have diets of higher nutritional quality compared to their food-insecure counterparts, which is consistent with the results reported in prior studies examining the relationship between food security status and dietary quality(Taylor et al., 2017; Leung and Tester, 2019; Mancino and Gregory, 2020).

Figure 2.1a displays the kernel density distribution of HEI-2010 scores by food security status [1]. Both distributions show a clear peak and follow a bell-shaped pattern, indicating that most households fall within a central range of scores. However, notable differences emerge between the two groups. A larger proportion of food-insecure households score between 30 and 50, while food-secure households more frequently fall in the 45 to 60 range. The rightward shift in the food-secure distribution suggests that these households generally have higher HEI-2010 scores on average, a result consistent with the data presented in Table 2.1.

1. While the density of HEI-2010 scores is lower in the tails, the unconditional quantile regression (UQR) approach defines each quantile using the same number of observations. For instance, the 10th and 90th percentiles are both calculated using 10% of the full sample, just as the median is calculated from the middle 10%. This ensures that the quantile estimates at the tails are based on equally sized portions of the data.

2.2.3 Demographics and household characteristics

This study does not aim to identify the causal effect of food security status on nutritional quality, as the objective is not to establish a cause-and-effect relationship. Rather, the aim is to highlight that nutrition security is a distinct concept from food security. The primary focus is to explore the factors that contribute to the observed disparities in nutritional quality between food-secure and food-insecure households, and to assess how these disparities differ across the distribution of nutritional quality. By incorporating a comprehensive set of controls for household demographics, food expenditures, and health behaviors, I may mitigate potential biases and self-selection issues in the estimated relationship between food security status and nutritional quality. These controls include: gender, age, education, race/ethnicity, marital status, employment status, household income, smoking status, expenditures on FAH per person, expenditures on FAFH per person, and number of children.[2] The variable *Income* represents the average monthly household income, calculated by aggregating the income per member and standardizing it in units of $1,000. I also compute *FAH expenditures per person* and *FAFH expenditures per person* by summing each household's weekly expenditures on FAH and FAFH, respectively, and dividing by the number of household members. To control for fixed effects, I include regional information in the model. Table 2.1 provides summary statistics for the sample. Food-secure households are more likely to have primary respondents who are male and older. These households also tend to have respondents with a college degree or higher, are more often married, employed, and report higher income levels compared to food-insecure households. In contrast, food-insecure households are more likely to have primary respondents who are divorced, smoke, and have children. The income disparity between the two groups is substantial, with food-secure households earning an average monthly income of $4,851, which is approximately twice that of food-insecure households, who earn an average of $2,613. This income gap is mirrored in their respective expenditures on FAH and

2. To assess potential multicollinearity among the independent variables, I calculate the variance inflation factor (VIF) for these covariates. All VIF values are below the threshold of 5, indicating that multicollinearity does not pose a significant concern in my model, as shown in Appendix Table A2.

FAFH, with food-secure households generally outspending their food-insecure counterparts in both categories.

Figure 2.2 provides a detailed view of the relationship between household food security status, food expenditure patterns, and Healthy Eating Index-2010 (HEI-2010) scores. The figure shows that expenditures on FAH increase with higher HEI-2010 quantiles for both food-secure and food-insecure households. The increase is more substantial among food-secure households, whose FAH expenditures rise steadily across the distribution. In contrast, food-insecure households also tend to increase FAH spending as HEI-2010 scores rise, but the magnitude of the increase is relatively smaller. The gap in FAH expenditures between the two groups becomes wider in the higher HEI quantiles, suggesting that food-insecure households may face growing financial barriers to achieving healthier diets. For FAFH, spending levels are consistently lower than FAH for both household types. Food-secure households spend more on FAFH across all HEI-2010 quantiles. The gap between the two groups remains relatively stable across the distribution, with a noticeable widening between the 50th and 60th quantiles. FAFH expenditures show a modest increase among food-secure households as nutritional quality improves, whereas FAFH spending among food-insecure households does not exhibit a clear upward trend and instead varies slightly across the HEI-2010 distribution. These findings suggest that higher HEI-2010 scores are associated with increased FAH expenditures, especially among food-secure households, reinforcing the idea that healthier eating patterns often require greater investment in groceries. The widening FAH spending gap between food-secure and food-insecure households at higher HEI quantiles highlights the potential role of financial constraints in limiting access to healthier diets. In contrast, FAFH expenditures appear less closely related to HEI-2010 scores, indicating that meals consumed away from home may contribute less directly to improvements in dietary quality.

2.2.4 Stochastic Dominance

Table 2.1 reveals that, on average, the HEI-2010 score for food-secure households is higher than that of food-insecure households, aligning with findings from previous research (Leung and

Tester, 2019; Gregory et al., 2019). To further compare the entire distributions of nutritional quality between these two groups, I take a comprehensive approach involving two critical steps. First, I ascertain if the observed differences in the HEI score distributions are statistically significant. Second, I explore whether there is any overlap between the two nutritional quality distributions.

To accomplish this, let *F(x)* and *G(x)* denote the cumulative distribution functions (cdfs) of the HEI scores for food-secure and food-insecure households, respectively. If $D_1(x) \equiv G(x) - F(x) \geq 0$ for all x, this implies that *F(x)* first-order stochastically dominates *G(x)*, signifying that *F(x)* is consistently to the right of *G(x)* without any overlap. Conversely, if $D_2(x) \equiv \int (G(x) - F(x))dx \geq 0$ for all x, then *F(x)* demonstrates second-order stochastic dominance over *G(x)*. In this case, the distributions may overlap, but the area under *G(x)* to the left of any given x is always greater than that under *F(x)*.

To asses the statistical significance of these distributional differences, I conduct a two-sample Kolmogorov-Smirnov test (Massey, 1951). The null hypothesis assumes that the two distributions are identical, i.e., *F(x)* = *G(x)* for all x. As illustrated in Figure 2.1b, the empirical cumulative distribution functions (CDFs) for the two groups exhibit visible differences. The test results reject the null hypothesis at the 5% significance level, providing statistical evidence that the distributions of HEI-2010 scores for food-secure and food-insecure households are significantly different. These findings suggest that the disparity between food-secure and food-insecure HEI-2010 scores is not uniform across the distributions. Rather, the gap in nutritional quality varies at different levels of the HEI-2010 score distribution. This highlights the importance of identifying the key factors contributing to these differences in nutritional quality across different levels of food security status.

2.3 Methods

2.3.1 Basic model

This study pursues three central objectives to deepen the understanding of nutritional quality disparities between food-secure and food-insecure households. First, I seek to identify the primary factors associated with variations in nutritional quality between these groups. Second, I assess whether the disparity in nutritional quality remains consistent or varies across different levels of nutritional quality. This approach allows me to determine if inequality is uniformly distributed or if it intensifies or diminishes at certain levels of nutritional quality. Finally, I aim to evaluate whether increasing disparities in nutritional quality are associated with widening socio-economic and demographic differences among households or changes in how these factors influence nutritional outcomes over time.

To conceptually frame the nutritional disparities between food-secure and food-insecure households, this study employs a reduced form household production function model, which builds upon the theoretical foundation introduced by Grossman (1972a,b). Grossman's model emphasizes how households combine market goods and their time to generate health outcomes. Extending this approach, I propose that households utilize available resources, including financial resources, time, and knowledge, to produce nutritional quality through their food acquisition behaviors

The reduced form household production function is defined as:

$$H_i = f(X_i, F_i, E_i, T_i, R_i) \tag{1}$$

where H_i denotes the nutritional quality outcome for household i, as measured by the HEI-2010. X_i includes household demographic characteristics such as gender, age, education, race/ethnicity, marital status, smoking status, and the number of children (Allcott et al., 2019; Leung and Tester, 2019; Smith et al., 2019; Hiza et al., 2013; Bouis et al., 2011). These factors reflect human capital, influencing nutritional knowledge and dietary decisions. F_i represents food expenditure per person on FAH and FAFH. E_i indicates household economic resources, primarily household income, which constrain food purchasing behaviors and dietary quality. T_i captures time resources available for activities related to food preparation and shopping, represented by employment status. R_i incorporates regional fixed effects, controlling for geographic disparities in food availability, pricing and dietary patterns. Time fixed effects are not incorporated in my model, as FoodAPS data was collected over a relatively short time span (one week per household), resulting in limited temporal variation that would significantly influence the results.

Households aim to achieve optimal nutritional quality under these constraints, leading to a reduced form equation:

$$H_{iFS} = \alpha_{FS} + \beta_{FS} X_{iFS} + \theta_{FS} F_{iFS} + \delta_{FS} E_{iFS} + \eta_{FS} T_{iFS} + \gamma_{FS} R_{iFS} \quad \text{for } FS \in \{0,1\} \tag{2}$$

where FS indicates food security status. Specifically, FS = 1 indicates food-secure households, while FS = 0 denotes food-insecure households. The coefficients α_{FS}, β_{FS}, θ_{FS}, δ_{FS}, η_{FS} and γ_{FS} quantify the marginal impacts of each resource and characteristic on nutritional quality outcome across different food security statuses.

2.3.2 Unconditional quantile regression

To investigate the variations in the relationships presented in Equation 2 across the distribution of nutritional quality, I employ the Unconditional Quantile Regression (UQR) estimation method, utilizing Recentered Influence Functions (RIFs) as developed by Firpo et al. (2009). The selection of UQR is motivated by two key reasons. First, UQR is particularly effective in assessing distributional effects, especially at the extremes of the distribution, which are often overlooked by methods that focus solely on average outcomes. Second, UQR is advantageous when comparing two distributions, as it avoids the limitations of Conditional Quantile Regression (CQR), which depends heavily on strong assumptions about covariates and their distribution. In contrast, UQR provides a more robust approach to examining disparities across distributions without these constraints.

The RIF of the τth quantile is defined by adding the statistics q_τ to the influence function (IF) (Firpo et al., 2009) as follows:

$$RIF(h_i, q_\tau) = q_\tau + IF(h_i, q_\tau) = q_\tau + \frac{\tau - I\{h_i \leq q_\tau\}}{f_h(q_\tau)} = c_{1,\tau} \cdot I\{h_i < q_\tau\} + c_{2,\tau} \tag{3}$$

where $c_{1,\tau} = \frac{1}{f_Y(q_\tau)}$ and $c_{2,\tau} = q_\tau - c_{1,\tau} \cdot (1-\tau)$. q_τ denotes the $\tau$the quantile of the nutritional quality distribution; I is the indicator function that equals one if $h_i \leq q_\tau$ and zero otherwise; $f_h(q_\tau)$ is the kernel density of nutritional quality evaluated at q_τ. Therefore, the conditional expectation of the RIF can be written as:

$$\mathbb{E}[RIF(h_i;q_\tau)|X=x] = c_{1,\tau} \cdot \mathbb{E}[I\{H > q_\tau\}|X=x] + c_{2,\tau} = c_{1,\tau} \cdot \Pr[H > q_\tau|X=x] + c_{2,\tau} \tag{4}$$

I specify the RIF regression to estimate Equation 2 as:

$$\mathbb{E}[RIF\{h_i,q_\tau\}|Z_{FS}] = \alpha_{FS} + \beta_{FS}X_{iFS} + \theta_{FS}F_{iFS} + \delta_{FS}E_{iFS} + \eta_{FS}T_{iFS} + \gamma_{FS}R_{iFS} \quad \text{for } FS \in \{0,1\} \tag{5}$$

For simplicity, I group X_{iFS}, F_{iFS}, E_{iFS}, T_{iFS} and R_{iFS} as elements of Z_{FS}. Following Firpo et al. (2009), the coefficient estimates from the RIF regression can be interpreted as the effect of a one-unit increase in the respective independent variable on the unconditional distributional statistic (e.g., the τ-th quantile).

2.3.3 Recentered influence function (RIF) decomposition

The Oaxaca-Blinder decomposition is a widely used econometric technique for analyzing differences in outcome means between two groups (Rios-Avila, 2020). This method allows researchers to decompose the gap in the mean outcome into two distinct components: a composition effect, which is associated with differences in characteristics or covariates between the groups, and a coefficient effect, which is linked to differences in the returns to those characteristics. While the Oaxaca-Blinder decomposition has been instrumental in understanding disparities in various fields, it has its limitations. The primary focus of this method is on average differences, which may not adequately capture the variations across different parts of the distribution. In other words, the Oaxaca-Blinder decomposition does not take into account how differences in characteristics and coefficients might vary across the range of the outcome distribution, such as at the lower or upper ends of nutritional quality distribution. To address this limitation and gain a more comprehensive understanding of the disparities, particularly in contexts like food insecurity and nutritional quality, researchers can employ an alternative approach called the RIF decomposition. This method, developed by Firpo et al. (2018), allows for a more thorough analysis of distributional statistics beyond the mean. The RIF decomposition extends the Oaxaca-Blinder framework by providing a way to decompose differences in various distributional statistics, such as quantiles, variance, and Gini coefficient.

The RIF decomposition is a two-stage process. In the first stage, the gap in outcomes between two groups is decomposed into coefficient and composition components using a reweighting approach. This reweighting procedure involves constructing a counterfactual distribution that combines the characteristics of one group with the returns to those characteristics of the other group. By comparing the actual and counterfactual distributions, researchers can estimate the relative contributions of differences in characteristics and differences in returns to the overall gap in outcomes. In the second stage of the RIF decomposition, the coefficient and composition effects are further broken down to understand the contribution of each individual covariate. This detailed decomposition provides valuable insights into how specific characteristics contribute to the overall gap in outcomes between the two groups.

To examine the nutritional inequality between food-secure and food-insecure households, the gap in the distributional statistic q_τ can be determined using the cumulative conditional distribution of HEI-2010 scores:

$$\Delta q_\tau = q_{\tau 1} - q_{\tau 0} = q_\tau(F_{H|FS=1}) - q_\tau(F_{H|FS=0}) \tag{6}$$

$$\Delta q_\tau = q_\tau\left(\int F_{H|Z,FS=1}\, dF_{Z|FS=1}\right) - q_\tau\left(\int F_{H|Z,FS=0}\, dF_{Z|FS=0}\right) \tag{7}$$

where $F_{H|FS=k}$ is the cumulative distribution of HEI-2010 scores conditional on the categorical variable FS.

The RIF regression can be estimated for each group:

$$q_{\tau 1} = \mathbb{E}\left[RIF\{h_i, q_\tau(F_{H|FS=1})\}\right] = \bar{Z}^{1'}\hat{\beta}_1 \tag{8}$$

$$q_{\tau 0} = \mathbb{E}\left[RIF\{h_i, q_\tau(F_{H|FS=0})\}\right] = \bar{Z}^{0'}\hat{\beta}_0 \tag{9}$$

To evaluate the association between the nutritional quality gap and the differences in household demographics and socio-economic factors, as well as the variances in returns derived from these factors, it is essential to employ a counterfactual framework. The counterfactual approach introduces a scenario in which the structural nutritional quality of food-insecure households remains stable while the distribution of their characteristics is conceptualized based on the food-secure households. This hypothetical distribution is not directly observable within the dataset. However, it can be approximated by multiplying a reweighting factor $w(Z)$ with the food-insecure households' observed characteristics distribution.

The reweighting factor $w(Z)$ is defined as:

$$w(Z) = \frac{1-P}{P}\frac{P(FS=1|Z))}{1-P(FS=1|Z)}, \tag{10}$$

where P is the probability of being in the food-secure group (FS = 1), and $1-P$ is the probability of belonging to the food-insecure group (FS = 0). The term $P($ FS = $1|Z)$ represents the conditonal probability of households having characteristics Z being in the food-secure group, calculated using a logit model. Coversely, $1-P(FS=1|Z)$ is the conditional probability for the food-insecure group, also derived from a logit model. Given the reweighting factor $w(Z)$, the counterfactual scenario can be defined as follows:

$$q_{\tau c} = q_\tau(F_H^C) = q_\tau\left(\int F_{H|Z,FS=0}\, dF_{Z|FS=1}\right) \cong q_\tau\left(\int F_{H|Z,FS=0}\, dF_{Z|FS=0} w(Z)\right) \tag{11}$$

where c is the notation for counterfactual. By fitting RIF regression with weighted least squares, I can obtain the counterfactual statistic as:

$$q_{\tau c} = \mathbb{E}\left[RIF\{h_i, q_\tau(F_H^C)\}\right] = \bar{Z}^{c'}\hat{\beta}_c \tag{12}$$

The nutritional quality gap between food-secure and food-insecure households can be written as:

$$\Delta q_\tau = q_{\tau 1} - q_{\tau 0} = q_{\tau 1} - q_{\tau c} + q_{\tau c} - q_{\tau 0} \tag{13}$$

$$\Delta q_\tau = \underbrace{\bar{Z}^{1'}(\hat{\beta}_1 - \hat{\beta}_c)}_{CoefficientEffect} + \underbrace{(\bar{Z}^1 - \bar{Z}^c)'\hat{\beta}_c}_{ReweightingError} + \underbrace{(\bar{Z}^c - \bar{Z}^0)'\hat{\beta}_0}_{CompositionEffect} + \underbrace{\bar{Z}^{c'}(\hat{\beta}_c - \hat{\beta}_0)}_{SepcificationError} \tag{14}$$

The composition effect refers to the differences in nutritional quality that can be attributed to the differences in the characteristics or attributes of the groups being compared. It explains how the unique makeup of each group contributes to the overall differences in nutritional quality. On the other hand, the coefficient effect refers to the differences in nutritional quality that arise due to how the characteristics of each group are valued or rewarded in terms of their impact on nutritional quality. In other words, it captures how the returns to certain attributes differ between the groups. These two effects collectively provide insights into the factors driving the nutritional quality inequality observed between different food security demographics. By decomposing the gap into the composition and coefficient effects, researchers can better understand the relative importance of differences in household characteristics and differences in the returns to these characteristics in shaping the observed disparities.

The reweighting error serves as an analytical tool to verify the precision of the reweighting approach used in the counterfactual analysis. In an ideal scenario, especially with large datasets, this error is expected to be marginal, nearing zero. If the reweighting error is statistically different from zero, it raises concerns about the efficacy of the reweighting procedure and implies that the employed reweighting strategy may not have been accurately identified or applied. The specification error is crucial for evaluating the integrity and effectiveness of model specifications and the RIF approximation. This error measurement is key in determining the precision of RIF regressions. When the specification error is notably large and significant, it raises concerns about the model's formulation. It could indicate that the RIF regression has been misspecified or that the RIF is inadequately approximating the distributional statistic of interest.

2.4 Results

To better understand how household characteristics contribute to variations in dietary quality across the distribution, I examine the UQR estimates separately for food-insecure and food-secure households in Appendix Tables A3 and A4, respectively. These tables include both ordinary least squares (OLS) estimates and UQR estimates at various quantiles of the HEI-2010 distribution. This approach allows us to capture heterogeneity in the associations between covariates and dietary quality that would be obscured by focusing solely on mean outcomes.

Several patterns across the distribution underscore the value of examining heterogeneous relationships between household characteristics and dietary quality. First, the estimated impact of educational attainment on dietary quality increases markedly across quantiles for both groups. For example, among food-insecure households (Appendix Table A3), having a college diploma is associated with a 2.76-point increase in HEI at the 15th quantile and a 6.72-point increase at the 85th quantile. Similarly, for food-secure households (Appendix Table A4), the coefficient on college diploma grows from 3.52 at Q15 to 8.08 at Q75. These results suggest that the link between education and dietary quality becomes stronger among households with higher HEI-2010 scores. Smoking status is also consistently associated with lower dietary quality. In both food-insecure and food-secure households, smoking is linked to significantly lower HEI-2010 scores across all quantiles, with estimated effects ranging from approximately -3.63 to -5.50 and from -3.38 to -5.88, respectively. The relatively stable coefficients suggest that smoking correlates with poorer dietary outcomes regardless of where households fall in the distribution, potentially offsetting the influence of more favorable attributes. Other covariates, such as marital status, show greater variability. For food-secure households, being married is positively and significantly related to HEI-2010 scores across all quantiles (e.g., approximately 2.79 to 4.09 points). In contrast, for food-insecure households, the estimated effects of marriage are generally small and statistically insignificant. This may reflect that the advantages associated with marital status are more fully realized when households also have access to stronger economic and social supports. In addition, FAH expenditures per person are positively and significantly associated with HEI-2010 scores across the distribution for both groups. The coefficients tend to be slightly larger at lower quantiles for food-insecure households, indicating that increased spending on food at home may have a stronger marginal benefit for households with initially lower dietary quality. These findings highlight that the effects of household characteristics on dietary quality vary across the HEI-2010 distribution. The heterogeneity observed across quantiles illustrates the usefulness of unconditional quantile regression in identifying distribution-specific patterns of association, particularly for subpopulations with very low or very high dietary quality. This motivates the subsequent RIF decomposition analysis, which explores how differences in characteristics and their associated effects contribute to the observed nutrition gap across the distribution.

2.4.1 RIF Decompositions

Table 2.2 presents the estimated differences in dietary quality between food-secure and food-insecure households, focusing on the role of composition and coefficient effects across different quantiles. The *Total difference* represents the gap in HEI-2010 scores between food-secure and food-insecure households, demonstrating a widening nutritional disparity as I move from lower to higher quantiles. In other words, food-secure households not only maintain healthier diets overall, but their nutritional advantage over food-insecure households grows larger among households with better diet quality. Specifically, at the 15th quantile (Q15), the difference in HEI-2010 scores is 2.827, increasing to 4.844 at the median (Q50), and reaching 6.441 at the 90th quantile (Q85). These differences are statistically significant at the 1% level across all quantiles, indicating a consistent gap in dietary quality between food-secure and food-insecure households.

Table 2.2 and Figure 2.3 show that at lower quantiles, the composition effect, which reflects differences in household characteristics, accounts for most of the total nutritional gap. A positive composition effect means that if food-insecure households had the same distribution of characteristics as food-secure households, their HEI-2010 score would improve by the estimated amount. At higher quantiles, the coefficient effect, which reflects differences in how household characteristics translate into dietary outcomes, accounts for a greater proportion of the total nutritional gap than at lower quantiles. A positive coefficient effect indicates that the returns on characteristics are greater for food-secure households, thus contributing to the widening nutritional quality gap. Conversely, a negative coefficient effect would imply that food-insecure households gain more nutritional benefits from these characteristics compared to food-secure households.

The composition effect follows an inverse U-shaped pattern across the distribution. It increases initially, then remains relatively stable, ranging from 3.92 to 3.71 points between the 40th and 85th percentiles, and eventually begins to decline beyond that range. The composition effect estimates in the upper half of the HEI-2010 distribution seem to be statistically the same. The gap between the composition and coefficient effects is largest at the 40th percentile, where it reaches 3.2 HEI-2010 points. However, from the 80th to the 95th percentiles, where households have the highest diet quality, the composition and coefficient effects are not statistically different. This indicates that for food-secure and food-insecure households with the best diets, both types of effects contribute equally to the nutritional quality gap.

Simply examining the magnitudes of the composition and coefficient effects does not fully capture how they contribute to the gap in HEI-2010 between the two groups, because the gap varies across quantiles. To provide a clearer picture, I assess the relative contributions to the nutritional quality gap by expressing the composition and coefficient effects as proportions of the total difference in HEI-2010 at each quantile. At the lower end of the distribution (Q15), the composition effect is 2.608 and explains almost all (92%) of the total difference in HEI-2010, which is 2.827. This suggests that when households have less healthy diets, the gap in nutritional quality between food-secure and food-insecure households is mainly due to differences in household characteristics. As I move to households with healthier diets, the role of these characteristics diminishes. At the median (Q50), the composition effect explains most (80%) of the nutritional quality gap, but its importance starts to decline. For households making the healthiest choices (Q85), the composition effect only accounts for 57% of the nutirtional quality gap. Conversely, the coefficient effect, which captures how household characteristics relate to nutritional quality, becomes more important as I look at healthier diets. It explains 28% of the gap at the median, increasing to 55% at Q85. Across most of the distribution, the composition effect accounts for a larger relative contribution to the nutritional quality gap than the coefficient effect and nearly fully explains the disparities among households with lower diet quality. However, the coefficient effect becomes increasingly influential for households with higher levels of nutritional quality, contributing a greater share of the gap at the upper end of the distribution.

2.4.2 Detailed decomposition of the composition effect

Table 2.3 presents the detailed decomposition of the composition effects of various covariates on the nutritional quality gap between food-secure and food-insecure households across different quantiles. The analysis shows that household demographics, health behaviors (smoking), and food expenditures contribute to the gap in nutritional quality, with their impacts varying across the distribution of dietary quality.

Household demographics, particularly educational attainment and marital status, play an important role in contributing to the nutritional quality gap, especially at lower quantiles of the HEI-2010 distribution. The composition effect associated with holding a college degree shows an increasing trend across quantiles, rising from 0.412 points at the 15th quantile to a peak of 0.946 points at the 75th quantile. However, the relative contribution of this effect to the nutritional quality gap, which is measured as the share of the total HEI-2010 difference attributed to differences in college attainment, does not increase consistently across the distribution. At Q15, the college composition effect accounts for 15% of the total gap, reaching a maximum relative contribution of 18% at Q30, after which it declines through Q85. A similar pattern is observed for the composition effect of holding a postgraduate degree. The effect increases from Q15 to Q30, peaking at 0.52 points and explaining up to 16% of the nutritional quality gap at Q30. Beyond this point, the relative contribution of the postgraduate education effect diminishes. These positive composition effects suggest that if food-insecure households had the same distribution of higher education levels as food-secure households, their HEI-2010 scores would improve by the estimated amounts. Marital status also contributes to the nutritional quality gap, particularly at the lower end of the distribution. The estimated composition effect of being married rises from 0.47 points at Q15 to a peak of 0.688 points at Q75, then declines to 0.521 points at Q85. Its relative contribution to the nutritional quality gap is highest at Q20, accounting for 19% of the total difference, and steadily decreases to 8% by Q85. This pattern suggests that differences in marital status matter more for households with lower dietary quality. This positive composition effects imply that if food-insecure households had the same marital status distribution as food-secure households, their HEI-2010 scores would improve. These findings highlight that differences in education and marital status between food-secure and food-insecure households contribute significantly to the nutritional quality gap. Education may influence dietary choices through increased knowledge about nutrition or better access to resources. Being married might provide additional support and shared resources that enhance dietary quality. Therefore, policies aimed at improving educational opportunities and supporting family stability could help reduce nutritional disparities.

Health behaviors, specifically smoking status, contribute to the nutritional quality gap, particularly within the lower quantiles. The composition effect of smoking remains relatively stable at the lower quantiles, then increases in the middle quantiles, reaching a peak of 1.03 points at Q60, and decreases toward Q85. The relative contribution of smoking to the nutritional quality gap is greatest at the lower quantiles, accounting for 28% of the total gap at Q15 and decreasing to 9% by Q85. This indicates that differences in smoking behavior between food-secure and food-insecure households have a larger impact on dietary quality among those with lower HEI-2010 scores. These results suggest smoking may be associated with poorer dietary choices or may divert household resources away from food expenditures, thereby exacerbating nutritional disparities among food-insecure households. The stronger explanatory power in lower quantiles suggests that addressing smoking behavior could be particularly effective in improving nutritional outcomes among households with the poorest dietary quality. The findings indicate that integrated health interventions targeting smoking cessation in food-insecure populations might serve as an effective strategy for reducing the nutritional quality gap, potentially both through direct health improvements and through the reallocation of household resources toward food expenditures.

Per capita FAH expenditure also contributes to the nutritional quality gap, with the composition effect increasing across quantiles. The effect rises from 0.315 points at Q15 to a maximum of 0.743 points at Q85. However, However, the relative contribution of FAH expenditure to the nutritional quality gap follows a different pattern: it increases at the lower end of the distribution, reaching its highest contribution of 18% at Q35, and then gently decreases across the higher quantiles. This indicates that differences in FAH expenditure between food-secure and food-insecure households have a greater impact on the nutritional quality gap among those with lower dietary quality scores. These findings imply that the amount of money dedicated to food plays an important role in achieving better dietary quality. Interventions aimed at increasing access to affordable, nutritious foods or providing financial assistance for food purchases could be especially effective among food-insecure households with lower dietary quality. This may help reduce disparities related to food expenditures and improve overall nutritional outcomes.

2.4.3 Detailed decomposition of the coefficient effect

Table 2.4 presents the detailed decomposition of the coefficient effects of various covariates on the nutritional quality gap between food-secure and food-insecure households across different quantiles. The analysis shows that household demographics, particularly marital status and holding a postgraduate degree, significantly contribute to the gap in nutritional quality, with their impacts varying across the distribution of dietary quality, especially from the middle to the upper quantiles of the nutritional quality distribution.

The coefficient effect of being married is positive and statistically significant from Q30 to Q85.

This positive coefficient effect indicates that being married is associated with a greater return in nutritional quality for food-secure households compared to food-insecure households. The coefficient effect increases across the distribution, reaching its peak at Q85, where being married is associated with an increase of 2.728 points in nutritional quality. The relative contribution of this coefficient effect to the nutritional quality gap is highest at Q40, where it accounts for 32% of the total gap. After Q40, the relative contribution gradually declines through Q70, suggesting that the nutritional advantages linked to marriage are most influential among households with moderate dietary quality. Additionally, the coefficient effect of holding a postgraduate degree is positive and statistically significant from the 45th to the 80th percentile, increasing from 0.134 points at the 45th percentile to 0.258 points at the 80th percentile. This indicates that having a postgraduate degree provides greater nutritional benefits for food-secure households compared to food-insecure households, especially in the middle to upper of the distribution. The relative contribution of this effect to the nutritional quality gap peaks at the 60th percentile, accounting for 5% of the total gap, and then declines to 2% at the 85th percentile. Although the magnitude of the coefficient effect of holding a postgraduate degree is relatively small across the 45th to 80th percentiles, it still contributes to reducing the nutritional quality gap between food-secure and food-insecure households.

2.5 Conclusions

In this study, I examine the impact of demographics, health behaviors, and food expenditures on the gap in nutritional quality across food security status. The research emphasizes analyzing inequalities in nutritional quality not solely at the mean level but across the entire nutrition distribution using RIF decomposition. My findings, illustrated in Figure 2.3, show that the gap in nutritional quality between food-secure and food-insecure households varies across the distribution. At lower quantiles, the difference in HEI-2010 scores is relatively small, about 2.7 points. This nutritional inequality grows to approximately 4.6 points in the middle range and further increases to around 6.2 points at the higher quantiles. These results suggests that maintaining a high-quality diet becomes more challenging for food-insecure households relative to food-secure households. Therefore, in addition to trying to obtain sufficient calories, it may be more challenging for food-insecure households to obtain adequate nutrition as well.

To analyze the inequality in nutritional quality, I decompose the inequality into composition and coefficient effects across the nutritional quality distribution. At the lower end of the distribution, where households mostly choose unhealthy foods, the composition effect explains 98% of the nutritional inequality. This indicates that differences in household characteristics are the primary contributors to the nutritional gap at this level. As I move to the higher end of the nutritional quality distribution, where households tend to make healthier food choices, the relative contribution of the composition effect declines, reaching 58%. Conversely, the coefficient effect, which reflects differences in how household characteristics translate into nutritional outcomes, becomes increasingly significant at higher quantiles. For example, households with similar levels of education may not experience the same improvements in diet quality due to differences in knowledge, preferences, local food environments, or time constraints. This growing importance of the coefficient effect at higher levels of diet quality suggests that it is not just who households are that matters, but also how effectively they are able to convert their resources into healthier food choices.

The findings indicate that the composition effect plays a bigger role for households with lower dietary quality, while the coefficient effect becomes more relevant for households with better diet quality. For those at the lower end of the distribution, disparities in household characteristics are the main drivers of the nutritional gap, which implies that addressing inequalities in access to resources could help reduce this gap. In contrast, for households with higher dietary quality, it is not only about having access to these resources but also about making the best use of them. Enhancing these returns may require improving food-related knowledge and decision-making skills, and fostering supportive environments that enable healthier behaviors. To narrow nutritional disparities across the entire distribution, policy efforts must focus not only on improving resource availability but also on enabling households to utilize those resources more effectively.

These findings have several important implications for food policy. Interventions aimed at reducing nutritional disparities should be tailored to different segments of the population based on their position within the nutritional quality distribution. For households with lower diet quality, policies that improve access to resources, such as education and financial support for food-insecure households, could help close the nutritional gap. This may involve expanding programs like SNAP and providing sufficient support for participants to access nutritious foods. For households with better diet quality, interventions should focus on improving the effectiveness of existing resources. This could include nutrition education programs that encourage healthier choices and initiatives that highlight the benefits of nutritious diets.

However, the study has two limitations. First, the relatively small sample size may limit the generalizability of the findings, as it might not fully represent the broader population. Second, while the analysis may be subject to endogeneity and potential residual confounding, it is important to clarify that the goal of this study is not to establish causal inference, but rather to highlight that nutrition security is a concept distinct from food security and to explore patterns of nutritional inequality across food security status. Although the decomposition accounts for key household characteristics, including education and smoking status, many of these factors are interrelated and may co-occur with unobserved influences such as food literacy or cultural dietary norms. This raises the possibility of residual confounding, meaning that some portion of the observed effects may reflect omitted or imperfectly measured variables. Nonetheless, the RIF decomposition remains a valuable tool for identifying where disparities are most pronounced and which factors contribute most to differences in nutritional outcomes across the distribution. The composition effects capture observed differences in household characteristics that help explain nutritional disparities, while the coefficient effects reveal how these characteristics translate differently into nutritional outcomes across groups. Even if selection into food security status is non-random, these patterns remain informative for identifying where nutritional gaps are largest and which factors matter most at various points in the distribution. Despite these methodological limitations, this study makes significant contributions to understanding nutritional quality inequalities across different food security statuses. By emphasizing how food security and nutrition security are related, rather than implying a one-cause-one-effect relationship, my research provides valuable insights for developing targeted interventions. The findings indicate that reducing disparities in household characteristics is crucial for addressing nutritional inequalities among households with lower nutritional quality, while ensuring effective use of these characteristics becomes increasingly important for improving diet quality among households with higher nutritional quality. These insights can guide policymakers and public health authorities in formulating strategies that meet the nutritional needs of those most at risk, ultimately creating more equitable and health-focused food environments.

Table 2.1: Summary Statistics By Food Security Status among Households

Variable	All	Food-secure households	Food-insecure households	Mean difference	
	(1)	(2)	(3)	(4)	
Healthy Eating Index-2010	50.783	52.926	48.410	4.516	***
	(0.194)	(0.276)	(0.264)		
Male	0.263	0.283	0.240	0.043	***
	(0.006)	(0.009)	(0.009)		
Age	46.305	48.719	43.631	5.087	***
	(0.244)	(0.363)	(0.313)		
White Hispanic	0.113	0.077	0.153	-0.077	***
	(0.005)	(0.005)	(0.008)		
Some college	0.329	0.323	0.335	-0.012	
	(0.007)	(0.009)	(0.010)		
College diploma	0.147	0.202	0.085	0.117	***
	(0.005)	(0.008)	(0.006)		
Post graduate degree	0.068	0.105	0.028	0.078	***
	(0.004)	(0.006)	(0.003)		
Married	0.427	0.501	0.345	0.156	***
	(0.007)	(0.010)	(0.010)		
Divorced	0.183	0.160	0.208	-0.048	***
	(0.006)	(0.007)	(0.009)		
Smoking	0.257	0.181	0.340	-0.158	***
	(0.006)	(0.008)	(0.010)		
Number of children	0.962	0.765	1.180	-0.416	***
	(0.019)	(0.023)	(0.030)		
Employment	0.687	0.705	0.666	0.040	**
	(0.007)	(0.009)	(0.010)		
Monthly household income ($1,000's)	3.789	4.851	2.613	2.238	***
	(0.054)	(0.087)	(0.050)		
FAH expenditures per person	41.336	46.585	35.521	11.064	***
	(0.605)	(0.884)	(0.801)		
FAFH expenditures per person	20.165	24.074	15.836	8.239	***
	(0.432)	(0.663)	(0.524)		
Number of obs	4721	2481	2240		

Note: The unit of observation is the household. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$ indicate significant differences in sample means between food-secure and food-insecure households. Standard deviations are reported in parentheses below the means.

Table 2.2: Estimated Differences in Dietary Quality Between Food-secure and Food-insecure Households

	RIF Decompositions								
Variable	Q15	Q25	Q35	Q45	Q50	Q55	Q65	Q75	Q85
Food-secure HEI-2010	38.453***	42.757***	46.528***	50.626***	52.465***	54.200***	57.986***	62.329***	67.738***
	(0.362)	(0.405)	(0.370)	(0.404)	(0.398)	(0.304)	(0.247)	(0.432)	(0.451)
Food-insecure HEI-2010	35.626***	39.841***	43.020***	46.252***	47.621***	49.373***	52.717***	56.284***	61.297***
	(0.392)	(0.363)	(0.273)	(0.338)	(0.367)	(0.336)	(0.362)	(0.367)	(0.506)
Total difference	2.827***	2.916***	3.508***	4.375***	4.844***	4.827***	5.270***	6.045***	6.441***
	(0.540)	(0.583)	(0.496)	(0.558)	(0.585)	(0.468)	(0.407)	(0.553)	(0.744)
Composition effect	2.608***	2.954***	3.631***	3.853***	3.862***	3.668***	3.827***	3.803***	3.708***
	(0.374)	(0.364)	(0.418)	(0.470)	(0.406)	(0.398)	(0.488)	(0.442)	(0.527)
Coefficient effect	0.358	0.362	0.715	0.864	1.378**	1.467**	1.944***	2.515***	3.583***
	(0.697)	(0.471)	(0.502)	(0.668)	(0.560)	(0.664)	(0.541)	(0.699)	(0.863)
Reweighting error	0.006	-0.030	-0.081	-0.037	-0.023	-0.021	-0.026	-0.128	-0.165
	(0.090)	(0.115)	(0.124)	(0.115)	(0.131)	(0.136)	(0.135)	(0.145)	(0.140)
Specification error	-0.144	-0.369	-0.757	-0.306	-0.373	-0.287	-0.475	-0.146	-0.685
	(0.310)	(0.383)	(0.557)	(0.389)	(0.349)	(0.407)	(0.415)	(0.495)	(0.459)

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors are shown in parentheses.

Table 2.3: Estimates of the Contributions of Covariates to the Composition Effect

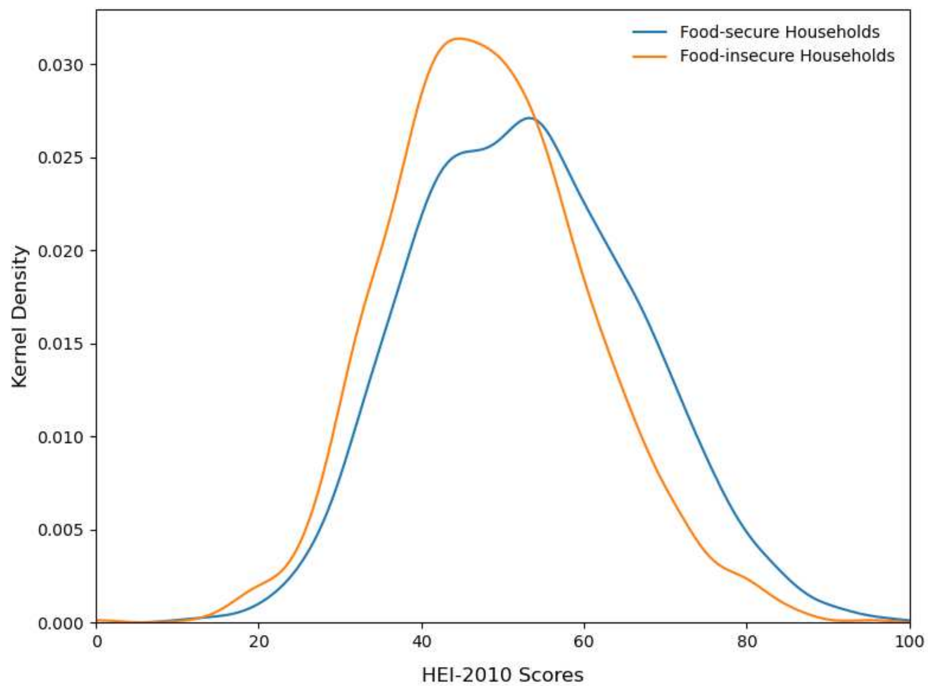
Variable	Q15	Q25	Q35	Q45	Q50	Q55	Q65	Q75	Q85
Male	-0.054	-0.070	-0.031	-0.025	-0.037	-0.007	-0.048	-0.014	-0.052
	(0.048)	(0.051)	(0.052)	(0.056)	(0.046)	(0.044)	(0.051)	(0.050)	(0.056)
Age	0.327**	0.375***	0.450***	0.345**	0.248**	0.139	-0.058	-0.123	-0.005
	(0.162)	(0.144)	(0.169)	(0.159)	(0.118)	(0.146)	(0.156)	(0.164)	(0.169)
White Hispanic	-0.255	-0.173	-0.338	-0.345	-0.310	-0.281	-0.235	-0.253	-0.021
	(0.116)	(0.146)	(0.160)	(0.174)	(0.137)	(0.145)	(0.138)	(0.151)	(0.197)
Some college	-0.011	-0.015	-0.015	-0.027	-0.016	-0.021	-0.017	-0.030	-0.028
	(0.027)	(0.034)	(0.027)	(0.033)	(0.040)	(0.034)	(0.030)	(0.041)	(0.041)
College diploma	0.412***	0.481***	0.598***	0.580***	0.636***	0.606***	0.722***	0.946***	0.784***
	(0.115)	(0.116)	(0.144)	(0.120)	(0.120)	(0.135)	(0.153)	(0.125)	(0.167)
Post graduate degree	0.352***	0.413***	0.513***	0.518***	0.571***	0.496***	0.560***	0.671***	0.615***
	(0.091)	(0.099)	(0.124)	(0.127)	(0.121)	(0.109)	(0.144)	(0.161)	(0.135)
Married	0.470***	0.531***	0.570***	0.574***	0.548***	0.578***	0.627***	0.688***	0.521***
	(0.137)	(0.180)	(0.188)	(0.179)	(0.168)	(0.176)	(0.171)	(0.160)	(0.184)
Divorced	-0.120	-0.050	-0.010	-0.057	-0.097	-0.080	-0.106	-0.178	-0.104
	(0.063)	(0.059)	(0.062)	(0.048)	(0.061)	(0.073)	(0.068)	(0.064)	(0.066)
Smoking	0.796***	0.768***	0.736***	0.846***	0.963***	0.960***	0.912***	0.659***	0.553***
	(0.191)	(0.176)	(0.165)	(0.223)	(0.173)	(0.151)	(0.159)	(0.155)	(0.165)
Number of children	-0.083	0.040	0.126	0.158	0.315	0.296	0.552***	0.598***	0.600***
	(0.233)	(0.236)	(0.234)	(0.227)	(0.176)	(0.231)	(0.196)	(0.208)	(0.181)
Employment	0.010	0.005	0.001	-0.007	-0.020	-0.023	-0.045	-0.050	-0.030
	(0.036)	(0.023)	(0.029)	(0.040)	(0.032)	(0.033)	(0.039)	(0.045)	(0.040)
Monthly household income ($1000's)	0.212	0.071	0.292	0.460	0.362	0.328	0.277	0.283	0.132
	(0.217)	(0.178)	(0.205)	(0.209)	(0.204)	(0.235)	(0.230)	(0.245)	(0.309)
FAH expenditures per person	0.315***	0.395***	0.619***	0.672***	0.676***	0.644***	0.708***	0.683***	0.743***
	(0.087)	(0.104)	(0.115)	(0.109)	(0.145)	(0.137)	(0.140)	(0.140)	(0.172)
FAFH expenditures per person	0.118	0.101	0.033	0.069	-0.034	-0.008	-0.055	-0.091	-0.050
	(0.107)	(0.095)	(0.095)	(0.103)	(0.098)	(0.090)	(0.101)	(0.120)	(0.132)

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ denote significance levels. Standard errors are shown in parentheses.
RIF decomposition was performed for each 5th quantile from Q15 to Q85; Selected quantiles are shown for brevity.

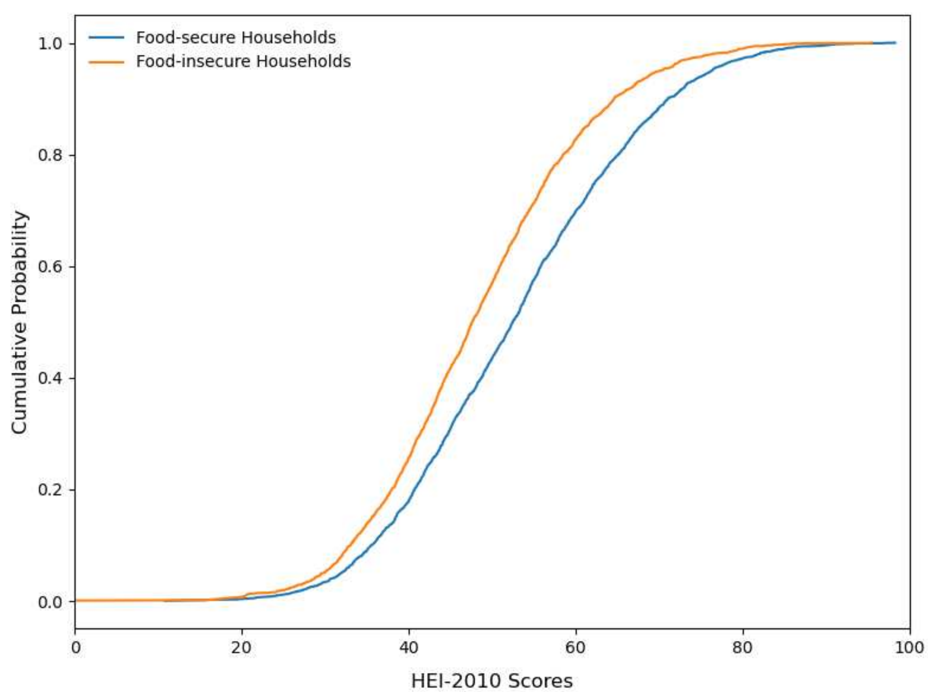
Table 2.4: Estimates of the Contributions of Covariates to the Coefficient Effect

Variable	Q15	Q25	Q35	Q45	Q50	Q55	Q65	Q75	Q85
Male	0.094	0.172	0.016	0.263	0.371	0.194	0.262	0.127	0.003
	(0.355)	(0.293)	(0.337)	(0.348)	(0.359)	(0.361)	(0.404)	(0.370)	(0.387)
Age	0.211	1.607	1.587	1.422	2.294	2.792**	0.464	-1.138	-4.671**
	(1.891)	(2.067)	(2.001)	(2.263)	(1.993)	(1.328)	(2.570)	(1.871)	(2.048)
White Hispanic	0.322	0.279	0.175	0.631*	0.825**	0.808***	0.409	0.232	0.350
	(0.322)	(0.316)	(0.319)	(0.338)	(0.396)	(0.301)	(0.374)	(0.415)	(0.478)
Some college	0.008	0.117	0.036	-0.488	-0.143	0.406	0.238	0.404	-0.001
	(0.447)	(0.517)	(0.505)	(0.457)	(0.496)	(0.448)	(0.455)	(0.455)	(0.663)
College diploma	-0.011	0.046	0.126	0.214	0.288	0.301	0.232	0.319	0.006
	(0.120)	(0.143)	(0.162)	(0.169)	(0.179)	(0.149)	(0.197)	(0.209)	(0.242)
Post graduate degree	0.000	-0.020	0.044	0.134*	0.172**	0.192**	0.245***	0.249**	0.149
	(0.059)	(0.049)	(0.067)	(0.070)	(0.083)	(0.088)	(0.085)	(0.109)	(0.103)
Married	0.812	0.763	1.029**	1.114*	1.044**	1.181**	1.355**	1.646***	2.728***
	(0.476)	(0.554)	(0.418)	(0.572)	(0.483)	(0.489)	(0.677)	(0.604)	(0.876)
Divorced	0.416	0.417	0.093	0.290	0.340	0.379	0.561	0.620	1.410
	(0.424)	(0.377)	(0.286)	(0.391)	(0.407)	(0.349)	(0.460)	(0.374)	(0.455)
Smoking	0.109	0.080	0.290	0.314	0.112	0.218	0.027	0.162	0.718
	(0.493)	(0.548)	(0.395)	(0.484)	(0.572)	(0.526)	(0.532)	(0.547)	(0.633)
Number of children	0.030	-0.092	0.065	0.188	0.636	0.815	1.254	0.632	-0.329
	(0.580)	(0.657)	(0.606)	(0.691)	(0.996)	(0.747)	(1.063)	(0.966)	(0.716)
Employment	-0.821	0.802	0.487	0.208	-0.177	0.401	-0.352	-1.260	-0.258
	(1.255)	(1.299)	(1.169)	(1.122)	(1.032)	(1.005)	(1.093)	(1.360)	(1.080)
Monthly household income ($1000's)	-0.165	0.353	-0.502	0.041	0.190	-0.010	-0.506	-0.670	-1.598
	(0.593)	(0.649)	(0.730)	(0.754)	(0.667)	(0.891)	(0.708)	(0.867)	(0.895)
FAH expenditures per person	-0.669	-0.217	0.286	0.694	0.754	1.060	1.016	0.758	1.003
	(0.619)	(0.631)	(0.594)	(0.756)	(0.652)	(0.649)	(0.602)	(0.470)	(0.669)
FAFH expenditures per person	-0.072	-0.034	0.056	-0.126	-0.065	-0.147	-0.411	-0.152	-0.803
	(0.272)	(0.267)	(0.361)	(0.326)	(0.343)	(0.359)	(0.327)	(0.431)	(0.502)

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ denote significance level. Standard errors are shown in parentheses.
RIF decomposition was performed for each 5th quantile from Q15 to Q85; Selected quantiles are shown for brevity.



(a) Kernel Density Distribution of HEI-2010 Scores



(b) Cumulative Distribution Functions of HEI-2010 Scores

Figure 2.1: Comparative Distributions of Healthy Eating Index-2010 Scores (HEI-2010) for Food-secure and Food-insecure Households: Kernel Density and Cumulative Distribution Functions

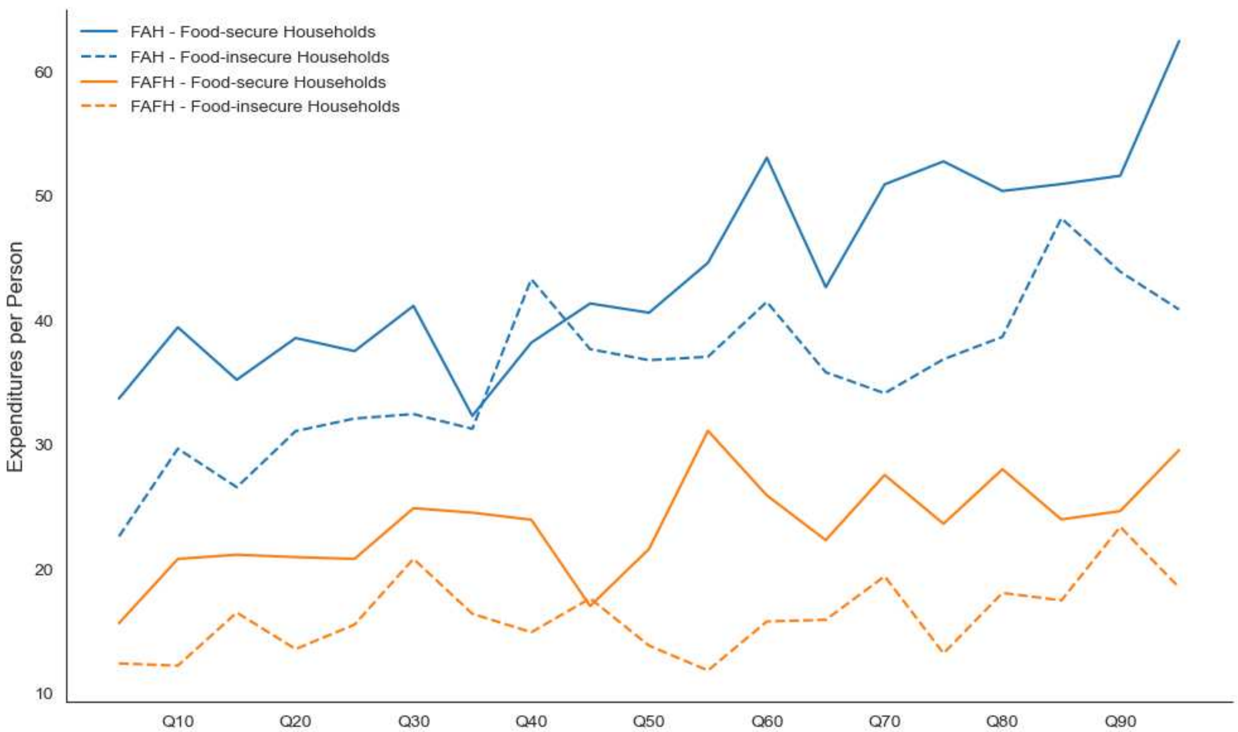


Figure 2.2: Average Food Expenditures Per Person for Each Inter-quantile Range of Healthy Eating Index-2010 (HEI-2010)

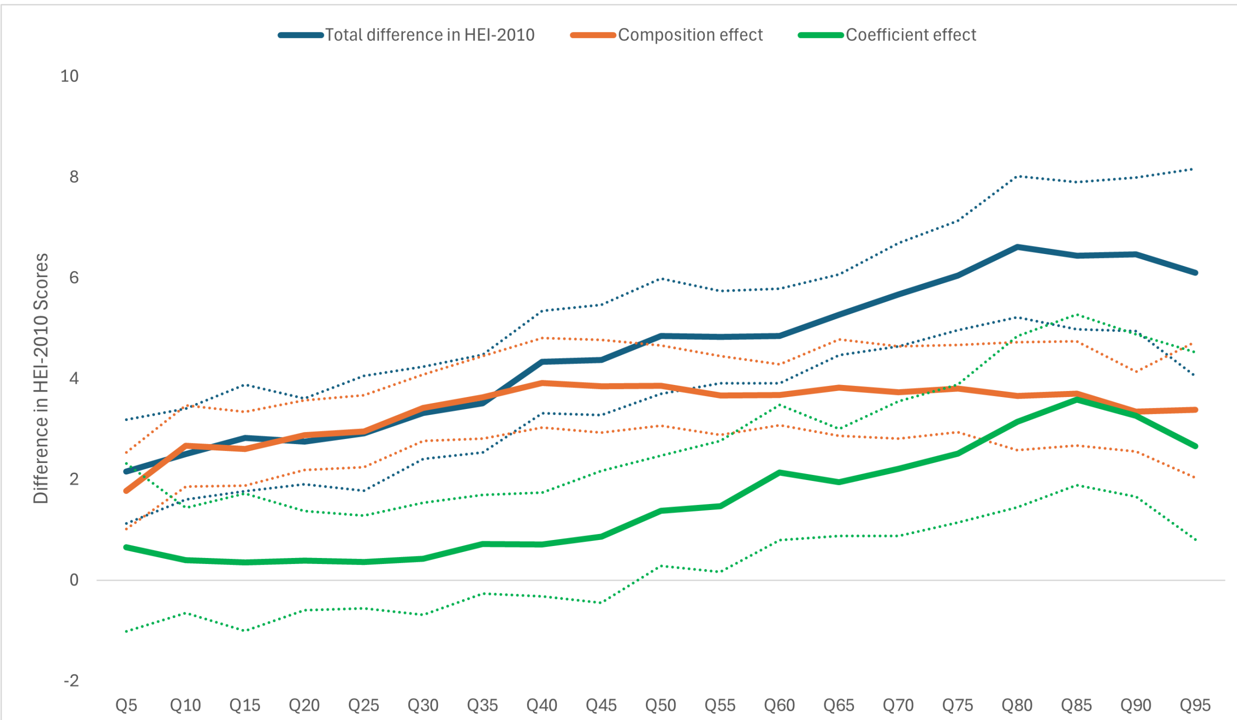


Figure 2.3: Differences in Healthy Eating Index-2010 (HEI-2010) Scores across Food Security Status

Note: This figure shows the tendency of total differences in HEI-2010 scores across food security status, the composition effect, and the coefficient effect. The differences are HEI-2010 scores of food-secure household minus those of food-insure household. Dotted lines represent the 95% confidence intervals.

Chapter 3 The Effect of SNAP on Black Households' Nutritional Quality of Food Purchases

3.1 Introduction

The Supplemental Nutrition Assistance Program (SNAP) plays a critical role in alleviating food insecurity and improving nutrition among low-income households in the United States. As the second largest means-tested food assistance program, SNAP served 12.5% of the U.S. population in an average month in fiscal year 2022. While SNAP aims to enhance food security and dietary quality by increasing food expenditures, its effectiveness in promoting healthier food choices remains inconclusive. Empirical studies have produced mixed evidence regarding SNAP's impact on the nutritional quality of food purchases. Previous research shows that households participating in SNAP tend to increase their food expenditures by more than the amount of benefits received (Hastings and Shapiro, 2018). This suggests that SNAP may help improve nutrition by enabling households to allocate more resources towards food. However, the extent to which increased food spending translates into better diet quality remains uncertain. Some studies find that greater food expenditures are modestly associated with improvements in dietary quality, particularly through increased consumption of fruits and vegetables (Mabli et al., 2010). Other evidence shows that SNAP is associated with improvements in specific aspects of dietary intake, such as higher consumption of vegetables and proteins like poultry and fish, and with reductions in fast food consumption and food insecurity. Despite these benefits, overall improvements in diet quality are modest (Anderson and Butcher, 2016). The nutritional impact of SNAP also differs across population groups. Among adults, SNAP participation is associated with higher intake of whole fruits and lower consumption of sodium and saturated fat, but these changes do not consistently translate into improvements in overall diet quality scores (Gregory et al., 2013). For children, SNAP alone does not enhance overall nutrient adequacy and may be linked to slightly lower fiber intake, with no significant effects observed for iron, potassium, or vitamin E (Yen, 2010).

Effectively evaluating the impact of SNAP participation on nutritional outcomes requires a deeper understanding of the racial disparities that shape both program utilization and health-related outcomes. In 2022, White, non-Hispanic individuals represented 62.7% of SNAP recipients, while comprising 75.5% of the total U.S. population. In contrast, Black, non-Hispanic individuals accounted for 27% of SNAP recipients but made up only 13.6% of the population, indicating overrepresentation. These patterns point to the unequal reliance on SNAP across racial groups and reflect structural inequities. At the same time, while overall diet quality in the U.S. has modestly improved in the past two decades, these improvements have not benefited all groups equally, gains have been smallest for Black and Hispanic individuals and largest for White individuals (Wang et al., 2014; Rehm et al., 2016). Further evidence suggests significant racial and ethnic differences in food, beverage, and nutrient purchases among low-income households, with many of these differences favoring White households (Grummon and Taillie, 2018). These disparities underscore the importance of conducting an analysis of SNAP's impact across racial groups. However, most research assumes a homogeneous response to SNAP participation across all racial groups, typically controlling for race with a simple categorical or binary variable that fails to account for interaction effects (Leung et al., 2012; Gregory et al., 2013; Zhang et al., 2018; Singleton et al., 2020; Hudak et al., 2021). This approach obscures how social, economic, and cultural contexts, particularly those faced by Black households, may shape both food purchasing behavior and SNAP's effectiveness. Higher poverty rates, lower educational attainment, and larger household sizes among Black families likely contribute to differential food purchasing behaviors and nutritional outcomes (French et al., 2019; Smed et al., 2007; Hiza et al., 2013). Moreover, historical and cultural dietary preferences, alongside broader social factors, also influence how BNH households respond to SNAP benefits (Airhihenbuwa et al., 1996; Semmes, 1996). When combined with systemic inequities and structural racism, these conditions may limit the program's potential to improve dietary quality for BNH households. To address these concerns, this study includes a binary indicator for Black households and interacting it with key demographic and socioeconomic variables.

This approach helps identify differential effects of SNAP participation across racial groups and provides a clearer understanding of how race and social context influence program outcomes.

Estimating the mean effects of SNAP participation can mask its heterogeneous impacts across different levels of dietary quality, as SNAP may affect households with lower dietary quality differently than those with higher dietary quality (Feng et al., 2023). This study aims to explore the distributional effects of SNAP on the dietary quality of Black and White households. To address this objective, I employ the instrumental variable unconditional quantile regression (IVUQR) approach (Imbens and Newey, 2009; Rothe, 2010b), which is particularly well-suited for identifying SNAP's impacts across the distribution of nutritional quality. The IVUQR model enables me to evaluate how SNAP affects dietary quality at different levels, providing insights into its impacts on households with low-quality diets separately from those with high-quality diets. Additionally, the IVUQR approach addresses the endogeneity of SNAP participation through the use of instrumental variables (IVs). Unobserved factors, such as household preferences or constraints, may affect both SNAP participation and dietary quality, making it challenging to establish a clear causal relationship. To address this issue, I use state-level welfare policies and SNAP administrative policies as IVs, which are further detailed in the data section. This approach allows me to identify the causal effects of SNAP on nutritional quality across different levels of dietary quality .

To quantify the causal effect of SNAP on dietary quality by race, I use data from the public access National Household Food Acquisition and Purchase Survey (FoodAPS), a nationally representative dataset that tracks U.S. households' food purchases over a 7-day period. Dietary quality is measured using the Healthy Eating Index (HEI)-2010, a widely accepted and validated metric for evaluating nutritional quality (Guenther et al., 2014). In this study, HEI-2010 scores are calculated based on combined food-at-home (FAH) and food-away-from-home (FAFH) acquisitions. On average, SNAP participation is associated with a significant reduction in dietary quality among Black households, while no statistically significant effect is observed for White households. Results from the instrumental variables unconditional quantile regression (IVUQR) model reveal heterogeneity across the distribution. Specifically, SNAP has negative and significant effects on nutritional quality at lower percentiles of the unconditional distribution for both Black and White households, primarily driven by an increased acquisition of empty calories. For households with higher nutritional quality, SNAP has no significant effect. Meanwhile, Black households with lower nutritional quality maintain more consistent dietary quality across food sources compared to White households receiving SNAP benefits.

This study makes several key contributions to the literature on SNAP and racial disparities in nutrition. First, I move beyond simplistic racial controls by explicitly modeling interactions between race and other factors, allowing me to capture the distinct socioeconomic and demographic contexts of Black households. Second, I employ the IVUQR approach to investigate how SNAP impacts dietary quality across the nutritional quality distribution, highlighting effects at the extremes where dietary quality is poorest or most robust (Onvani et al., 2017; Harmon et al., 2015). Third, I also disaggregate the HEI-2010 scores into its component scores to investigate which specific aspects of dietary quality, such as intake of empty calories, whole grains, or fruits and vegetables, are most affected by SNAP participation. By exploring distributional effects and unpacking the mechanisms behind observed disparities, this study advances our understanding of how food assistance programs can more equitably and effectively support dietary improvement across diverse populations.

The remainder of this paper is structured as follows: Section 3.2 details the data used in the analysis, focusing primarily on the public access FoodAPS dataset and the measures of nutritional quality. Section 3.3 outlines the empirical strategy, including the IVUQR approach. Section 3.4 discusses the results. Section 3.5 shows conclusion of my analysis.

3.2 Data

To quantify the impact of SNAP on the nutritional quality of food purchases among Black and White households, I utilize the public access National Household Food Acquisition and Purchase Survey (FoodAPS) as my primary data source. A key empirical challenge is the potential for selection bias: unobserved household characteristics, such as health consciousness, or food preferences, may jointly influence both SNAP participation and dietary quality. To address this endogeneity, I employ a set of state-level instrumental variables (IVs) that affect SNAP participation but not directly the nutritional quality of food acquisitions. These IVs include: (1) state SNAP administrative policies, such as online application availability, operation of statewide call centers, and certification periods for earning units, drawn from the USDA's SNAP Policy Database; (2) state outreach spending per capita, obtained from the USDA's SNAP State Outreach Spending Reports; and (3) the maximum weekly unemployment insurance benefits, sourced from the U.S. Department of Labor's Comparison of State Unemployment Insurance Laws. These instruments proxy for exogenous variation in the likelihood of SNAP enrollment while being plausibly unrelated to unobserved determinants of nutritional outcomes.

3.2.1 FoodAPS Data

The FoodAPS, a nationally representative survey, offers a comprehensive examination of food acquisition patterns in U.S. households, including both food at home (FAH) and food away from home (FAFH) purchases. The survey, conducted between April 2012 and January 2013, collected data from 4,826 households over a week-long period. During the survey period, the primary respondents of each household, typically those responsible for the majority of the household's food acquisitions, participated in two face-to-face interviews and up to three telephone interviews. These interviews captured demographic characteristics and records of all food acquisitions during the surveyed week. To ensure accurate data collection for FAH acquisitions, primary respondents were instructed to scan barcodes on packaged foods. For items without barcodes, such as fresh produce, they used generic codes and provided details regarding weight, quantity, and cost, as evidenced by store receipts. For FAFH acquisitions, receipts from restaurants and stores served as the primary data source, enabling researchers to track purchases made outside the home.

The FoodAPS dataset offers several advantages for researching the impact of SNAP on nutritional quality across racial groups. Firstly, it captures both FAH and FAFH acquisitions. While

SNAP benefits are restricted to FAH purchases, the fungibility of money allow recipients to potentially reallocate their food budget, affecting both FAH and FAFH spending. This creates a complex dynamic where SNAP participation may have differential effects on nutritional quality through both direct (FAH) and indirect (FAFH) channels. The income effect from SNAP benefits could increase overall food spending, including FAFH, while the substitution effect due to reduced FAH costs might decrease FAFH consumption. These effects may vary across racial groups due to cultural food preferences, neighborhood food environments, and socioeconomic factors. For instance, racial disparities in access to healthy food options, both for FAH and FAFH, could influence how SNAP impacts nutritional quality for different groups. By including both FAH and FAFH spending in my model, I can capture a more comprehensive picture of SNAP's influence on dietary patterns and nutritional outcomes across racial categories. Secondly, FoodAPS uses SNAP administrative files to validate self-reported participation, significantly reducing measurement errors common in self-reported data. Thirdly, the deliberate oversampling of SNAP and low-income non-SNAP households increases statistical power for comparing SNAP participants to eligible non-participants. Fourthly, FoodAPS provides insight into household-level food choices and purchasing decisions, rather than individual dietary intake. This household-level focus aligns perfectly with studying SNAP, which operates at the household level.

Between 2002 and 2008, several states modified their SNAP eligibility criteria by raising gross income thresholds above the standard 130% of the Federal Poverty Level (FPL). Eleven states implemented such changes: Arizona, Delaware, Massachusetts, Maryland, North Carolina, Washington, and Wisconsin increased their thresholds to 200% FPL; Maine and Oregon to 185% FPL; and Minnesota and Texas to 165% FPL. To account for these policy adjustments, my analysis includes households with monthly gross incomes up to 200% FPL (Feng et al., 2023) [3].

3. Although several states expanded SNAP gross income thresholds up to 200% of the Federal Poverty Level (FPL) through Broad-Based Categorical Eligibility (BBCE), actual eligibility continued to depend on meeting specific categorical requirements, such as participation in Temporary Assistance for Needy Families (TANF)-related or other non-cash assistance programs. To avoid excluding potentially eligible households, particularly in BBCE states, I include in the sample all households with gross monthly incomes up to 200% FPL. This threshold is used to ensure comprehensive coverage of the SNAP-eligible population, not to define eligibility per se, which varied by state and programmatic linkage.

The public access FoodAPS dataset originally includes 4,826 households. For this study, I focus on households with incomes below 200% of the Federal Poverty Level (FPL), specifically targeting those in which the primary respondent identifies as White or Black and is between 18 and 64 years old [4]. Households with a primary respondent over age 64 are excluded because older adults face distinct participation barriers, such as complex enrollment procedures and stigma, that contribute to persistently lower SNAP take-up compared to younger eligible populations (Dean et al., 2022). Moreover, they are eligible for additional nutrition support programs (e.g., Meals on Wheels, the Commodity Supplemental Food Program), which can modify both their food access and behavioral responses to SNAP (Hake and Dawes, 2024). Applying these criteria narrows the dataset to 1,619 households, of which 877 participate in SNAP, while 742 do not. Households are classified as SNAP participants if they reported receiving SNAP benefits within the 30 days preceding the survey.

3.2.2 Healthy Eating Index

In this study, the Healthy Eating Index (HEI) is utilized as the primary tool for assessing dietary quality of households' food acquisitions. The HEI is widely recognized as a robust and reliable measure, designed to evaluate how closely dietary patterns adhere to the Dietary Guidelines for Americans (DGA) (Guenther et al., 2013, 2014). I specifically used the HEI-2010 since my data comes from the public access FoodAPS survey that was collected from 2012 to 2013 (Guenther et al., 2014).

The HEI-2010 is structured around twelve dietary components, which are categorized into two main groups: adequacy and moderation. Adequacy components include total fruits, whole fruits, total vegetables, greens and beans, whole grains, dairy, total protein foods, seafood and plant proteins, and fatty acids. These components are primarily assessed based on their nutrient density per 1,000 calories, except for fatty acids, which are evaluated through the ratio of polyunsaturated

4. For robustness, the analysis is replicated using a subsample consisting solely of non-Hispanic Black and non-Hispanic White primary respondents aged 18 years or older. The corresponding results are presented in the Appendix.

and monounsaturated fats to saturated fats (Guenther et al., 2013; Berube et al., 2017). Moderation components consist of refined grains, sodium, and empty calories, where scoring is similarly based on nutrient density per 1,000 calories, with the exception of empty calories, which are measured by their proportion of total caloric intake. The HEI employs a scoring scale of 1 to 100, where better nutrition is reflected by higher scores in adequacy categories and lower scores in moderation categories. If the HEI is above 80, it means your diet has high nutritional quality. If it falls between 51 and 80, you need to improve your nutritional quality. If the score is below 51, it means your diet has poor nutritional quality (Beatty et al., 2014; Berube et al., 2017).

Since FoodAPS dataset does not provide direct HEI-2010 scores, I calculated these scores for each household by examining their food acquisitions over a seven-day period, following the HEI-2010 scoring criteria outlined in Table A1. This approach allowed me to construct both FAH+FAFH HEI-2010 scores, incorporating both FAH and FAFH acquisitions, and FAH HEI-2010 scores, based solely on FAH acquisitions, for a sample of 1,619 households. Figure 3.1 presents kernel density distributions of FAH+FAFH HEI-2010 scores by household race and SNAP participation. Figure 3.1a focuses on racial differences between White and Black households. For FAH+FAFH HEI-2010 scores, the density distributions of White and Black households largely overlap. However, White households are more concentrated in the 60-90 point range, possibly reflecting stronger emphasis on diet quality. Figure 3.1b narrows the analysis to Black households, distinguishing between SNAP participants and non-SNAP participants. It shows that very few Black SNAP households fall above the 80th percentile of the FAH+FAFH HEI-2010 distribution, and there are also few with extremely low HEI-2010 scores. This sparsity may lead to imprecise estimates at the nutritional quality distribution tails. However, because I use unconditional quantile regression on the pooled sample of Black and White households, each quantile includes the same number of observations from the full sample. As a result, the estimates reflect relationships across the joint distribution, which helps mitigate the influence of sparse observations in any single subgroup. For FAH+FAFH HEI-2010 scores, Black SNAP households show a higher peak around scores of 40, whereas the distribution for Black non-SNAP households is more spread out, with greater density in both the lower and upper tails. This pattern may suggest that SNAP influences Black households' nutritional quality across the entire FAH+FAFH HEI-2010 score distribution. Figure 3.1c compares White SNAP and non-SNAP households. Across both FAH+FAFH HEI-2010 scores, White non-SNAP households show a shift toward higher values. This pattern reflects generally better diet quality among White non-SNAP households [5]

3.2.3 Household and state socio-economic characteristics

The sample consists of 1,619 households, including 1,4269 White households and 350 Black households. Table 3.1 shows demographic and socioeconomic differences between White and Black households. Black households have higher SNAP participation rates, are less likely to be married, and have lower food-at-home expenditures per person and lower household incomes. They are also less likely to reside in rural areas and have shorter driving distances to the nearest SNAP-authorized supermarket or superstore compared to White households.

Tables 3.2 and 3.3 present detailed comparisons between SNAP and non-SNAP participants within each racial group. Several patterns are consistent across both racial groups. Among both Black and White households, non-SNAP participants are more likely to have male primary respondent, married household heads, and be non-smokers. They also have higher household incomes, higher FAFH expendtirue per person, and fewer children. Additionally, non-SNAP households across both racial groups have lower participation rates in other assistance programs compared to their SNAP-participating counterparts.

To account for factors that may influence both SNAP participation and dietary quality, I include a range of state-level variables. The first category comprises economic indicators that reflect the broader economic environment. These include the monthly unemployment rate, annual per capita GDP, and annual per capita income, sourced from the U.S. Bureau of Labor Statistics, the U.S. Bureau of Economic Analysis, and the 2008–2012 American Community Survey. These indica-

5. Kernel density distributions of FAH HEI-2010 scores by household race and SNAP participation are presented in Appendix Figure A1.

tors are important because states experiencing economic hardship may be more likely to adopt less restrictive SNAP policies as a strategy to mitigate food insecurity and support low-income residents. Moreover, economic conditions can directly impact household employment opportunities, income, and food purchasing behaviors, which in turn affect dietary quality. Additionally, I incorporate educational attainment levels from the U.S. Census Bureau, as higher community education levels could enhance program awareness and promote better nutritional knowledge, both of which are likely to influence dietary choices and SNAP participation.

3.2.4 Instruments

One major challenge in estimating SNAP's impact on dietary quality is selection bias. To address this, I employ an instrumental variables (IV) approach. Specifically, I use exogenous variations in state-level welfare policies and SNAP administrative policies as instruments to control for selection into the program. The welfare policy instruments include state's maximum weekly unemployment insurance (UI) benefits and state SNAP outreach spending per capita. The SNAP administrative policy instruments include three measures: whether the state allows online SNAP applications, whether the state operates statewide call centers, and the average certification period in months for SNAP households with earnings.

For the IV strategy to be valid, each instrument must meet two conditions. First, the instrument must be relevant, meaning it is correlated with SNAP participation. Second, it must satisfy the exclusion restriction, meaning it influences dietary quality only through its effect on SNAP participation, and not through any other channel. To structure the identification strategy, I treat maximum weekly UI benefits as the primary instrument that satisfies the exclusion restriction. UI is a collaborative federal-state program that provides temporary financial assistance to eligible unemployed workers (Feng et al., 2023). The level of UI benefits, as measured by the maximum weekly benefit amount, varies across states and is determined by state UI laws. Previous research has shown that UI benefit levels can influence SNAP participation through two potential channels. First, receiving UI benefits may increase awareness of other safety net programs like SNAP, leading to a positive information effect" on SNAP participation (Finifter and Prell, 2013). Second, higher UI benefits may increase household income and reduce SNAP eligibility, resulting in a negative effect on SNAP participation (Reich and West, 2015). Importantly, the maximum weekly UI benefits, set by states, are likely unrelated to individual dietary quality, as they are determined by state-specific factors such as labor market conditions and policy preferences rather than individual characteristics. Several studies have utilized variation in UI benefit levels as a source of exogenous variation to estimate the causal effects of UI on various outcomes, such as job search behavior (Krueger and Mueller, 2010), and health (Kuka, 2020). Based on this reasoning, I treat the maximum weekly UI benefits as satisfying the exclusion restriction and use them as the basis for identification.

The remaining instruments are evaluated under the assumption that the exclusion restriction holds for maximum weekly UI benefits. SNAP outreach spending per capita reflects a state's investment in promoting program participation. These funds are generally allocated toward informational campaigns and application assistance, not toward changing individual dietary behavior. Outreach spending is primarily shaped by political and budgetary decisions that are external to household nutritional choices. Previous research has successfully employed this instrument in related contexts, such as estimating SNAP's effect on food insecurity and childhood obesity (Ratcliffe et al., 2011; Schmeiser, 2012), lending empirical support to its validity. The other three instruments capture administrative aspects of SNAP enrollment. Online application availability, call center operation, and certification period length influence the ease of accessing the program. These features are not designed to change what households eat but to reduce the transaction costs of applying for and maintaining benefits. For instance, whether a state offers online applications or operates a centralized call center affects the burden of applying or recertifying for SNAP but does not change what foods are available to households or what they choose to purchase. Similarly, longer certification periods reduce the administrative burden for recipients but do not change the relative prices or preferences for healthy foods. To account for the possibility that these administrative features are correlated with unobserved factors affecting food choice, I test their validity empirically.

Assuming the exclusion restriction holds for the UI benefit variable, I use Hansen's J test to assess whether the additional instruments satisfy the exogeneity condition. A failure to reject the null hypothesis in this test supports the claim that the remaining instruments are valid. Furthermore, the first-stage regression confirms the relevance of the full set of instruments, with F-statistics exceeding the standard threshold of 10, indicating strong predictive power for SNAP participation. Details of these diagnostic tests are provided in the Results section.

I obtain data on maximum weekly UI benefit levels from the U.S. Department of Labor's Comparison of State UI Laws (U.S. Department of Labor., 2012, 2013), state SNAP outreach spending data and snap administrative policy from the USDA's SNAP Policy Database (Economic Research Service (ERS), U.S. Department of Agriculture (USDA)., 2019). These variables are merged with the public access FoodAPS data by region, year, month.

3.3 Methods

SNAP is a crucial policy intervention aimed at improving nutritional outcomes for low-income households. However, investigating the impact of SNAP on nutritional quality is complicated by the endogeneity of SNAP participation. Endogeneity arises when there are unobserved factors that influence both the decision to participate in SNAP and the nutritional quality of an individual's diet. Failure to account for this endogeneity can lead to biased estimates of the effect of SNAP on nutritional quality.

To address the endogeneity of SNAP participation and investigate the quantile effect of SNAP on nutritional quality across race, I employ the Instrumental Variable Unconditional Quantile Regression (IVUQR) approach (Imbens and Newey, 2009; Rothe, 2010a). IVUQR is an extension of the Instrumental Variable Quantile Regression (IVQR) method (Chernozhukov and Hansen, 2005), which allows for the estimation of the unconditional quantile treatment effects in the presence of endogeneity.

The choice of IVUQR over IVQR is motivated by several advantages that make IVUQR more suitable for my research. Firstly, IVUQR estimates the unconditional quantile treatment effects, which are the effects of SNAP participation on the quantiles of the marginal distribution of nutritional quality, without conditioning on the control variables. In contrast, IVQR estimates the conditional quantile treatment effects, which are the effects of SNAP participation on the quantiles of nutritional quality conditional on the control variables included in the model. The unconditional quantile treatment effects estimated by IVUQR are more interpretable and policy-relevant because they directly answer questions about the impact of SNAP on the overall distribution of nutritional quality, rather than the distribution conditional on specific values of the control variables.

Moreover, the unconditional quantile treatment effects estimated by IVUQR have a clear interpretation as the impact of SNAP on the quantiles of the marginal distribution of nutritional quality. For example, if IVUQR estimates that SNAP participation increases the 25th percentile of the nutritional quality distribution by a certain amount, this can be directly interpreted as the effect of SNAP on the nutritional quality of individuals at the lower end of the distribution, regardless of their characteristics. This interpretation is more useful for policymakers, as it provides insights into how SNAP affects the overall population and specific subgroups, rather than the effect conditional on specific values of the control variables.

The IVUQR model is estimated in two stages:

$$SNAP_{ist} = \alpha_0 + \alpha_1 Z_{st} + \beta_1 Black_{ist} + \beta_2 X_{ist} + \beta_3 Black_{ist} * X_{ist} + u_{ist} \tag{15}$$

$$\begin{aligned} Q_{h_{ist}}(\tau) = {} & \gamma_0(\tau) + \gamma_1(\tau) Black_{ist} + \gamma_2(\tau)\widehat{SNAP}_{ist} + \gamma_3(\tau) Black_{ist} * \widehat{SNAP}_{ist} \\ & + \gamma_4(\tau) X_{ist} + \gamma_5(\tau) Black_{ist} * X_{ist} + w_{ist} \end{aligned} \tag{16}$$

where $SNAP_{ist}$ represents whether a household i, located in region s, participates in SNAP at time t. The vector Z_{st} includes IVs. The vector X_{ist} includes household and individual characteristics and state socio-economic characteristics. u_{ist} and w_{ist} are error terms. In the second stage, $Q_{h_{ist}}(\tau)$ denotes the τ-th quantile of the unconditional distribution of dietary quality for household i residing in region s at time t, and $\widehat{SNAP}_{ist}$ is the predicted value of SNAP participation derived from Equation 15. The α, β, and $\gamma(\tau)$ are parameters to be estimated. $\gamma_2(\tau)+\gamma_3(\tau)$ represents the estimated quantile impact of SNAP participation on the nutritional quality of purchases made by Black households, while $\gamma_2(\tau)$ represents the corresponding effect for White households.

The vector of household characteristics, X_{ist}, includes variables such as age, gender, marital status, smoking status, ethnicity, education, household size, number of children, food-at-home expenditures per person, food-away-from-home expenditures per person, rural residence status, household income, driving distance to the nearest SNAP-authorized supermarket or superstore, participation in other nutrition assistance programs, including the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), the School Breakfast Program (SBP), and the National School Lunch Program (NSLP) and state socio-economic characteristics. *Income* is measured as the average monthly household income, aggregating the income per member and standardizing it in units of $1,000. I also calculate *FAH expenditures per person* and *FAFH expenditures per person* by summing each household's weekly expenditures on FAH and FAFH, respectively, and dividing by the number of household members. To account for potential regional correlations, standard errors are clustered at the region level.

Following the control function approach proposed by Imbens and Newey (2009) and further developed by Rothe (2010a), the IVUQR model is estimated in a three-step procedure. First, I estimate the first-stage regression (Equation 15) and obtain the residuals. Second, I compute the empirical cumulative distribution functions (CDFs) of these residuals. Since the CDFs are monotonic transformations of the first-stage errors, they capture the dependence between the endogenous regressor and the unobserved components of the outcome equation. If the instruments are valid, including the CDFs helps isolate exogenous variation relevant for identification (Rothe, 2010a; Baylis et al., 2019). In the final step, the estimated CDFs are included as control variables in the second-stage quantile regression (Equation 16) [6].

6. All estimation procedures were implemented in Python using custom code, without relying on off-the-shelf econometric packages. The implementation uses standard Python libraries including pandas, numpy, statsmodels for OLS estimation, and scikit-learn for kernel density estimation. The IVUQR estimator follows the logic of the control function approach and employs bootstrapping to compute standard errors and confidence intervals.

3.4 Results

This section presents and discusses the estimated effects of SNAP on nutritional quality using three approaches: an OLS model, a two-stage least squares (2SLS) model (referred to as IV regression in this paper), and the IVUQR model, given by Equations 15 and 16. Additionally, I examine the quantile effects of SNAP on dietary quality for FAH+FAFH food consumption, using pooled data for White and Black households.

3.4.1 Nutritional Quality Differences Between SNAP and Non-SNAP Households

Table 3.4 reports the mean differences in FAH+FAFH HEI-2010 total and the 12 individual HEI-2010 component scores between Black SNAP and non-SNAP households [7]. Dietary quality from combined FAH and FAFH acquisitions is lower among Black SNAP households compared to their non-SNAP counterparts. On average, the FAH+FAFH HEI-2010 score for Black SNAP households is 3.22 points lower than that of Black non-SNAP households. Both groups fall into the poor nutritional quality category, as HEI-2010 scores below 51 are considered poor (Beatty et al., 2014; Berube et al., 2017). For individual adequacy components, where higher scores indicate better nutritional quality, Black SNAP households have significantly lower scores for total vegetables, greens and beans, whole fruit, and whole grains. These components are essential for a nutritionally balanced diet, and deficiencies in them may partly explain the lower FAH+FAFH HEI-2010 scores observed among Black SNAP participants.

7. Appendix Table A5 provides corresponding mean differences for the FAH HEI-2010 and component scores between Black SNAP and non-SNAP households.

3.4.2 Average Effects of SNAP

Table 3.5 displays the OLS and IV regression estimates for Black and White households, with the dependent variable being the FAH+FAFH HEI-2010 score [8]. The OLS results, which do not account for potential self-selection bias, suggest that SNAP participation is not significantly associated with dietary quality among Black households when considering combined FAH and FAFH acquisitions. In contrast, among White households, SNAP participation is statistically associated with lower dietary quality. Specifically, White SNAP participants have, on average, a 2.278-point lower FAH+FAFH HEI-2010 score compared to non-participants. However, these OLS estimates may be biased due to unobserved factors that influence both SNAP participation and diet quality. To address this endogeneity, IV regression is employed.

To assess the validity of the instrumental variables employed in this analysis, I conduct two diagnostic tests: a relevance test and Hansen's J test for overidentifying restrictions. The relevance test evaluates whether the instruments are sufficiently correlated with the endogenous variable, SNAP participation, by testing the joint significance of the instruments in the first-stage regression. According to Table 3.5, the first-stage F-statistic is 24.155, which exceeds the conventional threshold of 10. This provides strong evidence that the instruments are relevant and have substantial explanatory power for predicting SNAP participation. To assess the exogeneity of the instruments, I use Hansen's J test for overidentifying restrictions. The test produces a p-value of 0.321, indicating that the null hypothesis—that the instruments are uncorrelated with the structural error term—cannot be rejected. This supports the validity of the overidentifying restrictions and the assumption that the instruments are exogenous. The IV estimates, reported in Table 3.5, suggest that SNAP participation significantly reduces the combined FAH and FAFH HEI-2010 dietary quality score among Black households by 14.907 points. In contrast, the estimated effect for White households is not statistically significant, highlighting potential racial heterogeneity in the nutritional impacts of SNAP participation.

8. Appendix Table A6 provides OLS and IV regression estimates using the FAH HEI-2010 score as the dependent variable for robustness checks.

The HEI-2010 score combines 12 components. Examining the total HEI score alone may overlook the heterogeneous effects of SNAP on individual components. Table 3.6 reports the IV estimates for SNAP's impact on the 12 FAH+FAFH HEI-2010 components. Among Black households, SNAP participation is associated with acquiring fewer total vegetables from FAH and FAFH consumption, reducing the FAH+FAFH HEI-2010 score by 1.988 points. In contrast, for White households, SNAP participation is linked to acquiring fewer greens and beans from FAH and FAFH consumption, leading to a 2.284-point decrease in the FAH+FAFH HEI-2010 score.

Table 3.7 presents the effects of SNAP participation on the 12 FAH HEI-2010 component scores. Among Black households, SNAP participants acquire less seafood/plant proteins for their FAH consumption, reducing the FAH HEI-2010 by 2.587 points. However, no negative impact is observed on the FAH+FAFH HEI-2010 score for this component. This suggests that SNAP increases or maintains seafood/plant protein acquisitions from FAFH, resulting in no significant difference in the combined FAH+FAFH HEI-2010 score for this component between SNAP and non-SNAP households. These findings imply that SNAP participation shifts seafood/plant protein purchases from FAH to FAFH among Black households.

Similarly, Black households participating in SNAP acquire fewer total vegetables for their FAH consumption, which reduces the FAH HEI-2010 score by 2.536 points. This negative effect is larger than the effect in the FAH+FAFH HEI-2010 total vegetable score (1.988 points), suggesting that SNAP households acquire more total vegetables from FAFH while purchasing fewer from FAH. This pattern may reflect a shift in consumption behavior enabled by SNAP benefits, whereby increased flexibility in the household food budget allows for greater access to FAFH options that include vegetables. However, further data would be needed to confirm this mechanism. For White households, SNAP participation has no effect on greens and beans acquisition from FAH. However, it is associated with acquiring fewer greens and beans from FAH and FAFH combined. This suggests that the reduction in greens and beans acquisition among SNAP participants primarily comes from FAFH consumption.

3.4.3 Distribution effects of SNAP

Appendix Figures A2 and A3 illustrate estimates derived from unconditional quantile regression (UQR), which do not address potential endogeneity issues related to SNAP participation. These results highlight that the impact of SNAP varies significantly across different levels of dietary quality. For the FAH+FAFH HEI-2010 scores, SNAP participation negatively influences Black households primarily at higher dietary quality quantiles, while for White households, negative impacts occur across most quantiles, except at the 90th quantile. Specifically, for Black households with higher dietary quality, SNAP participation is associated with reductions in HEI-2010 scores ranging from approximately 4.321 to 6.424 points. For White households, estimated reductions range from 1.789 to 3.209 points across most quantiles. However, for both racial groups, the UQR estimates are not statistically different from the corresponding OLS point estimates. It is important to note that because SNAP participation is endogenous, and the UQR framework does not control for unobserved factors that may influence both program participation and dietary quality, these estimates may be biased.

According to Table 3.5, the IV regression results indicate a significant negative average effect of SNAP participation on FAH+FAFH dietary quality for Black households. However, this average effect does not necessarily imply that SNAP has the same negative effect across all Black households, regardless of their initial nutritional quality. Similarly, while the IV regression results show no significant average effect of SNAP participation on FAH+FAFH or FAH dietary quality for White households, this does not mean that SNAP has no effect on White households with lower or higher nutritional quality. To explore whether households with different levels of dietary quality are affected differently by SNAP, I employ the IVUQR model. Table 3.5 reports the distributional effects of SNAP participation on FAH+FAFH HEI-2010 scores, along with 90% confidence intervals. These effects are also visualized in Figure 3.2 [9].

As shown in Table 3.5, I find significant negative effects of SNAP participation on the uncon-

9. Appendix Table A6 presents the IVUQR estimates for FAH HEI-2010 scores, with 90% confidence intervals. Appendix Figures A4 provide corresponding visualizations.

ditional distributions of HEI-2010 scores, particularly among households with low-to-intermediate nutritional quality. For Black households at the 20th, 30th, 40th, and 50th percentiles of the FAH+FAFH HEI-2010 distribution, SNAP participation decreases dietary quality by 16.697, 13.712, 16.567, and 14.121 points, respectively. Given average FAH+FAFH HEI-2010 scores of 37.951, 41.313, 44.324, and 47.918 points at the 20th, 30th, 40th, and 50th percentiles, respectively, the estimated reductions correspond to 44%, 33%, 37% , and 29% of their respective averages. For White households, SNAP participation reduces FAH+FAFH HEI-2010 scores by 17.251 points (52%) at the 10th percentile, 19.788 points (52%) at the 20th percentile, 17.643 points (43%) at the 30th percentile, 19.296 points (44%) at the 40th percentile, 15.772 points (33%) at the 50th percentile, and 10.208 points (20%) at the 60th percentile. These findings align with prior research. Smith and Valizadeh (2024) documented significant declines in diet quality (10.5% – 21.5%) for children losing WIC access, particularly among those with lower-quality diets, while children with higher-quality diets were unaffected. Similarly, Feng et al. (2023) found that SNAP participation significantly decreases diet quality (17% – 23%) for households at the 30th–50th percentiles of the unconditional distribution of diet quality. Consistent with these studies, my analysis shows that SNAP participation leads to at least a 20% decrease in HEI-2010 scores, out of a total score of 100, for Black and White households in the 10th–60th percentiles of the unconditional distribution of diet quality. However, as shown in Figure 3.2, the IVUQR estimates are not statistically different from the average treatment effects estimated by IV regression. This may be because the sample size is not large enough to detect significant differences across quantiles [10].

SNAP participation decreases nutritional quality for Black households, particularly for those with lower nutritional quality. My initial hypothesis was that SNAP benefits might encourage Black households to switch from cheaper frozen vegetables to more expensive fresh vegetables.

10. As a robustness check, I also estimated the models using a pooled sample restricted to Black non-Hispanic and White non-Hispanic households with household heads aged 18 or older. Appendix Figure A5 and Figure A6 present unconditional quantile regression estimates of the effect of SNAP participation on HEI-2010 scores for FAH+FAFH and FAH, respectively, among Black and White non-Hispanic households. These results show that the IVUQR estimates are statistically equivalent to the IV regression estimates.

While fresh produce is often perceived as higher quality, it is also more expensive and perishable compared to frozen or canned alternatives. With a fixed budget, Black households might purchase fewer vegetables overall or reduce other healthy food items to afford fresh produce. This could lead to a decline in total vegetable acquisition and, consequently, lower nutritional quality. However, when I examine the data, they do not support this hypothesis. As shown in Figure 3.3, SNAP participants acquire fewer fresh and frozen/canned vegetables for FAH overall. While they obtain more fresh dark green vegetables, frozen dark green vegetables, and canned red and orange vegetables, their acquisition of other vegetable groups is lower. This pattern is further supported by Table A5, which shows that the average component score for total vegetables is lower for Black SNAP households than for Black non-SNAP households. This suggests that Black SNAP households acquire fewer total vegetables than their non-SNAP counterparts.

The negative effect of SNAP on the nutritional quality of Black households is driven by increased acquisition of sugary foods and beverages. As shown in Figure 3.4, SNAP participants acquire more sweetened coffee and tea, caloric beverages, sweeteners, candies, cake mixes, milk drinks and desserts, and other desserts compared to non-participants. These items are major sources of added sugar, a key contributor to empty calories, which in turn lowers overall diet quality among Black SNAP households.

Table 3.8 provides estimates of Black households' heterogeneous responses to SNAP across quantiles of the unconditional distributions of 12 individual FAH+FAFH HEI-2010 component scores. The results indicate that SNAP reduces the FAH+FAFH HEI-2010 component score for total vegetables at the 10th, 20th, 60th, and 70th percentiles, suggesting that SNAP generally leads to lower vegetable acquisitions from FAH+FAFH. Additionally, Black SNAP households at the intermediate-to-high quantiles (50th–80th) acquire fewer greens and beans from FAH+FAFH. Meanwhile, those at the 30th and 50th percentiles acquire more empty calories from FAH+FAFH than Black non-SNAP households, further contributing to lower diet quality.

Table 3.9 presents estimates of Black households' heterogeneous responses to SNAP across quantiles of the unconditional distributions of 12 individual FAH HEI-2010 component scores.

The results show no significant effects of SNAP on the FAH HEI-2010 component scores for total vegetables at the 10th and 20th percentiles, suggesting that for Black households with the lowest dietary quality, SNAP does not influence FAH vegetable purchases but instead reduces vegetable acquisitions from FAFH. For Black households with middle dietary quality (30th, 40th, 60th, and 70th percentiles), the negative effect of SNAP on the FAH HEI-2010 component score for total vegetables is larger than its effect on the FAH+FAFH HEI-2010 score. This suggests that SNAP households reduce their FAH purchases of total vegetables while increasing their FAFH acquisitions of this component. For greens and beans, I do not observe significant negative effects of SNAP on FAH HEI-2010 component scores from the 50th to 80th percentiles. This suggests that for Black households with intermediate-to-high dietary quality, SNAP does not impact FAH acquisitions of greens and beans, but it reduces their FAFH consumption of these components. Lastly, for Black households with middle dietary quality (40th–70th percentiles), the negative effect of SNAP on the FAH HEI-2010 component score for empty calories is larger than its effect on the FAH+FAFH HEI-2010 score. This indicates that SNAP households increase their FAH purchases of empty calories while decreasing their FAFH consumption of these items [11]

3.5 Conclusions

In this paper, I address a critical gap in the literature by examining racial differences in the impact of SNAP on dietary quality. Rather than using a simple categorical control for race, I model interactions between race and other factors to better capture the distinct socioeconomic and demographic contexts of Black and White households. To estimate the causal effect of SNAP participation, I employ an IVUQR approach, which allows me to examine SNAP's effects across the entire distribution of dietary quality. This distributional approach provides insights into how SNAP affects households with different levels of nutritional quality, particularly those at the lower

11. Appendix Table A7 and Appendix Table A8 shows estimates of White households' heterogeneous responses to SNAP across quantiles of the unconditional distributions of the 12 individual FAH+FAFH HEI-2010 component scores and the FAH component scores, respectively.

end of the nutritional quality distribution.

My analysis shows that, on average, SNAP has a significant negative impact on the dietary quality of Black households but no significant effect on White households. This null effect may result from different household responses to SNAP. While some households may improve their dietary quality by acquiring healthier and more nutritious foods, others may experience a decline by increasing their consumption of less nutritious options. As these opposing effects balance out, the overall average impact of SNAP on dietary quality appears negligible. Additionally, SNAP benefits may enable households to expand their overall food purchases, including both healthy and unhealthy items, potentially offsetting any positive dietary improvements.

However, IVUQR results indicate substantial heterogeneity in SNAP's impact. At the lower percentiles of the dietary quality distribution, SNAP significantly reduces FAH+FAFH HEI-2010 scores for both Black and White households. For both groups, the IVUQR estimates are statistically equivalent to the IV regression estimates. For Black households, the negative effect of SNAP on nutritional quality suggests that the program's current design may not effectively promote healthier food choices. A key potential driver of this negative effect is an increase in empty calorie acquisitions, including sugar-sweetened beverages, candies and desserts. Another potential explanation is household time constraints. Previous research (Davis and You, 2010) finds that SNAP participants spend about 26% more time preparing FAH meals that meet the Thrifty Food Plan target compared to non-SNAP participants. This additional time cost can create a significant barrier, leading households, especially those with lower nutritional quality, to rely more on convenience foods that are often less nutritious. Devine et al. (2006) describe how low-wage, working parents routinely adopt food-choice coping strategies that prioritize convenience when balancing demanding work and family schedules, often at the expense of nutritional quality. Devine et al. (2003) detail how "sandwiching" food preparation into fragmented, unpredictable time frames leads many households to depend on highly convenient or ready-to-eat meals, even when these are nutritionally inferior. Together, these findings suggest that while SNAP increases food purchasing power, it does not address the time demands of preparing nutritious meals. For households with limited time, this can lead to a heavier reliance on convenience foods, which are often lower in nutritional value.

To improve the nutritional quality of food purchases among Black households, one potential policy approach is restricting SNAP purchases of foods high in empty calories, such as sugar-sweetened beverages, candies and desserts. Research indicates that households are more likely to purchase food and beverages when receiving SNAP benefits compared to receiving cash assistance (Smith et al., 2016; Hastings and Shapiro, 2018). Limiting SNAP-eligible items to healthier options could help shift consumption patterns toward more nutritious foods. Another policy implication arises from the observed differences in how SNAP affects FAH versus FAFH acquisitions across racial groups. Black SNAP households maintain relatively stable diet quality across sources, while White households show greater substitution from FAFH to FAH without improvement in nutritional outcomes. This suggests that increasing access to affordable, healthy FAH options, such as through subsidies for fresh produce at grocery stores or support for mobile markets in underserved neighborhoods, may be particularly effective.

Table 3.1: Summary Statistics By Race

Variable	All (1)	White households (2)	Black households (3)	Mean difference (4)	
FAH+FAFH HEI-2010	48.780	48.986	48.035	0.951	
	(0.320)	(0.365)	(0.662)		
FAH HEI-2010	47.578	48.094	45.704	2.390	**
	(0.367)	(0.415)	(0.782)		
SNAP participation	0.542	0.519	0.623	-0.104	***
	(0.012)	(0.014)	(0.026)		
Age	40.946	40.990	40.789	0.201	
	(0.291)	(0.331)	(0.606)		
Male	0.231	0.235	0.217	0.018	
	(0.010)	(0.012)	(0.022)		
Married	0.348	0.389	0.197	0.192	***
	(0.012)	(0.014)	(0.021)		
Non-smoker	0.629	0.630	0.629	0.001	
	(0.012)	(0.014)	(0.026)		
Hispanic	0.190	0.234	0.029	0.205	***
	(0.010)	(0.012)	(0.009)		
Some college	0.340	0.330	0.374	-0.044	
	(0.012)	(0.013)	(0.026)		
Household size	3.317	3.336	3.249	0.087	
	(0.047)	(0.054)	(0.099)		
Number of children	1.368	1.346	1.446	-0.100	
	(0.037)	(0.042)	(0.082)		
FAH expenditure per person	37.718	39.233	32.228	7.005	**
	(0.916)	(1.046)	(1.863)		
FAFH expenditure per person	13.179	12.891	14.223	-1.332	
	(0.462)	(0.442)	(1.411)		
Rural	0.285	0.316	0.171	0.145	***
	(0.011)	(0.013)	(0.020)		
Monthly household income ($1,000's)	1.805	1.858	1.613	0.245	***
	(0.028)	(0.031)	(0.060)		
Driving distance (miles) to the nearest SNAP-authorized supermarket or superstore	2.736	2.907	2.118	0.789	***
	(0.098)	(0.114)	(0.179)		
The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)	0.161	0.164	0.151	0.012	
	(0.009)	(0.010)	(0.019)		
National School Lunch Program (NSLP)	0.464	0.454	0.500	-0.046	
	(0.012)	(0.014)	(0.027)		
School Breakfast Program (SBP)	0.405	0.395	0.443	-0.048	
	(0.012)	(0.014)	(0.027)		
Monthly unemployment rate	7.252	7.256	7.237	0.019	
	(0.015)	(0.017)	(0.034)		
Annual per capita GDP ($1,000's)	53.856	53.803	54.047	-0.245	**
	(0.032)	(0.037)	(0.064)		
Annual per capita income ($1,000's)	27.610	27.680	27.359	0.321	**
	(0.041)	(0.048)	(0.077)		
Proportion of adults with some college or higher (%)	57.385	57.651	56.420	1.232	***
	(0.084)	(0.095)	(0.171)		
Observations	1619	1269	350		

Note: The unit of observations is the household. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ indicate significant differences in sample means between White and Black households. Standard deviations are reported in parentheses.

Table 3.2: Summary Statistics for Black Households Across SNAP Status

Variable	Black (1)	Black SNAP households (2)	Black NonSNAP households (3)	Mean difference (4)	
FAH+FAFH HEI-2010	48.035	46.820	50.042	-3.222	*
	(0.662)	(0.760)	(1.212)		
FAH HEI-2010	45.704	44.260	48.090	-3.830	*
	(0.782)	(0.915)	(1.399)		
Age	40.789	40.642	41.030	-0.388	
	(0.606)	(0.755)	(1.017)		
Male	0.217	0.179	0.280	-0.101	*
	(0.022)	(0.026)	(0.039)		
Married	0.197	0.156	0.265	-0.109	*
	(0.021)	(0.025)	(0.039)		
Non-smoker	0.629	0.550	0.758	-0.207	***
	(0.026)	(0.034)	(0.037)		
Hispanic	0.029	0.028	0.030	-0.003	
	(0.009)	(0.011)	(0.015)		
Some college	0.374	0.353	0.409	-0.056	
	(0.026)	(0.032)	(0.043)		
Household size	3.249	3.376	3.038	0.338	
	(0.099)	(0.128)	(0.156)		
Number of children	1.446	1.587	1.212	0.375	*
	(0.082)	(0.106)	(0.126)		
FAH expenditure per person	32.228	34.686	28.167	6.520	
	(1.863)	(2.570)	(2.498)		
FAFH expenditure per person	14.223	11.730	18.339	-6.608	*
	(1.411)	(1.353)	(2.974)		
Rural	0.171	0.174	0.167	0.008	
	(0.020)	(0.026)	(0.033)		
Monthly household income ($1,000's)	1.613	1.390	1.981	-0.590	***
	(0.060)	(0.067)	(0.105)		
Driving distance (miles) to the nearest SNAP-authorized supermarket or superstore	2.118	2.137	2.087	0.050	
	(0.179)	(0.237)	(0.272)		
The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)	0.151	0.193	0.083	0.109	**
	(0.019)	(0.027)	(0.024)		
National School Lunch Program (NSLP)	0.500	0.532	0.447	0.085	
	(0.027)	(0.034)	(0.043)		
National School Lunch Program (NSLP)	0.443	0.477	0.386	0.091	
	(0.027)	(0.034)	(0.043)		
Monthly unemployment rate	7.237	7.259	7.201	0.058	
	(0.034)	(0.043)	(0.055)		
Annual per capita GDP ($1,000's)	54.047	54.129	53.913	0.216	
	(0.064)	(0.078)	(0.109)		
Annual per capita income ($1,000's)	27.359	27.335	27.399	-0.064	
	(0.077)	(0.098)	(0.127)		
Proportion of adults with some college or higher (%)	56.420	56.168	56.837	-0.669	
	(0.171)	(0.209)	(0.293)		
observations	350	218	132		

Note: The unit of observations is the household. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ indicate significant differences in sample means between Black SNAP and Black NonSNAP households. Standard deviations are reported in parentheses.

Table 3.3: Summary Statistics for White Households Across SNAP Status

Variable	White (1)	White SNAP households (2)	White NonSNAP households (3)	Mean difference (4)	
FAH+FAFH HEI-2010	48.986	47.657	50.421	-2.764	***
	(0.365)	(0.492)	(0.537)		
FAH HEI-2010	48.094	46.793	49.501	-2.708	**
	(0.415)	(0.552)	(0.621)		
Age	40.990	40.345	41.686	-1.341	*
	(0.331)	(0.458)	(0.479)		
Male	0.235	0.203	0.269	-0.066	**
	(0.012)	(0.016)	(0.018)		
Married	0.389	0.325	0.459	-0.134	***
	(0.014)	(0.018)	(0.020)		
Non-smoker	0.630	0.564	0.700	-0.136	***
	(0.014)	(0.019)	(0.019)		
Hispanic	0.234	0.243	0.225	0.018	
	(0.012)	(0.017)	(0.017)		
Some college	0.330	0.319	0.343	-0.024	
	(0.013)	(0.018)	(0.019)		
Household size	3.336	3.539	3.116	0.422	***
	(0.054)	(0.075)	(0.077)		
Number of children	1.346	1.533	1.144	0.388	***
	(0.042)	(0.058)	(0.058)		
FAH expenditure per person	39.233	39.048	39.433	-0.385	
	(1.046)	(1.446)	(1.516)		
FAFH expenditure per person	12.891	10.364	15.621	-5.257	***
	(0.442)	(0.531)	(0.703)		
Rural	0.316	0.307	0.326	-0.020	
	(0.013)	(0.018)	(0.019)		
Monthly household income ($1,000’s)	1.858	1.639	2.095	-0.455	***
	(0.031)	(0.042)	(0.045)		
Driving distance (miles) to the nearest SNAP-authorized supermarket or superstore	2.907	2.776	3.049	-0.273	
	(0.114)	(0.154)	(0.171)		
The Special Supplemental Nutrition Program for Women, Infants, and Children (WIC)	0.164	0.217	0.107	0.110	***
	(0.010)	(0.016)	(0.013)		
National School Lunch Program (NSLP)	0.454	0.504	0.400	0.104	***
	(0.014)	(0.019)	(0.020)		
National School Lunch Program (NSLP)	0.395	0.432	0.354	0.078	**
	(0.014)	(0.019)	(0.019)		
Monthly unemployment rate	7.256	7.307	7.202	0.104	**
	(0.017)	(0.022)	(0.025)		
Annual per capita GDP ($1,000‘s)	53.803	53.786	53.821	-0.034	
	(0.037)	(0.053)	(0.053)		
Annual per capita income ($1,000‘s)	27.680	27.767	27.585	0.181	
	(0.048)	(0.069)	(0.066)		
Proportion of adults with some college or higher (%)	57.651	57.888	57.396	0.492	**
	(0.095)	(0.132)	(0.136)		
observations	1269	659	610		

Note: The unit of observations is the household. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ indicate significant differences in sample means between White SNAP and White NonSNAP households. Standard deviations are reported in parentheses.

Table 3.4: FAH+FAFH HEI-2010 Component Scores for Black Households

Component	Average Component Score for SNAP Participation	Average Component Score for Non-SNAP Participants	Mean difference	
	(1)	(2)	(3)	
FAH+FAFH HEI-2010	46.82	50.04	-3.22	*
	(0.76)	(1.21)		
Adequacy Components:				
Total Vegetables	2.29	2.80	-0.50	***
	(1.40)	(1.62)		
Greens and Beans	1.11	1.62	-0.51	***
	(1.59)	(1.84)		
Total Fruit	1.68	1.98	-0.30	
	(1.63)	(1.77)		
Whole Fruit	1.67	2.07	-0.40	*
	(1.82)	(1.99)		
Whole Grains	1.54	2.14	-0.60	**
	(2.12)	(2.75)		
Dairy	4.51	4.73	-0.21	
	(2.99)	(3.03)		
Total Protein Foods	4.13	4.21	-0.08	
	(1.31)	(1.23)		
Seafood and Plant Proteins	1.75	1.93	-0.18	
	(1.80)	(1.92)		
Fatty Acids	5.87	5.34	0.53	
	(3.31)	(3.32)		
Moderation Components:				
Sodium	5.47	5.31	0.16	
	(3.82)	(3.65)		
Refined Grains	6.06	6.17	-0.10	
	(3.37)	(3.65)		
Empty Calories	10.73	11.74	-1.01	
	(5.63)	(5.82)		

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ indicate significant differences in average FAH+FAFH HEI-2010 component score between Black SNAP and Black NonSNAP households. Standard deviations are reported in parentheses.

Table 3.5: The Effects of SNAP on FAH+FAFH HEI-2010 Scores Among Black and White Households with Monthly Gross Income <= 200% of the Federal Poverty Level

	OLS	IV Regression	Percentiles of FAH+FAFH HEI-2010 Score								
			10th	20th	30th	40th	50th	60th	70th	80th	90th
SNAP Black	-2.303	-14.907**	-11.169	-16.697**	-13.712**	-16.567**	-14.121**	-7.950	-9.018	-10.256	-11.236
	(1.553)	(6.703)	(7.017)	(6.787)	(5.835)	(3.925)	(6.935)	(6.396)	(6.743)	(8.430)	(9.006)
SNAP White	-2.278***	-11.852	-17.251***	-19.788***	-17.643***	-19.296***	-15.772**	-10.208*	-8.296	-9.658	-8.448
	(0.766)	(8.869)	(6.440)	(6.766)	(5.900)	(6.615)	(6.783)	(6.040)	(6.323)	(7.870)	(8.563)
First-stage F-statistic		24.155									
First-stage P-value		0.000									
J Test statistic		9.261									
J Test P-value		0.321									

Note: All regressions include controls for household, individual, and state-level characteristics. The instrumental variables are state-level welfare and SNAP administrative policies. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are in parentheses and clustered at the region level.

Table 3.6: IV Estimates of the Effects of SNAP on 12 FAH+FAFH HEI-2010 Component Scores Among Black and White Households with Monthly Gross Income <= 200% of the Federal Poverty Level

Variable	Total vegetables	Greens and Beans	Total Fruit	Whole Fruit	Whole Grains	Dairy
SNAP Black	-1.988**	-1.657	-0.219	-1.277	-0.816	-1.239
	(0.865)	(1.034)	(0.927)	(1.059)	(1.517)	(1.646)
SNAP White	-1.752	-2.284*	-0.020	-0.797	-1.595	1.981
	(1.249)	(1.299)	(1.260)	(1.430)	(1.752)	(1.975)
Variable	Total protein foods	Seafood and plant protein	Fatty Acid	Sodium	Refined Grains	Empty calories
SNAP Black	-1.005	-1.079	0.178	-2.515	-0.035	-3.254
	(0.770)	(1.027)	(1.976)	(2.275)	(1.805)	(3.167)
SNAP White	-1.220	1.439	-1.839	-3.177	1.331	-3.918
	(0.959)	(1.502)	(2.490)	(3.058)	(2.535)	(3.928)

Note: All regressions include controls for household, individual, and state-level characteristics. The instrumental variables are state-level welfare and SNAP administrative policies. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are in parentheses and clustered at the region level.

Table 3.7: IV Estimates of the Effects of SNAP on 12 FAH HEI-2010 Component Scores Among Black and White Households with Monthly Gross Income <= 200% of the Federal Poverty Level

Variable	Total vegetables	Greens and Beans	Total Fruit	Whole Fruit	Whole Grains	Dairy
SNAP Black	-2.536***	-1.180	0.065	-1.209	-1.291	-1.493
	(0.979)	(0.955)	(1.052)	(1.146)	(1.634)	(1.897)
SNAP White	-1.130	-1.580	-0.410	-1.472	-1.597	-0.553
	(1.421)	(1.441)	(1.464)	(1.698)	(2.117)	(2.495)
Variable	Total protein foods	Seafood and plant protein	Fatty Acid	Sodium	Refined Grains	Empty calories
SNAP Black	-1.234	-2.587**	-0.465	-2.941	0.364	-4.944
	(1.115)	(1.078)	(2.068)	(2.303)	(2.194)	(3.494)
SNAP White	-0.691	1.066	1.219	-3.103	0.884	-5.268
	(1.239)	(1.606)	(2.625)	(2.697)	(2.688)	(4.322)

Note: All regressions include controls for household, individual, and state-level characteristics. The instrumental variables are state-level welfare and SNAP administrative policies. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are in parentheses and clustered at the region level.

Table 3.8: IVUQR Estimates of the Effects of SNAP on 12 FAH+FAFH HEI-2010 Component Scores Among Black Households with Monthly Gross Income <= 200% of the Federal Poverty Level

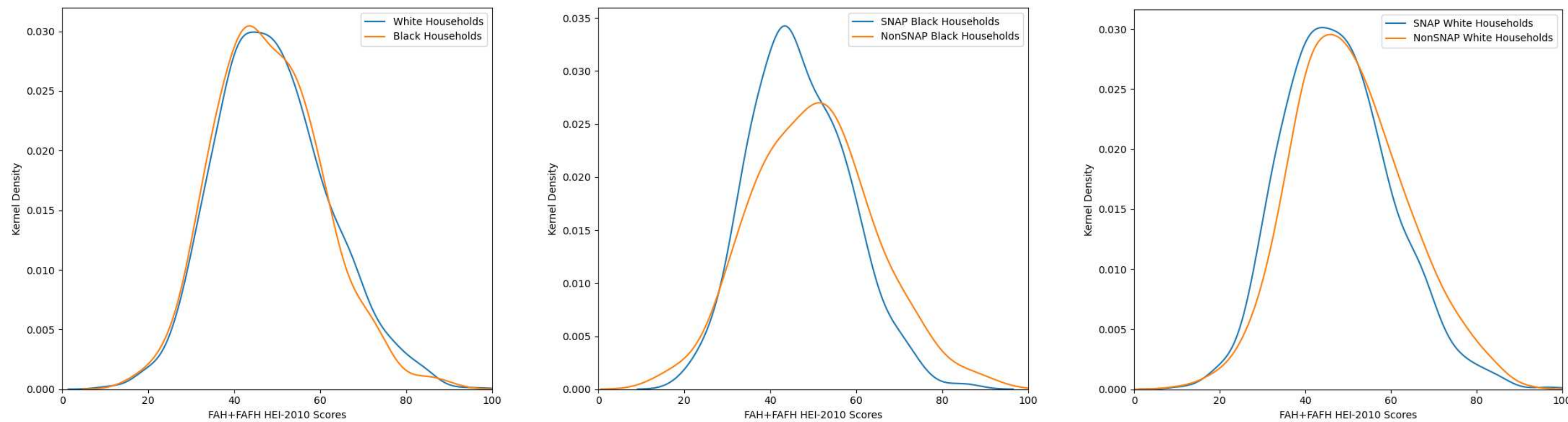
	Percentiles of FAH+FAFH HEI-2010 Component Score								
	10th	20th	30th	40th	50th	60th	70th	80th	90th
Total vegetables	-1.88**	-1.24*	-1.18	-0.84	-1.20	-2.08**	-2.97**	-1.74	0.00
	(0.78)	(0.74)	(0.74)	(0.75)	(0.82)	(0.99)	(1.31)	(1.17)	(0.00)
Greens and beans	-0.27	-0.27	-0.27	-0.37	-1.23**	-2.36*	-3.94**	-5.81*	0.00
	(0.23)	(0.23)	(0.23)	(0.24)	(0.60)	(1.22)	(1.85)	(3.00)	(0.00)
Total fruits	-0.12	-0.01	-0.15	-0.53	-0.40	-0.00	0.45	0.16	0.27
	(0.27)	(0.30)	(0.59)	(0.78)	(0.86)	(1.02)	(1.26)	(1.90)	(0.79)
Whole fruits	-0.25	-0.35	-0.59	-1.59	-2.26	-1.28	-1.43	-1.14	0.00
	(0.28)	(0.29)	(0.53)	(1.07)	(1.44)	(1.71)	(2.13)	(1.57)	(0.00)
Whole grains	-0.20	-0.20	-0.20	-0.58	-1.06	-1.65	-2.17	-2.71	-7.51*
	(0.26)	(0.26)	(0.26)	(0.50)	(0.79)	(1.08)	(1.60)	(2.32)	(4.08)
Dairy	0.09	0.29	-0.27	0.07	-1.34	-1.48	-0.58	0.91	0.00
	(1.49)	(1.80)	(1.79)	(1.78)	(1.73)	(1.83)	(2.27)	(2.24)	(0.00)
Total protein foods	-1.85	-1.00	-1.14	-0.75	-0.19	0.00	0.00	0.00	0.00
	(1.40)	(1.27)	(1.45)	(1.18)	(0.36)	(0.00)	(0.00)	(0.00)	(0.00)
Seafood and plant proteins	0.21	0.27	0.40	0.42	0.49	0.93	-1.42	-1.78	0.00
	(0.27)	(0.28)	(0.46)	(0.88)	(1.14)	(1.60)	(2.17)	(1.76)	(0.00)
Fatty acids	-0.29	0.05	0.84	0.62	0.20	0.83	0.36	0.91	0.00
	(0.52)	(2.22)	(1.78)	(1.65)	(2.00)	(2.52)	(2.91)	(3.59)	(0.00)
Sodium	-0.96	-4.86	-2.43	-1.33	0.61	0.30	-0.76	0.00	0.00
	(0.61)	(3.36)	(3.34)	(2.38)	(2.00)	(2.16)	(2.49)	(0.10)	(0.00)
Refined grains	0.81	4.60	0.30	-0.15	0.33	1.19	-1.15	0.00	0.00
	(0.53)	(3.20)	(2.66)	(2.94)	(2.33)	(1.83)	(2.38)	(0.06)	(0.00)
Empty calories	-5.61	-5.51	-5.57*	-5.31	-5.96*	-3.99	-4.69	-4.05	-1.51
	(6.67)	(4.10)	(3.38)	(3.46)	(3.17)	(2.84)	(3.08)	(3.57)	(4.28)

Note: All regressions include controls for household, individual, and state-level characteristics. The instrumental variables are state-level welfare and SNAP administrative policies. *** p < 0.01, ** p < 0.05, * p < 0.1. The standard errors are in parentheses and clustered at the region level.

Table 3.9: IVUQR Estimates of the Effects of SNAP on 12 FAH HEI-2010 Component Scores Among Black Households with Monthly Gross Income <= 200% of the Federal Poverty Level

	Percentiles of FAH HEI-2010 Component Score								
	10th	20th	30th	40th	50th	60th	70th	80th	90th
Total vegetables	-0.24	-0.78	-2.35**	-1.79*	-1.56	-3.00**	-3.13*	-1.01	0.00
	(0.38)	(0.69)	(0.98)	(1.03)	(1.00)	(1.24)	(1.72)	(1.02)	(0.00)
Greens and beans	-0.12	-0.12	-0.12	-0.12	-0.12	-0.12	-2.37	-3.67	-0.00
	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(2.48)	(2.68)	(0.24)
Total fruits	0.01	0.01	-0.00	0.26	-0.64	-0.20	-0.68	1.11	0.00
	(0.27)	(0.27)	(0.34)	(0.95)	(1.10)	(1.32)	(1.74)	(2.46)	(0.00)
Whole fruits	-0.40	-0.40	-0.40	-0.46	-2.24	-3.27	-3.72	0.00	0.00
	(0.25)	(0.25)	(0.25)	(0.61)	(1.69)	(2.22)	(2.70)	(0.57)	(0.00)
Whole grains	-0.28	-0.28	-0.28	-0.28	-2.14*	-2.81*	-3.60	-5.13	-7.41
	(0.21)	(0.21)	(0.21)	(0.21)	(1.11)	(1.56)	(2.23)	(3.25)	(5.77)
Dairy	-0.28	-1.36	-0.30	-0.37	-0.67	-3.00	-1.79	-0.31	0.00
	(0.57)	(2.02)	(2.36)	(2.28)	(2.15)	(2.54)	(2.95)	(0.97)	(0.00)
Total protein foods	-0.64	-1.98	-1.57	-0.43	0.65	0.00	0.00	0.00	0.00
	(0.74)	(2.02)	(1.94)	(1.92)	(1.52)	(0.14)	(0.00)	(0.00)	(0.00)
Seafood and plant proteins	-0.11	-0.11	-0.11	-0.13	-0.39	-1.05	-2.94	-1.41	0.00
	(0.23)	(0.23)	(0.23)	(0.27)	(1.34)	(1.98)	(2.57)	(1.62)	(0.00)
Fatty acids	0.02	0.09	1.15	0.70	0.47	1.64	2.13	0.00	0.00
	(0.39)	(0.46)	(2.51)	(2.19)	(2.66)	(3.38)	(3.79)	(0.28)	(0.00)
Sodium	-0.59	-5.01	-2.18	-0.43	-0.48	-0.21	0.00	0.00	0.00
	(0.53)	(5.28)	(3.28)	(3.10)	(2.74)	(1.60)	(0.00)	(0.00)	(0.00)
Refined grains	0.88*	7.03	0.82	0.92	1.77	1.15	0.00	0.00	0.00
	(0.46)	(4.50)	(3.46)	(3.84)	(2.72)	(2.45)	(0.01)	(0.00)	(0.00)
Empty calories	-0.27	-3.87	-6.29	-7.58*	-7.06*	-11.72**	-10.16**	-4.82	0.00
	(0.54)	(4.73)	(5.20)	(4.51)	(4.09)	(4.78)	(4.04)	(3.61)	(0.00)

Note: All regressions include controls for household, individual, and state-level characteristics. The instrumental variables are state-level welfare and SNAP administrative policies. *** p < 0.01, ** p < 0.05, * p < 0.1. The standard errors are in parentheses and clustered at the region level.



(a) White and Black Households

(b) Black Households by SNAP Status

(c) White Households by SNAP Status

Figure 3.1: Distributions of FAH + FAFH HEI-2010 Scores by Household Race and SNAP Participation

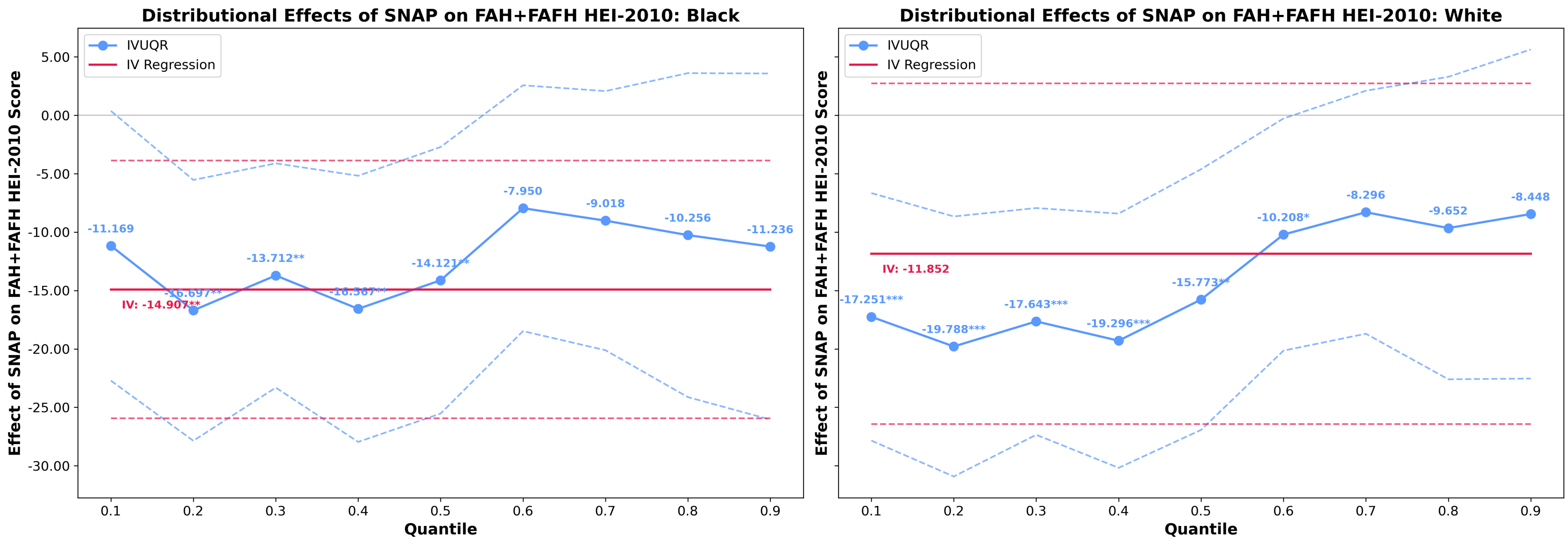


Figure 3.2: The Distributional Effects of SNAP on FAH+FAFH HEI-2010 Scores Among Black and White Households with Monthly Gross Income <= 200% of Federal Poverty Level

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are bootstrapped and clustered at the regional level. Confidence intervals are based on a 90% confidence level.

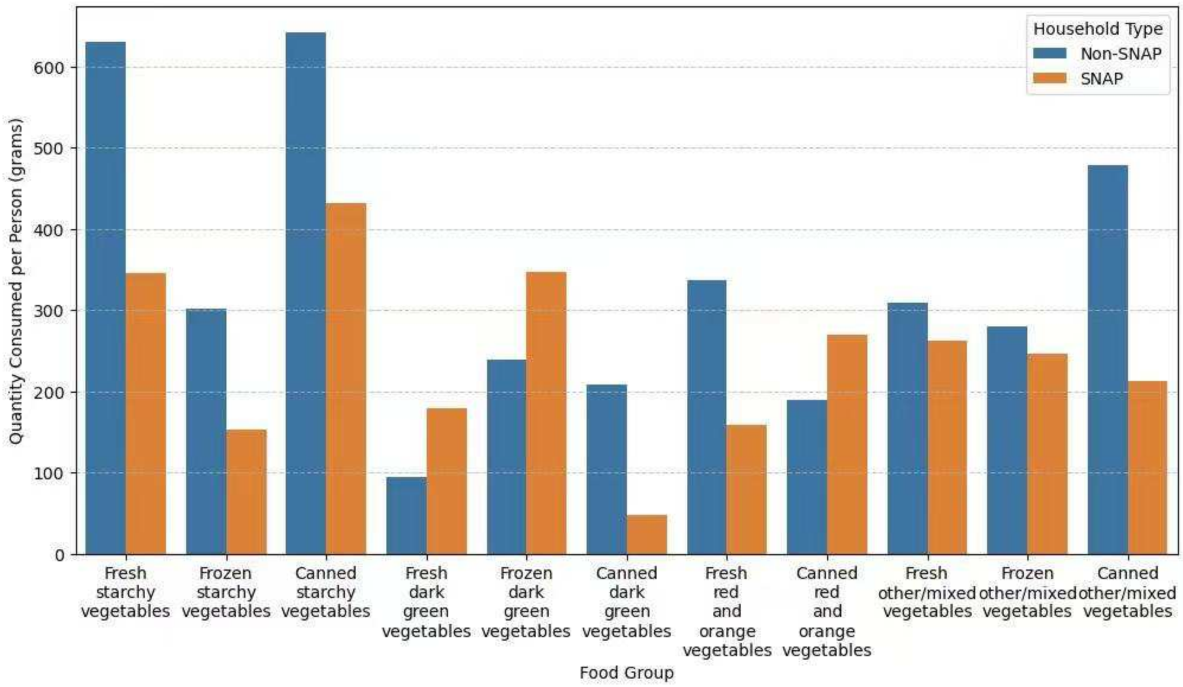


Figure 3.3: Comparison of Vegetable Consumption from FAH between Black SNAP and Black non-SNAP Households

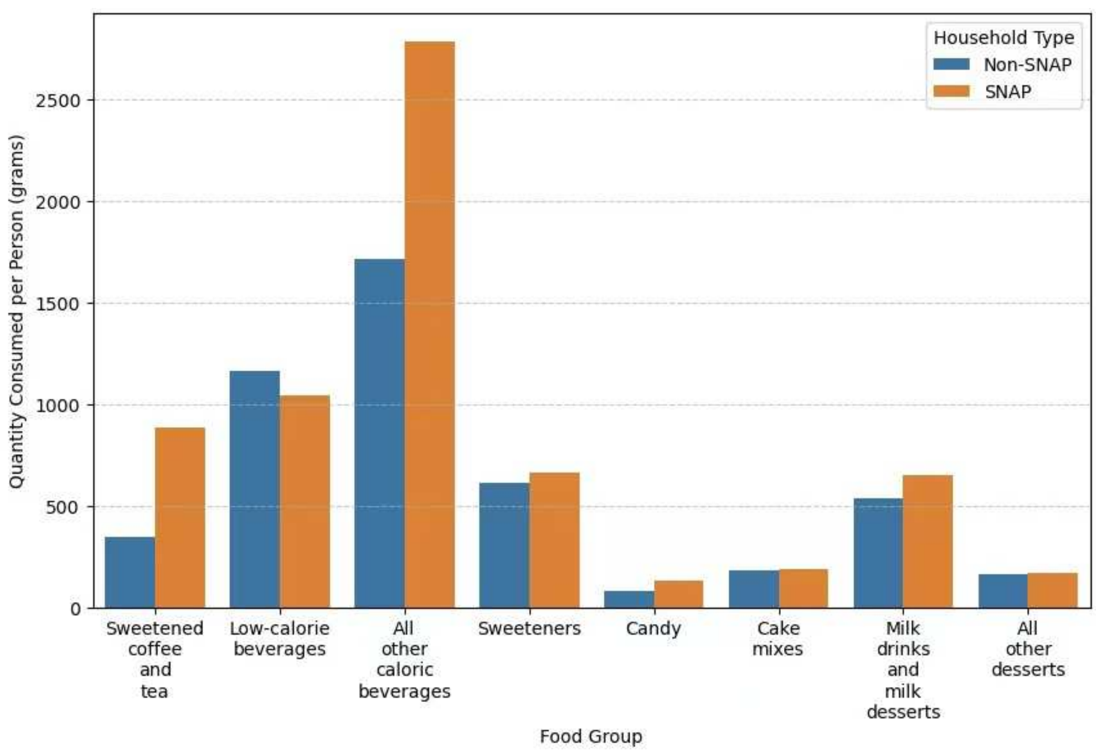


Figure 3.4: Comparison of Sugary Food and Beverage Consumption from FAH between Black SNAP and Black non-SNAP Households

Chapter 4 Evaluating the Effect of an Industry-wide Sugar Content Reduction on the U.S. Sugar-Sweetened Beverage Market [12]

4.1 Introduction

The prevalence of obesity among the adult U.S. population has increased from 30.5% in 1999-2000 to 41.9% in 2017-2020 (Centers for Disease Control and Prevention, 2021), making it a critical public health concern. Obesity is associated with several adverse health outcomes, including diabetes, metabolic syndrome, and cardiovascular disease (Malik et al., 2010; Gao et al., 2007; Koning et al., 2012). The rise in the consumption of sugar-sweetened beverages (SSBs) is often cited as a contributing factor to the increase in obesity rates (Lopez and Fantuzzi, 2012; Falbe et al., 2016), as it far exceeds the American Heart Association's (AHA) recommendations.[13] Despite a slight decline in soft drink consumption, SSBs remain the largest contributor to added sugar intake in the U.S. diet (Haeck et al., 2022; Valizadeh et al., 2020; Malik and Hu, 2012; Welsh et al., 2011). In 2023, Americans consumed an average of 15 ounces of soft drinks daily, far exceeding the AHA's recommendation of limiting sugary beverage consumption to no more than 8 ounces per week. Among the various interventions considered to curb the high consumption of SSBs,

12. The findings and conclusions in this dissertation are those of the author and should not be construed to represent any official USDA or U.S. Government determination or policy. The analysis, findings, and conclusions expressed in Chapter 3 of this dissertation should also not be attributed to Circana (formerly IRI). Circana data access was granted via a Third Party Access Agreement between Colorado State University and the Economic Research Service of the United States Department of Agriculture, currently pending renewal.

13. From 1999 to 2008, mean intakes were approximately 22 ounces per day, with 10% of individuals consuming at least 45 ounces daily (Leung et al., 2014),

policymakers have focused their efforts on taxing sugary drinks.[14] SSB taxes are often applied to added sugars, not natural sugars, so to exempt natural juices and dairy drinks from taxation.

As the goal of "soda taxes" is to reduce consumption among high-risk households, typically with high SSB consumption, SSB taxes may be less effective in reducing demand among their target group. In fact, tax-induced reduction in SSB consumption, may disproportionately affect low SSB-consuming households, who are more sensitive to price changes. Households with high SSB consumption may exhibit reduced price sensitivity especially among lower-income brackets, and may even increase consumption shortly after the implementation of the tax (Falbe et al., 2016; Debnam, 2017; Valizadeh and Ng, 2021). Additionally, excise taxes may be too small to significantly discourage SSB consumption (Zhen et al., 2014), and may prompt consumers to switch to other calorie-rich, untaxed beverages (Fletcher et al., 2010), and some consumers (particularly low-income households) may bypass taxation by purchasing SSBs outside of the taxed jurisdiction, (Cawley and Frisvold, 2017).

Interestingly, widespread improvements in the average diet quality of U.S. consumers seem instead to be largely attributable to "supply-side" product reformulation rather than changes in consumer behavior or health habits (Beatty et al., 2014). Rojas and Jaenicke (2024) identifies a significant decline in the proportion of products containing sugar beginning in 2018, the year several SSB taxed were implemented. Specifically, the proportion of such products decreased from approximately 75% to just over 50% between 2018 and 2021. However, while sugar content per grams of product decreased from 2010 to 2017, it began increasing again from 2017 to 2021, hinting that, without appropriate oversight, manufacturers may revert to higher sugar formulations. While voluntary reformulation represents a promising approach, its effectiveness in reducing sugar intakes may be limited if consumer preferences remain oriented toward higher-sugar alternatives. Cengiz and Rojas (2024) find that, between 2007 and 2015, if consumers continued purchasing

14. Several jurisdictions have approved SSB taxes: Boulder, Colorado (2017); Philadelphia, Pennsylvania (2017); Seattle, Washington (2017); and four California cities: Albany (2017), Berkeley (2015), Oakland (2016), and San Francisco (2016).

their 2007 shopping baskets, voluntary reformulation and product turnover (the withdrawal of high-sugar products and/or the introduction of low-sugar products) could have resulted in a reduction in consumer sugar intake by 52.8%. However, as consumers responded to reformulations by switched to products with higher sugar content, the actual decrease in sugar intake was only 15.6%.

This study evaluates the potential impact of an industry-wide sugar content reduction mandate on equilibrium outcomes in the U.S. SSB market. To do so, I simulate changes in product market shares and per capita sugar consumption using a random coefficients logit model developed by Berry et al. (1995). The analysis integrates point-of-sale scanner data from Circana, detailed product-level nutritional information, and demographic data from the U.S. Census. This study makes two key contributions. First, it examines the impact of product reformulation on consumer sugar purchased. Second, it explores the feasibility of implementing a sugar content ceiling from an industry perspective, assessing whether such a policy could be adopted without compromising industry profitability. The results indicate that consumer demand is downward-sloping with respect to price. Higher sugar and sodium contents reduce utility, while caffeine slightly increases it. Simulation results suggest that sugar reformulation policies could reduce average sugar purchased per 12-ounce serving. The baseline value of 13.07 grams falls to 11.30 grams, 10.26 grams, and 9.17 grams under the 10%, 20%, and 30% sugar reduction scenarios, respectively. These correspond to reductions of 13.5 percent, 21.5 percent, and 29.9 percent. Market shares shift only modestly, implying that such policies could achieve meaningful health improvements without requiring drastic changes in consumer preferences or market dynamics.

The remainder of this paper is organized as follows: Section 4.2 describes the Circana Data employed in this study. Section 4.3 presents the methodology, focusing on the BLP random coefficients logit model and covering demand-side, supply-side, and simulation techniques. Section 4.4 discusses the results, and Section 4.5 concludes with a discussion of our findings.

4.2 Data

In this analysis, I use data from four different sources. The primary dataset is drawn from Circana Point of Sales (PoS) scanner data, which offers detailed Universal Product Code (UPC) level weekly retail sales figures and quantities sold for the years 2016 and 2017. The analysis focuses on three beverage categories: carbonated soft drinks, sports drinks, and fruit beverages. To ensure consistency in the analysis, retail prices are standardized as average weekly prices, expressed per 12 oz serving. Markets are defined as unique combinations of state and month.

In addition to sales data, the study incorporates nutrient information from the Purchases to Plate Crosswalk (PPC) dataset. This dataset links Circana UPCs to USDA food composition databases, enabling the inclusion of detailed nutritional profiles for grocery items.

To account for demographic and socioeconomic heterogeneity, the study incorporates state-level variables for 2017 from the U.S. Census Bureau's American Community Survey Public Use Microdata Sample (ACS PUMS). These variables include household income, marital status, and gender.

Finally, to address potential endogeneity in retail prices, instrumental variables are employed. Specifically, I employ weekly state-level wages in the retail sector from the U.S. Bureau of Labor Statistics and monthly national sugar prices from the USDA Economic Research Service.

4.2.1 Circana Data

To investigate the impact of regulations on sugar content in SSBs on consumer purchasing behavior, I utilize Circana Data, formerly kown as Information Resource Incorporated (IRI) InfoScan Retail Data. This comprehensive dataset provides store-based scanner data, including weekly retail sales figures for both revenue and quantity, at UPC level. The specificity of UPCs is noteworthy, as even identical food items may possess distinct UPCs due to variations in size, flavor, or packaging materials. Furthermore, identical packages from different retail outlets might carry unique UPCs. My investigation focuses on the top 70% of volume sold for SSBs, including carbonated soft drinks (both regular and diet varieties), sports drinks, and fruit beverages. As these products are sold in various package sizes (8, 12, and 16 ounces per unit and in packages of 1, 6, 12, and 24 packs), I standardize their weekly prices and their nutritional content to refer to a 12-ounce serving. This standardization is used solely for normalization purposes, as 12 ounces corresponds to the size of a typical can of carbonated soft drink.

Each market is defined as a unique combination of state and month. The total market size is quantified in terms of 12-ounce servings consumed per month, calculated by multiplying the population by the per capita monthly consumption, where consumption is expressed in these standardized units. Market shares are determined by dividing the volume sold (in 12-ounce servings) by the total market size. The market share for the outside good is defined as one minus the sum of the observed market shares for all products in the choice set. To ensure a balanced panel and comparability across markets, only those markets where all product types are sold are included in the analysis.

Circana Data provides information at the store-week level, while my analysis is conducted at the state-month level. This difference in data levels brings challenges in aggregation because products are sampled differently across stores and weeks. For example, in a given state and week, data on carbonated soft drinks may come from fewer stores than data on sports drinks, which could lead to misleading results if sales are simply added together. Furthermore, although information for carbonated soft drinks and sports drinks is available in every state and month, data on fruit drinks appears only in some months and states. To address these differences, I use a multi-step aggregation process in which every measurement is based on a 12-ounce serving. First, I calculate the average volume sold per store for each product during each week. I then combine these weekly averages to form a monthly average for each state. To estimate the total monthly volume sold for each product, I multiply the monthly average by four, assuming four weeks in a month. The same process is applied to standardize product prices.

To account for variations over time and across regions, I include fixed effects for month, year, and state in my analysis. These fixed effects control for seasonal fluctuations (such as increased carbonated soft drink sales in summer), year-specific trends (like price increases for some products in late 2017), and state-specific factors (including higher sales volumes and prices in states with larger populations and stronger economies).

My research focuses on SSB consumption in the United States during 2016 and 2017. I selected this period for two main reasons. First, I aim to investigate consumer reactions to sugar content mandates during a non-pandemic period. The COVID-19 pandemic, which emerged in December 2019, significantly altered purchasing behaviors through lockdown restrictions, shifts to online ordering, and heightened health concerns (Ulpiano J. Vázquez-Martínez et al., 2021). These changes affected beverage choices in ways not directly related to sugar policies. By concentrating on the pre-pandemic years, my study isolates consumer responses to sugar-reduction policies without the confounding effects introduced by the pandemic. Second, manufacturers were required to implement the new nutrition facts panel in 2016. This mandate ensured that consumers were informed about the added sugar content in the products they purchased (U.S. Food and Drug Administration, 2016). By studying consumer behavior during this period, I can assess how the availability of added sugar information on nutrition labels influenced their purchasing decisions.

4.2.2 Purchases to Plate Crosswalk Data

The Circana Data provides detailed information on purchases, prices, and stores but lacks detailed nutrient data. To address this, ERS researchers collaborated with the USDA Center for Nutrition Policy and Promotion (CNPP) and the USDA Agricultural Research Service (ARS) to create the Purchases to Plate Crosswalk (PPC) [15]. The PPC links grocery items from the Cricana Product Dictionary (PD) with food items in the USDA Food and Nutrient Database for Dietary Studies (FNDDS), National Nutrient Database for Standard Reference (SR), and the Food Pattern Equivalents Databases (FPED/FPID). I use the PPC as the primary source for nutrient data in my analysis. To ensure comparability across different product sizes and formulations, I standardize all

15. For more information about the Purchases to Plate Crosswalk Data, please refer to this link.

nutrient data to a 12-ounce serving.

4.2.3 American Community Survey Public Use Microdata Sample

The Public Use Microdata Sample (PUMS) is released by the U.S. Census Bureau and provides detailed individual and household-level information collected through the American Community Survey (ACS). The ACS is a continuous survey that gathers data on a wide range of demographic, social, economic, and housing characteristics of the U.S. population. In this study, I focus on the period from 2016 to 2017 by employing individual-level PUMS data aggregated at the state level. Although my analytical period covers two years, I exclusively use data from 2017 under the assumption that population demographics remain relatively stable over short intervals. The analysis incorporates the following variables from the ACS PUMS dataset: the inverse of household income, marital status, and gender. The inclusion of inverse household income reflects the economic principle that lower income households are more price sensitive, particularly regarding SSBs. Empirical evidence supports this, showing that lower-income households significantly reduce SSB consumption in response to price increases, such as those induced by taxation (Falbe et al., 2016). Marital status is incorporated to account for differences in household composition and health-conscious purchasing behavior. Studies indicate that married individuals may exhibit different consumption patterns for SSBs compared to unmarried individuals (Lundeen et al., 2018). The gender indicator captures taste heterogeneity, as research has found that men are significantly more likely than women to consume SSBs daily (Lundeen et al., 2018) and may respond differently to sugar reduction initiatives.

To implement the numerical integration required by Equation 25 of the BLP model, I employ random draws from the ACS PUMS dataset. Consistent with the approach outlined by Berry et al. (1995), I use identical draws across all markets. As Nevo (2001) suggests, using actual individual data is preferable to imposing arbitrary distributional assumptions when prior information on the demographic distribution is limited. Moreover, Reynaert and Verboven (2014) recommend using more than 200 draws for Monte Carlo integration to reduce estimation bias. Following these guidelines, I sample 200 households from the ACS PUMS data and apply the same set of draws consistently across each market.

4.2.4 Instrumental Variables

Estimating the demand for SSBs involves addressing endogeneity that arises from the correlation between price and unobserved product characteristics. Local advertising and promotions may affect both price and demand, yet these factors are not captured in the data (Zhang and Palma, 2021). To mitigate this issue, I use two instrumental variables for price. The first is the average weekly wage in the retail sector, measured in units of $1,000, by state and quarter from the U.S. Bureau of Labor Statistics (U.S. Department of Labor, Bureau of labor statistics: Quarterly census of employment and wages, 2020). The second is the monthly national cost of sugar from the USDA Economic Research Service (U.S. Department of Agriculture, Economic Research Service., 2024a), standardized to reflect the cost for 39 grams of sugar, the amount typically contained in a 12 oz regular carbonated soft drink. The retail wage data include only private retail establishments, which is consistent with the data available in the Circana dataset.

The average weekly wage captures changes in retailer costs: when wages rise, retailers face higher operating expenses, which leads to higher prices. These wage changes reflect broad economic conditions and do not directly influence consumer taste for SSBs. Similarly, the national cost of sugar affects production costs because sugar is a key ingredient in SSBs. Sugar prices, driven by global market forces, change beverage prices without impacting local consumer demand. Thus, both instruments influence prices through cost channels while remaining unrelated to the unobserved factors that determine consumer preferences. I do not use Hausman instruments, which rely on the prices of the same product in other markets, because local demand shocks can spread across geographically proximate markets due to factors such as advertising, promotions, or consumer mobility. The assumption that these demand shocks are independent of the prices in neighboring markets may not hold in practice, especially when the markets share similar characteristics.

4.3 Methods

In this analysis, I use the random coefficients logit model introduced by Berry et al. (1995), henceforth BLP. On the demand side, the model allows consumer taste parameters to vary across individuals. This variation helps address the Independence of Irrelevant Alternatives (IIA) problem and leads to realistic substitution patterns among products. The BLP model is widely used in markets with differentiated products. For example, it has been applied to study breakfast cereals (Nevo, 2001; Chidmi and Lopez, 2007), ketchup (Rennhoff, 2008), ready-to-heat meals (Zhang and Gallardo, 2022), instant packaged noodles (Chen and Zhen, 2022), milk products (Hirsch et al., 2018; Lopez and Lopez, 2009), and canned tuna (Shreay et al., 2015). In the soft drinks market, Lopez and Fantuzzi (2012) used this model to analyze the effects of a caloric content-based tax, Zhang and Palma (2021) examined changes in sugary drinks consumption in Berkeley following a tax, and Lopez et al. (2015) evaluated the impact of television advertising on demand.

On the supply side, firms set prices to maximize their profits in a competitive market. Each firm faces the demand conditions described by the random coefficients logit model. Firms choose prices based on product characteristics and production costs, which may vary with factors such as added sugar content. The pricing decisions follow from the first-order conditions of profit maximization, yielding a system of equations that determine the equilibrium prices. In this study, I simulate a scenario where the industry voluntarily complies with a predetermined added sugar content ceiling. This setup lets me examine the effects of the voluntary standard on market shares and the overall amount of added sugar purchased by consumers.

The estimation of the random coefficients logit demand model was conducted in Python using the pyblp package (Conlon and Gortmaker, 2020), a widely adopted open-source routine designed to implement the BLP framework. The package enables the recovery of both mean utility parameters and individual-level taste heterogeneity through a generalized method of moments estimator. In addition to parameter estimation, pyblp computes own- and cross-price elasticities of demand based on the estimated random coefficients distribution. The package also includes built-in functionality for conducting counterfactual simulations, which I used to evaluate changes in market outcomes under different sugar content ceilings. Specifically, I simulated market shares and prices under alternative product formulations by leveraging the equilibrium-solving routines within pyblp. After estimating the structural parameters of the model, including mean utilities, random coefficients, marginal costs, and firm markups, I retained all parameter estimates and product characteristics unchanged, except for the sugar content variable. To simulate the policy scenario in which a sugar content ceiling is imposed, I multiplied the sugar content of affected products by a sugar reduction factor to reflect reformulated versions of those products. These modified product characteristics were then passed to the pyblp simulation routine, which recomputes equilibrium market shares by solving for the new set of prices and consumer demand given the updated product attributes. For further details on implementation, see the pyblp user manual.

4.3.1 The Demand Side

Consider $t = 1,\ldots,T$ markets, with each market t defined as a unique combination of a geographic area and time period (i.e. state-month combinations). Within each market, I observe $i = 1,\ldots,I_t$ consumers and $j = 1,\ldots,J$ products. The indirect utility u_{ijt} obtained by consumer i from consuming a product j in market t can be expressed as $u_{ijt} = U(x_{jt}, p_{jt}, \xi_{jt}, \tau_i; \theta)$. This utility function incorporates observed x_{jt}, and unobserved product characteristics ξ_{jt}, a product's price p_{jt}, individual consumer characteristics τ_i and a vector of unknown parameters θ, and takes the form:

$$u_{ijt} = -\alpha_i p_{jt} + x_{jt}\beta_i + \xi_{jt} + \varepsilon_{ijt} \quad \text{for } i = 1,\ldots,I, j = 1,\ldots,J, t = 1,\ldots,T \tag{17}$$

where x_{jt} is a $1 \times K$ vector of observable product characteristics, such as added sugar, caffeine, sodium and binary indicators for diet and low-calorie formulations; ξ_{jt} captures unobservable product attributes; ε_{ijt} is an idiosyncratic error term assumed to be independently and identically distributed (i.i.d) with zero mean; α_i represents the price sensitivity of consumer i, representing how changes in p_{jt} affect their utility; β_i is a vector of parameters capturing the taste for the observable product characteristics x_{jt}. The specification of the utility function also includes category-specific dummy variables to account for potential variations across product types. The unobserved product characteristics' term, is decomposed as $\xi_{jt} = \xi_j + \xi_s + \xi_m + \xi_y + \Delta\xi_{jt}$, where ξ_j represents brand-specific unobservables, ξ_s, ξ_m ,and ξ_y are unobservables at the state, month and year-level, which are captured by brand-level, and market-level fixed effects, respectively. The term $\Delta\xi_{jt}$ serves as the econometric error term, accounting for brand-market specific variations.

Consumers choose either one unit of a product from the available choice set, or opt for an alternative outside the choice set (the "outside option"), whichever leads to the highest utility level. The outside option is denoted as $j = 0$, and the resulting indirect utility for consumer i at time t is $u_{i0t} = \varepsilon_{i0t}$ (Nevo, 2001), which is normalized to zero.

Consumer's heterogeneity is assumed to be function of both observable (demographic) characteristics D_i and unobservable characteristics v_i. The distribution of taste parameters is assumed taking the following form:

$$\begin{aligned} \begin{pmatrix} \alpha_i \\ \beta_i \end{pmatrix} &= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \Pi D_i + \Sigma v_i \qquad D_i \sim P_D(D), \qquad v_i \sim P_v(v) \\ &= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} \Pi_\alpha \\ \Pi_\beta \end{pmatrix} D_i + \begin{pmatrix} \Sigma_\alpha \\ \Sigma_\beta \end{pmatrix} v_i \end{aligned} \tag{18}$$

where D_i is a $d \times 1$ vector of consumer characteristics, such as income, age and race. The matrix Π, with dimensions $(K+1) \times d$, captures the effect of observed consumer characteristics on α_i and the K elements of β_i. v_i is a $(K+1) \times 1$ vector of unobserved consumer characteristics; Σ is a $(K+1) \times (K+1)$ matrix of parameters associated with them. $P_D(D)$, represents the empirical distribution of consumers' observable characteristics, and $P_v(v)$ is the distribution of unobserved consumer characteristics, which is assumed to follow a multivariate standard normal distribution.

Combining Equations (17) and (18), I obtain:

$$\begin{aligned} u_{ijt} &= \delta_{jt}(x_{jt}, p_{jt}, \xi_{jt}; \theta_1) + \mu_{ijt}(x_{jt}, p_{jt}, v_i, D_i; \theta_2) + \varepsilon_{ijt} \\ \delta_{jt} &= x_{jt}\beta - \alpha p_{jt} + \xi_{jt} \\ \mu_{ijt} &= [-p_{jt}, x_{jt}](\Pi D_i + \Sigma v_i) \end{aligned} \tag{19}$$

where δ_{jt} is the mean utility derived from the product attributes, constant across consumers; μ_{ijt} represents the mean utility arising from consumer heterogeneity; ε_{ijt} is i.i.d error term. The term $\mu_{ijt} + \varepsilon_{ijt}$ constitutes a mean-zero heteroskedastic deviation from the mean utility δ_{jt}, effectively capturing the role of consumers' taste heterogeneity. Let $\theta = (\theta_1, \theta_2)$ denote a vector encompassing all model parameters. Specifically, $\theta_1 = (\alpha, \beta)$ contains the parameters being part of the mean utility, whereas $\theta_2 = (\Pi, \Sigma)$ comprises nonlinear parameters.

Consumers are assumed to purchase one unit of the good that delivers the highest utility. In the model, an individual is characterized by a vector that includes observable demographics D_i and unobservable product-specific shocks ν_i as well as $\varepsilon_{i0t}, \varepsilon_{i1t}, \ldots, \varepsilon_{iJt}$. This characterization implicitly defines the set of individual attributes that lead to the choice of a particular product j. More formally, define the set of consumers who choose product j in market t as

$$A_{jt}(x_{\cdot t}, p_{\cdot t}, \delta_{\cdot t}; \theta_2) = \Big\{(D_i, \nu_i, \varepsilon_{i0t}, \ldots, \varepsilon_{iJt}) : u_{ijt} \geq u_{ilt} \quad \forall l = 0, 1, \ldots, J\Big\} \tag{20}$$

where the vectors

$$x_{\cdot t} = \begin{pmatrix} x_{1t} \\ \vdots \\ x_{Jt} \end{pmatrix}, \quad p_{\cdot t} = \begin{pmatrix} p_{1t} \\ \vdots \\ p_{Jt} \end{pmatrix}, \quad \text{and} \quad \delta_{\cdot t} = \begin{pmatrix} \delta_{1t} \\ \vdots \\ \delta_{Jt} \end{pmatrix}$$

represent the observed characteristics, prices, and mean utilities of all products, respectively.

Assuming no ties, the market share of product j in market t is the probability that it provides the highest utility. So, the market share is given by

$$s_{jt}(x_{\cdot t}, p_{\cdot t}, \delta_{\cdot t}; \theta_2) = \int_{A_{jt}} dP^*(D, \nu, \varepsilon) \tag{21}$$

where $P^*(\cdot)$ denotes the joint population distribution of (D, ν, ε). By applying Bayes' rule, I can decompose this integral as

$$
\begin{aligned}
s_{jt}(x_{\cdot t}, p_{\cdot t}, \delta_{\cdot t}; \theta_2) &= \int_{A_{jt}} dP^*(\varepsilon \mid D, \nu)\, dP^*(\nu \mid D)\, dP^*_D(D) && (22)\\
&= \int_{A_{jt}} dP^*_\varepsilon(\varepsilon)\, dP^*_\nu(\nu)\, d\hat{P}^*_D(D) && (23)
\end{aligned}
$$

Here, the equality in Equation 22 is obtained by applying Bayes' rule to factor the joint distribution. The final expression in Equation 23 results from the independence assumptions imposed in the model.

The term ε_{ijt} is assumed to be distributed Type-I extreme value, so that the market share of product j for consumer i is given by:

$$
s_{ijt} = \frac{e^{\delta_{jt}+\mu_{ijt}}}{1+\sum_{k=1}^{J} e^{\delta_{kt}+\mu_{ikt}}} \tag{24}
$$

The overall market share of product j in market t is obtained integrating individual consumers' market share function in Equation 24 across all consumer types:

$$
\begin{aligned}
s_{jt} &= \int_v \int_D s_{ijt} dP_D(D) dP_v(v)\\
&= \int_v \int_D \left[\frac{e^{\delta_{jt}+\mu_{ijt}}}{1+\sum_{k=1}^{J} e^{\delta_{kt}+\mu_{ikt}}} \right] dP_D(D) dP_v(v) && (25)
\end{aligned}
$$

Since market shares are determined by a combination of the mean utility of each product and the distribution of consumer taste paramters for the attributes in each products, heterogeneity in consumer preferences and the availability of substitutes.

The expressions of own- and cross-price elasticity formulas (with respect to the price of product j) are:

$$
\begin{aligned}
\eta_{jjt} &= \frac{\partial s_{jt}}{\partial p_{jt}} \frac{p_{jt}}{s_{jt}} = -\frac{p_{jt}}{s_{jt}} \int_v \int_D \alpha_i s_{ijt}(1-s_{ijt}) dP_D(D) dP_v(v) && (26)\\
\eta_{jkt} &= \frac{\partial s_{jt}}{\partial p_{kt}} \frac{p_{kt}}{s_{jt}} = \frac{p_{kt}}{s_{jt}} \int_v \int_D \alpha_i s_{ijt} s_{ikt} dP_D(D) dP_v(v) && (27)
\end{aligned}
$$

4.3.2 Simulation of Market Share Changes

To evaluate the effects of sugar content reformulation on market outcomes, I simulate counterfactual equilibria in which products reduce their sugar content by a fixed percentage. A key step in this simulation involves estimating marginal costs, which are not observed but can be recovered using observed prices and estimated demand elasticities. This is done using the Nevo Menu approach (Nevo, 2001), based on firms' profit-maximizing behavior in a differentiated product market.

On the supply side, assume there are F firms, with each firm f producing a subset $\mathscr{J}_f$ of the total J products in the market. The profit for firm f in market t is given by

$$\pi_{ft} = \sum_{j \in \mathscr{J}_f} \left[(p_{jt} - mc_{jt}) M s_{jt}(p_t) - C_{jt} \right] \tag{28}$$

where p_{jt} is the price of product j in market t; mc_{jt} is the constant marginal cost of producing product j; M denotes the market size; $s_{jt}(p_t)$ is the market share of product j, and C_{jt} is the fixed cost associated with product j.

Under a multi-product Bertrand-Nash equilibrium, each firm maximizes its profit by setting the prices for all products it offers while taking competitors' prices as given. The first-order conditions for the profit maximization problem with respect to the price of product j is

$$\frac{\partial \pi_{ft}}{\partial p_{jt}} = s_{jt}(p_t) + \sum_{r \in \mathscr{J}_f} (p_{rt} - mc_{rt}) \frac{\partial s_{rt}(p_t)}{\partial p_{jt}} = 0 \tag{29}$$

The first-order conditions from Equation 28 can be rearranged into matrix form and inverted to solve for marginal costs. The resulting expression, known as the Nevo Menu approach, is:

$$mc_t = p_t - \underbrace{\Delta_t(p_t)^{-1} s_t(p_t)}_{\eta_t(p_t, s_t, \theta_2)} \tag{30}$$

where p_t is the vector of observed prices, and $s_t(p_t)$ is the vector of estimated shares. $\eta_t(p_t, s_t, \theta_2)$ is the Bertrand-Nash markup vector (dimension $J \times 1$), and $\Delta_t(p_t)$ is a $J \times J$ matrix defined as

$\Delta_t(p_t) \equiv -\mathscr{H}_t \odot \frac{\partial s_t}{\partial p_t}(p_t)$. Here, $\mathscr{H}_t$ is a firm ownership matrix with entries $\mathscr{H}_{jr} = 1$ if products j and r are owned by the same firm (i.e., $\exists f : \{r, j\} \subset \mathscr{F}_f$), and $\mathscr{H}_{jr} = 0$ otherwise. $\frac{\partial s_t}{\partial p_t}(p_t)$ is the cross-price derivative of demand from the BLP model, and $\odot$ denotes the Hadamard (element-wise) product.

To model marginal costs structurally, I specify them as a linear function of observable product characteristics:

$$mc_{jt} = x_{jt}\gamma + \varepsilon_{jt} \tag{31}$$

where mc_{jt} is the marginal cost of product j in market t. x_{jt} is a vector of product attributes, such as sugar, sodium and caffeine. γ is a vector of cost parameters. ε_{jt} captures unobserved cost shocks.

In the counterfactual scenario, I simulate a sugar reformulation by proportionally reducing the sugar content of each product (e.g., by 10%, 20%, or 30%). Given that sugar is the only input assumed to change under the policy, the change in marginal cost simplifies to:

$$\Delta mc_{jt} = mc_{jt} - mc_{jt,\text{new}} = (Sugar_{jt} - Sugar_{jt,\text{new}}) \cdot \hat{\gamma}_{\text{sugar}} = -\Delta Sugar_{jt} \cdot \hat{\gamma}_{\text{sugar}} \tag{32}$$

where $Sugar_{jt}$ is the original sugar content of product j at time t, and $Sugar_{jt,\text{new}}$ is the sugar content after reformulation. The negative sign in $-\Delta Sugar_{jt}$ reflects the reduction in sugar content, implying a decrease in marginal cost if sugar is a cost-increasing ingredient. $\hat{\gamma}_{\text{sugar}}$ is the estimated marginal cost coefficient associated with sugar content.

The new marginal cost is simply:

$$mc_{jt,new} = mc_{jt} + \Delta mc_{jt} \tag{33}$$

Then, I solve for the new equilibrium prices and market shares under the reformed product attributes. On the supply side, prices follow the Bertrand-Nash equilibrium condition:

$$s_{jt,new}(p_{jt,new}) = (p_{jt,new} - mc_{jt,new})\Delta_t(p_{t,new}) \tag{34}$$

On the demand side, market shares $s_{jt,new}$ are derived from the Equation 25:

$$s_{jt,new} = \int_v \int_D \frac{e^{\delta_{jt,new}+\mu_{ijt,new}}}{1+\sum_{k=1}^J e^{\delta_{kt,new}+\mu_{ikt,new}}} dP_D(D) dP_v(v) \tag{35}$$

where $\delta_{jt,\text{new}} = x_{jt,\text{new}}\beta - \alpha p_{jt,\text{new}} + \xi_{jt}$ and $\mu_{ijt,\text{new}} = \begin{bmatrix} -p_{jt,\text{new}}, & x_{jt,\text{new}} \end{bmatrix} (\Pi D_i + \Sigma v_i)$.

The system defined by Equations 34 and 35 is solved jointly and iteratively. At each step, prices are updated based on simulated marginal costs and markups, and then new shares are computed from the demand system until convergence is achieved for both $p_{jt,new}$ and $s_{jt,new}$.

The predicted changes in market shares resulting from product reformulation can be expressed as:

$$s_{jt}(p_t) - s_{jt,new}(p_{t,new}) = (p_{jt} - mc_{jt})\Delta_t(p_t) - (p_{jt,new} - mc_{jt,new})\Delta_t(p_{t,new}) \tag{36}$$

Using the cost identity from Equation 33, the Equation 36 can be rewritten as:

$$\begin{aligned} s_{jt}(p_t) - s_{jt,new}(p_{t,new}) &= (p_{jt} - mc_{jt})\Delta_t(p_t) - (p_{jt,new} - mc_{jt} + \Delta Sug_{jt})\Delta_t(p_{t,new}) \\ &= (p_{jt} - mc_{jt})\Delta_t(p_t) - (p_{jt,new} - mc_{jt})\Delta_t(p_{t,new}) - \Delta Sug_{jt}\Delta_t(p_{t,new}) \end{aligned} \tag{37}$$

Assuming $\Delta_t(p_{t,new}) = \Delta_t(p_t)$,

$$s_{jt}(p_t) - s_{jt,new}(p_{t,new}) = (p_{jt} - p_{jt,new} - \Delta Sug_{jt})\Delta_t(p_t) \tag{38}$$

This equation is used to compute the change in market share for each product under different sugar reformulation scenarios. Using information on the average sugar content per product, I can calculate how each sugar reduction scenario translates into a decrease in the total quantity of sugar purchased by consumers. This allows me to evaluate not only the impact on market performance but also the potential public health benefits associated with reduced sugar consumption.

4.4 Results

Figure 4.1 displays the distribution of price coefficients. The distribution is centered around a median is -1.246, with a standard deviation of 1.285. This negative median value reflects the expected inverse relationship between price and quantity demanded, confirming the downward-sloping nature of demand. Approximately 4% of simulated consumers exhibit a positive price coefficient, indicating heterogeneity in price sensitivity. These cases may reflect consumers who place relatively less importance on price, possibly due to strong brand loyalty or a consistent preference for certain product characteristics regardless of cost.

Table 4.2 and Tabe 4.3 shows the BLP estimates. I included category, state, year, and month fixed effects to capture category, geographic and seasonal variations in consumers' preferences. Category fixed effects control for inherent differences across beverage types that may be associated with varying levels of sugar, while state fixed effects account for regional differences such as local dietary habits, or socioeconomic factors that might shape consumers' sensitivity to sugar. Additionally, year and month fixed effects help capture time-related variations in consumption patterns, such as seasonal fluctuations in beverage consumption during hot summer months or increased sales during festive periods.

The majority of parameter estimates are statistically significant at the 5% level. The coefficient for sugar is –1.171 and statistically significant, indicating that higher sugar content reduces consumer utility. While sugar is essential for the taste that defines many SSBs, this negative sign suggests that, in the sample, consumers are penalizing higher sugar levels. One plausible explanation is that, following the implementation of detailed nutrition labels in 2016, consumers have become increasingly health-conscious. With clearer information about sugar content, many are now actively avoiding high-sugar products despite the traditional taste appeal of sugar. The coefficient for sodium is –0.002 and statistically significant. Although sodium is not a primary ingredient in most SSBs, its presence in beverages such as sports drinks means that even slight increases can reduce utility. In contrast, caffeine has a positive coefficient (0.019), suggesting that higher caffeine content increases utility. This is consistent with the appeal of caffeinated beverages like colas and energy drinks, where the stimulant effect of caffeine is a valued attribute that enhances the overall consumption experience. Both the Diet (–0.526) and Low Calories (–0.023) indicators exhibit negative coefficients, implying that these reformulated options are less preferred. One possible explanation is that artificial sweeteners or reduced-calorie formulations may not replicate the satisfaction provided by products containing full sugar.

Table 4.4 reports the median price elasticities of demand for all products in market share. The diagonal entries represent each product's own-price elasticity, while the off-diagonal entries capture the cross-price elasticities between pairs of products. All own-price elasticities along the main diagonal are negative, confirming that an increase in a product's own price reduces its demand. The elasticities range roughly from -0.819 to -1.776, implying that most products in this category are relatively price-elastic. The larger (in absolute value) the elasticity, the more sensitive consumers are to price changes for that particular product. Products 6, 8, 13 exhibit the highest own-price elasticities (in absolute value), suggesting that their consumers are especially responsive to price changes. These might be products with many close substitutes in the market or highly competitive brands that rely heavily on promotions and discounts. Conversely, products 10, 11, 14, 16 show relatively smaller elasticities (in absolute value), indicating that their demand is less sensitive to price. These products may be more differentiated, due to unique flavors, brand loyalty, or specific attributes. All cross-price elasticities are positive, indicating that these products are substitutes. However, the magnitude of these cross-price elasticities is relatively small, implying that, while consumers do switch to alternative products, the degree of substitutability may be modest for certain pairs.

Table 4.5 reports the effects of three sugar reduction scenarios (10%, 20%, and 30% reformulations) on both product-level market shares and sugar purchased per serving. The column *Share*Sugar* represents the contribution of each product to the average grams of sugar purchased per 12 oz serving, calculated by multiplying the product's market share by its sugar content. At the baseline, consumers purchase an average of 13.07 grams of sugar per serving from the products included in the analysis. Individual product contributions vary significantly. Product 3 contributes the most at 1.52 grams, while several products such as Product 4 through Product 7 and Product 21 contribute zero due to having no sugar content. When sugar content is reduced by 10 percent, this average falls to 11.30 grams, which is a 13.5 percent decrease. The magnitude of reduction differs across products. Products with both high sugar content and sizable market shares, such as Product 3 and Product 17, experience significant declines in their sugar contributions. In contrast, some products such as Product 13 show very little change in contribution despite slight increases in market share. A 20 percent sugar reduction brings the average down to 10.26 grams, a 21.5 percent decrease. The 30 percent scenario leads to the largest reduction, with average sugar purchased declining to 9.17 grams per serving, a 29.9 percent drop from baseline. Changes in market share across scenarios are generally small, so the reduction in total sugar purchased is mainly due to the lower sugar content. This simulation shows that a policy targeting sugar reformulation could significantly reduce average sugar purchased from SSBs without requiring large shifts in consumer behavior or preferences.

4.5 Conclusions

This study evaluates the efficacy of implementing a sugar content ceiling for SSBs as a strategy to reduce sugar consumption in the United States. Using the BLP random coefficients logit model and a comprehensive dataset from Circana, I simulate a scenario in which all manufacturers voluntarily comply with a predetermined added sugar content standard, not exceeding a specified threshold. This approach allows me to assess the impact of such a policy on prices, market shares, consumer demand, and overall sugar purchased from SSBs.

The BLP model, which accounts for consumer heterogeneity and flexible substitution patterns, provides robust insights into how changes in product characteristics affect demand. My analysis highlights that higher sugar content reduces consumer utility, indicating that consumers penalize higher sugar levels. This trend is likely influenced by the implementation of detailed nutrition labels in 2016, which have made consumers more aware of the sugar content in the products they purchase. A randomized experiment found that after the updated Nutrition Facts label was introduced in 2016, 75–87% of U.S. consumers across five product categories actively used the added-sugar information, and those who accessed it chose lower-sugar options (Kim et al., 2021). The findings also indicate that the sugar ceiling policy can lead to a meaningful reduction in the average sugar content of beverages. However, the effects on consumer demand and market shares are not uniform. Reformulated products generally experience modest declines in market share, with the extent of the decline varying by the degree of reformulation and brand loyalty. These shifts in market share are accompanied by price adjustments, which differ across products as a result of changes in marginal costs and competitive conditions. Despite these adjustments, the overall outcome is a significant reduction in total sugar purchased from SSBs, which supports the policy's nutritional objectives. The analysis also reveals important economic trade-offs. Some consumers substitute toward lower-sugar products, while others maintain their previous consumption patterns due to relatively inelastic preferences. This demonstrates that a sugar ceiling can enhance dietary quality at the population level without substantially reducing consumer choice or placing excessive burdens on producers. The demand-side estimates, which account for differences in consumer preferences, show that individual taste variation plays a role in shaping responses to reformulated products. From a policy perspective, these findings have several implications. First, a sugar ceiling may be a more politically and economically feasible intervention than outright taxes or bans, particularly given its ability to preserve product availability while improving dietary quality. Second, reformulation mandates may be most effective when complemented by consumer education initiatives that encourage acceptance of lower-sugar options. Third, policymakers should remain attentive to possible unintended consequences, such as increased demand for other high-sugar products not subject to the regulation or compensatory behaviors in other food categories. This study contributes to the broader literature on food and nutrition policy by providing empirical evidence on how supply-side interventions can shift market dynamics. Future research could build on this work by exploring substitution patterns across product categories or evaluating the distributional impact of reformulation policies across different demographic groups.

Table 4.1: Summary Statistics for Product-Level Beverage Attributes

Product	Market Share	Standard Error	Sugar (g)	Standard Deviation
Product 1	0.0143	(0.0053)	31.002	(0.011)
Product 2	0.0240	(0.0079)	21.000	(0.000)
Product 3	0.0485	(0.0092)	31.305	(0.575)
Product 4	0.0397	(0.0063)	0.000	(0.000)
Product 5	0.0262	(0.0059)	0.000	(0.000)
Product 6	0.0300	(0.0121)	0.000	(0.000)
Product 7	0.0269	(0.0072)	0.000	(0.000)
Product 8	0.0358	(0.0135)	33.242	(0.117)
Product 9	0.0118	(0.0025)	31.000	(0.000)
Product 10	0.0229	(0.0051)	32.263	(4.386)
Product 11	0.0364	(0.0091)	14.248	(2.318)
Product 12	0.0238	(0.0063)	40.420	(3.013)
Product 13	0.0303	(0.0115)	42.907	(0.297)
Product 14	0.0240	(0.0054)	36.256	(0.334)
Product 15	0.0295	(0.0066)	30.359	(0.171)
Product 16	0.0201	(0.0068)	25.145	(0.524)
Product 17	0.0375	(0.0090)	30.784	(1.452)
Product 18	0.0296	(0.0052)	34.906	(0.108)
Product 19	0.0305	(0.0113)	27.489	(0.878)
Product 20	0.0161	(0.0031)	14.285	(0.920)
Product 21	0.0239	(0.0034)	0.000	(0.000)

Note: Market shares are expressed as proportions. Standard errors of market share estimates are reported in parentheses in the adjacent column. Sugar content is standardized to a 12 oz serving. Values in parentheses under the sugar column indicate the standard deviation across different flavors within each product.

Table 4.2: Parameter Estimates of Demand Side

Mean Utility			**Random Utility**		
Variable	Estimate	Standard Error	Variable	Estimate	Standard Error
Product attributes					
Constant	-1.075***	(0.049)	Price Std Deviation	-0.069***	(0.007)
Sugar	-1.171***	(0.044)	Price × 1/Household Income ($10,000)	-3.687***	(0.013)
Sodium	-0.002***	(0.000)	Price × Married	-0.975***	(0.040)
Caffeine	0.019***	(0.000)	Price × Male	-0.519***	(0.037)
Diet	-0.526***	(0.016)			
Low Calories	-0.023***	(0.015)			
Categories (Ref: Category1)					
Category2	-0.460***	(0.008)			
Category3	0.480***	(0.013)			
Years (Ref: Year1)					
Year2	0.020***	(0.005)			
Months (Ref: Month1)					
Month2	0.000	(0.012)			
Month3	-0.012	(0.012)			
Month4	-0.014	(0.012)			
Month5	-0.029**	(0.012)			
Month6	-0.059***	(0.012)			
Month7	-0.070***	(0.012)			
Month8	-0.055***	(0.012)			
Month9	-0.035***	(0.012)			
Month10	-0.023*	(0.012)			
Month11	-0.011	(0.012)			
Month12	-0.012	(0.012)			

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors are in parentheses. State-level controls are included but omitted here

Table 4.3: Parameter Estimates of Supply Side

Variable	Marginal Cost	Standard Error
Cost factors		
Constant	0.044***	(0.013)
Price of sugar $\times$ Sugar	0.448***	(0.100)
Retailer wage	0.050***	(0.011)
Sodium	0.000	(0.000)
Caffeine	0.000	(0.000)
Firms (Ref: Firm1)		
Firm2	-0.039***	(0.011)
Firm3	-0.024***	(0.003)
Firm4	-0.035***	(0.009)
Firm5	0.021***	(0.006)

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors are in parentheses.

Table 4.4: Price Elasticities of Demand for Products

Product	Product1	Product2	Product3	Product4	Product5	Product6	Product7	Product8	Product9	Product10
Product 1	-1.10881	0.02539	0.05836	0.04526	0.03243	0.03913	0.03183	0.04475	0.01503	0.02255
Product 2	0.01613	-1.4805	0.05864	0.04550	0.03258	0.03929	0.03199	0.04498	0.01510	0.02273
Product 3	0.01590	0.02517	-1.1295	0.04493	0.03223	0.03894	0.03162	0.04449	0.01495	0.02233
Product 4	0.01599	0.02534	0.05826	-1.0970	0.03238	0.03907	0.03178	0.04468	0.01501	0.02249
Product 5	0.01584	0.02507	0.05791	0.04476	-1.1831	0.03882	0.03151	0.04439	0.01492	0.02224
Product 6	0.01573	0.02487	0.05744	0.04444	0.03197	-1.2303	0.03133	0.04413	0.01484	0.02204
Product 7	0.01591	0.02520	0.05800	0.04494	0.03225	0.03893	-1.1515	0.04450	0.01496	0.02235
Product 8	0.01584	0.02505	0.05777	0.04473	0.03214	0.03883	0.03151	-1.1758	0.01491	0.02222
Product 9	0.01577	0.02496	0.05760	0.04460	0.03204	0.03785	0.03140	0.04422	-1.2304	0.02212
Product 10	0.01622	0.02577	0.05896	0.04581	0.03273	0.03940	0.03217	0.04515	0.01517	-0.9864
Product 11	0.01647	0.02617	0.05967	0.04642	0.03309	0.03977	0.03254	0.04567	0.01531	0.02341
Product 12	0.01608	0.02553	0.05845	0.04543	0.03246	0.03910	0.03194	0.04474	0.01505	0.02266
Product 13	0.01568	0.02479	0.05728	0.04430	0.03187	0.03857	0.03124	0.04400	0.01480	0.02195
Product 14	0.01633	0.02594	0.05933	0.04609	0.03292	0.03961	0.03233	0.04544	0.01524	0.02315
Product 15	0.01586	0.02510	0.05784	0.04480	0.03216	0.03886	0.03156	0.04439	0.01492	0.02226
Product 16	0.01644	0.02614	0.05958	0.04637	0.03304	0.03972	0.03251	0.04558	0.01529	0.02335
Product 17	0.01604	0.02544	0.05846	0.04535	0.03247	0.03917	0.03187	0.04481	0.01506	0.02237
Product 18	0.01587	0.02512	0.05789	0.04484	0.03218	0.03889	0.03157	0.04442	0.01493	0.02227
Product 19	0.01586	0.02513	0.05787	0.04483	0.03218	0.03884	0.03155	0.04441	0.01493	0.02228
Product 20	0.01607	0.02554	0.05866	0.04549	0.03256	0.03929	0.03194	0.04495	0.01511	0.02273
Product 21	0.01592	0.02521	0.05804	0.04498	0.03226	0.03895	0.03165	0.04452	0.01497	0.02264

Note: The elasticities presented are the median price elasticities across all markets. The own-price elasticities appear on the diagonal, while the cross-price elasticities are off the diagonal.

Table 4.4: Price Elasticities of Demand for Products (Continued)

Product	Product11	Product12	Product13	Product14	Product15	Product16	Product17	Product18	Product19	Product20	Product21
Product 1	0.02915	0.02497	0.04040	0.02166	0.03602	0.01674	0.04075	0.03625	0.03627	0.01700	0.02831
Product 2	0.02945	0.02512	0.04055	0.02184	0.03618	0.01692	0.04099	0.03643	0.03643	0.01711	0.02845
Product 3	0.02888	0.02475	0.04022	0.02141	0.03580	0.01654	0.04042	0.03604	0.03604	0.01685	0.02813
Product 4	0.02905	0.02492	0.04034	0.02159	0.03596	0.01668	0.04067	0.03620	0.03622	0.01696	0.02827
Product 5	0.02861	0.02465	0.04011	0.02131	0.03569	0.01645	0.04026	0.03593	0.03593	0.01678	0.02804
Product 6	0.02829	0.02445	0.03996	0.02108	0.03550	0.01626	0.03995	0.03571	0.03569	0.01664	0.02785
Product 7	0.02881	0.02477	0.04021	0.02143	0.03582	0.01655	0.04045	0.03604	0.03606	0.01686	0.02814
Product 8	0.02859	0.02463	0.04012	0.02128	0.03569	0.01643	0.04023	0.03591	0.03589	0.01677	0.02802
Product 9	0.02843	0.02453	0.04004	0.02118	0.03557	0.01634	0.04008	0.03581	0.03580	0.01671	0.02793
Product 10	0.02982	0.02536	0.04066	0.02208	0.03637	0.01711	0.04133	0.03661	0.03667	0.01726	0.02862
Product 11	-0.8189	0.02570	0.04099	0.02261	0.03676	0.01763	0.04195	0.03703	0.03702	0.01755	0.02896
Product 12	0.02934	-1.0656	0.04036	0.02178	0.03610	0.01679	0.04099	0.03635	0.03647	0.01708	0.02840
Product 13	0.02814	0.02436	-1.2527	0.02099	0.03541	0.01620	0.03982	0.03562	0.03559	0.01659	0.02777
Product 14	0.03021	0.02548	0.04085	-0.9157	0.03655	0.01735	0.04159	0.03683	0.03682	0.01740	0.02879
Product 15	0.02867	0.02469	0.04015	0.02133	-1.1725	0.01647	0.04031	0.03595	0.03596	0.01680	0.02806
Product 16	0.03058	0.02571	0.04094	0.02254	0.03672	-0.8563	0.04188	0.03698	0.03702	0.01757	0.02829
Product 17	0.02929	0.02502	0.04044	0.02173	0.03607	0.01682	-1.0647	0.03632	0.03634	0.01705	0.02837
Product 18	0.02868	0.02470	0.04017	0.02135	0.03574	0.01648	0.04034	-1.1673	0.03598	0.01682	0.02808
Product 19	0.02869	0.02472	0.04012	0.02137	0.03572	0.01646	0.04037	0.03596	-1.1666	0.01682	0.02808
Product 20	0.02945	0.02509	0.04055	0.02186	0.03616	0.01692	0.04098	0.03644	0.03643	-1.0552	0.02845
Product 21	0.02883	0.02480	0.04023	0.02145	0.03584	0.01656	0.04048	0.03607	0.03608	0.01687	-1.1502

Note: The elasticities presented are the median price elasticities across all markets. The own-price elasticities appear on the diagonal, while the cross-price elasticities are off the diagonal.

Table 4.5: Impact of Sugar Reduction on Market Share and Sugar Purchased by Product

Product	Baseline		Scenario 1 – 10% Sugar Reduction			Scenario 2 – 20% Sugar Reduction			Scenario 3 – 30% Sugar Reduction		
	Share	Share*Sugar	Share	Share*Sugar	% Change	Share	Share*Sugar	% Change	Share	Share*Sugar	% Change
Product 1	0.0143	0.4433	0.0153	0.4269	-3.7%	0.0156	0.3869	-12.7%	0.0159	0.3451	-22.1%
Product 2	0.0240	0.5040	0.0246	0.4649	-7.8%	0.0249	0.4183	-17.0%	0.0251	0.3690	-26.8%
Product 3	0.0485	1.5183	0.0464	1.3073	-13.9%	0.0472	1.1821	-22.1%	0.0481	1.0540	-30.6%
Product 4	0.0397	0.0000	0.0344	0.0000	—	0.0337	0.0000	—	0.0331	0.0000	—
Product 5	0.0262	0.0000	0.0277	0.0000	—	0.0272	0.0000	—	0.0268	0.0000	—
Product 6	0.0300	0.0000	0.0314	0.0000	—	0.0309	0.0000	—	0.0303	0.0000	—
Product 7	0.0269	0.0000	0.0268	0.0000	—	0.0264	0.0000	—	0.0259	0.0000	—
Product 8	0.0358	1.1901	0.0393	1.1758	-1.2%	0.0402	1.0691	-10.2%	0.0411	0.9564	-19.6%
Product 9	0.0118	0.3658	0.0122	0.3404	-6.9%	0.0124	0.3075	-15.9%	0.0126	0.2734	-25.3%
Product 10	0.0229	0.7388	0.0196	0.5691	-23.0%	0.0200	0.5162	-30.1%	0.0205	0.4630	-37.3%
Product 11	0.0364	0.5186	0.0221	0.2834	-45.4%	0.0221	0.2519	-51.4%	0.0222	0.2214	-57.3%
Product 12	0.0238	0.9620	0.0196	0.7130	-25.9%	0.0201	0.6500	-32.4%	0.0207	0.5857	-39.1%
Product 13	0.0303	1.3001	0.0335	1.2936	-0.5%	0.0346	1.1877	-8.6%	0.0358	1.0752	-17.3%
Product 14	0.0240	0.8701	0.0240	0.7831	-10.0%	0.0247	0.7164	-17.7%	0.0254	0.6446	-25.9%
Product 15	0.0295	0.8956	0.0312	0.8525	-4.8%	0.0318	0.7723	-13.8%	0.0324	0.6885	-23.1%
Product 16	0.0201	0.5054	0.0126	0.2851	-43.6%	0.0128	0.2575	-49.1%	0.0129	0.2271	-55.1%
Product 17	0.0375	1.1544	0.0319	0.8838	-23.4%	0.0325	0.8004	-30.7%	0.0332	0.7154	-38.0%
Product 18	0.0296	1.0332	0.0291	0.9142	-11.5%	0.0298	0.8322	-19.5%	0.0304	0.7428	-28.1%
Product 19	0.0305	0.8384	0.0317	0.7843	-6.5%	0.0323	0.7103	-15.3%	0.0328	0.6311	-24.7%
Product 20	0.0161	0.2300	0.0173	0.2224	-3.3%	0.0174	0.1988	-13.6%	0.0174	0.1740	-24.3%
Product 21	0.0239	0.0000	0.0217	0.0000	—	0.0213	0.0000	—	0.0209	0.0000	—
Overall		13.0681		11.2998	-13.5%		10.2573	-21.5%		9.1667	-29.9%

Note: Share*Sugar is calculated as the product of market share (expressed as a proportion) and sugar content (grams per 12 oz), using values from Table 4.1. In each reformulation scenario (10%, 20%, 30%), sugar content is proportionally reduced. The percentage change in Share*Sugar is computed relative to the baseline as: $(\text{Share*Sugar}_{\text{scenario}} - \text{Share*Sugar}_{\text{baseline}}) / \text{Share*Sugar}_{\text{baseline}}$. "—" indicates undefined percentage changes due to zero baseline sugar content. The final row reports the average sugar purchased per serving across all products.

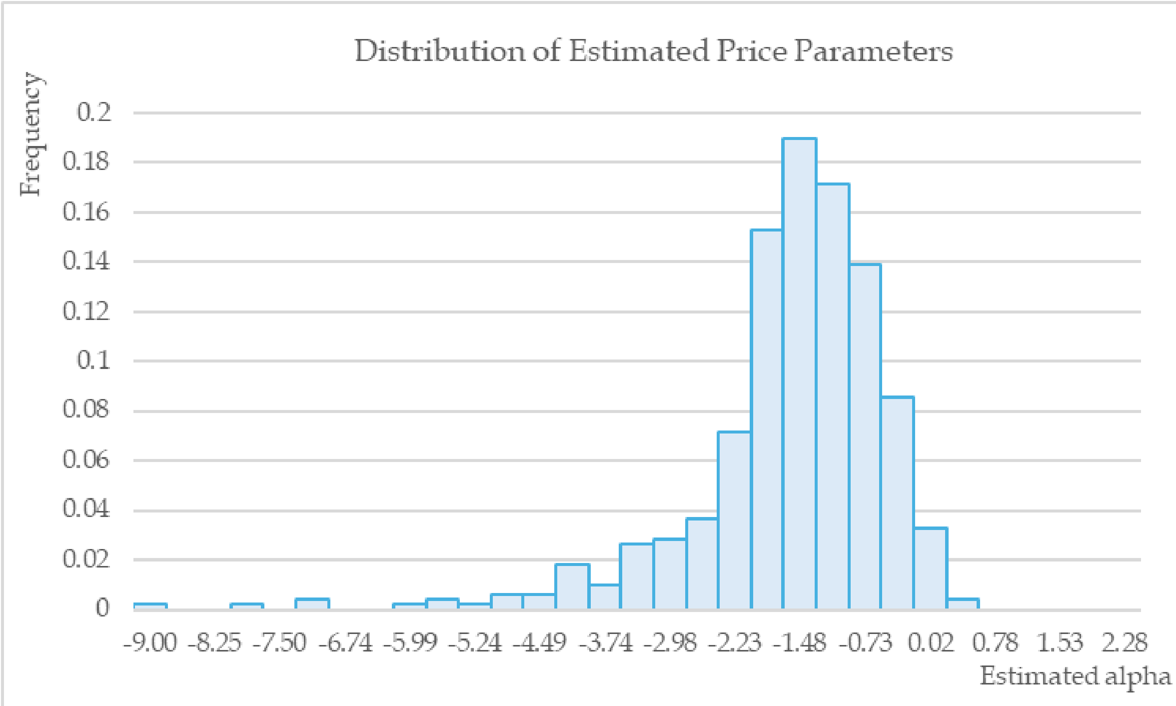


Figure 4.1: Distribution of Price Parameter

Bibliography

Airhihenbuwa, C. O., Kumanyika, S., Agurs, T. D., Lowe, A., Saunders, D., and Morssink, C. B. (1996). Cultural aspects of african american eating patterns. *Ethnicity health*, 1(3):245–260.

Alkerwi, A., Baydarlioglu, B., Sauvageot, N., Stranges, S., Lemmens, P., Shivappa, N., and Hébert, J. R. (2017). Smoking status is inversely associated with overall diet quality: Findings from the oriscav-lux study. *Clinical Nutrition*, 36(5):1275–1282.

Allcott, H., Diamond, R., Dubé, J.-P., Handbury, J., Rahkovsky, I., and Schnell, M. (2019). Food deserts and the causes of nutritional inequality. *The Quarterly Journal of Economics*, 134(4):1793–1844.

Anderson, P. M. and Butcher, K. F. (2016). The relationships among snap benefits, grocery spending, diet quality, and the adequacy of low-income families' resources. *Center for Budget and Policy Priorities Policy Futures*, 20.

Armour, B. S., Pitts, M. M., and Lee, C.-W. (2007). Cigarette smoking and food insecurity among low-income families in the United States, 2001. *American Journal of Health Promotion*, 22(6):386–390.

Baylis, K., Fan, L., and Nogueira, L. (2019). Agricultural market liberalization and household food security in rural china. *American journal of agricultural economics*, 101(1):250–269.

Beatty, T. K. M., Lin, B.-H., and Smith, T. A. (2014). Is diet quality improving? distributional changes in the United States, 1989–2008. *American Journal of Agricultural Economics*, 96(3):769–789.

Berry, S., Levinsohn, J., and Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica*, 63(4):841.

Berube, L. T., Kiely, M., Yazici, Y., and Woolf, K. (2017). Diet quality of individuals with rheumatoid arthritis using the Healthy Eating Index (HEI)-2010. *Nutrition and Health*, 23(1):17–24.

Bouis, H. E., Eozenou, P., and Rahman, A. (2011). Food prices, household income, and resource allocation: socioeconomic perspectives on their effects on dietary quality and nutritional status. *Food and Nutrition Bulletin*, 32(1):S14–S23.

Cawley, J. and Frisvold, D. E. (2017). The pass-through of taxes on sugar-sweetened beverages to retail prices: The case of berkeley, california. *Journal of policy analysis and management*, 36(2):303–326.

Cengiz, E. and Rojas, C. (2024). Are food manufacturers reducing sugar content? Evidence from scanner data. *Agribusiness*, 40(3):571–595.

Centers for Disease Control and Prevention (2021). Adult obesity facts. Retrieved from https://data.bls.gov/cew/apps/data_views/data_views.htm#tab=Tables.

Chen, Y. and Zhen, C. (2022). The potential impact of reducing sodium in packaged food: The case of the Chinese instant noodles market. *Agribusiness*, 38(1):3–20.

Chernozhukov, V. and Hansen, C. (2005). An iv model of quantile treatment effects. *Econometrica*, 73(1):245–261.

Chidmi, B. and Lopez, R. A. (2007). Brand-supermarket demand for breakfast cereals and retail competition. *American Journal of Agricultural Economics*, 89(2):324–337.

Conlon, C. and Gortmaker, J. (2020). Best practices for differentiated products demand estimation with pyblp. *The Rand journal of economics*, 51(4):1108–1161.

Crews, D. C., Kuczmarski, M. F., Miller, E. R., Zonderman, A. B., Evans, M. K., and Powe, N. R. (2015). Dietary habits, poverty, and chronic kidney disease in an urban population. *Journal of Renal Nutrition*, 25(2):103–110.

Davis, G. C. and You, W. (2010). The time cost of food at home: general and food stamp participant profiles. *Applied economics*, 42(20):2537–2552.

Dean, O., Gothro, A., Bleiweiss-Sande, R., Navarro, S., and Reynolds, M. (2022). Older adult supplemental nutrition assistance program (snap) participation series. Washington, DC: AARP Public Policy Institute, July 13, 2022. https://doi.org/10.26419/ppi.00166.000.

Debnam, J. (2017). Selection effects and heterogeneous demand responses to the berkeley soda tax vote. *American Journal of Agricultural Economics*, 99(5):1172–1187.

Devine, C. M., Connors, M. M., Sobal, J., and Bisogni, C. A. (2003). Sandwiching it in: spillover of work onto food choices and family roles in low- and moderate-income urban households. *Social science medicine (1982)*, 56(3):617–630.

Devine, C. M., Jastran, M., Jabs, J., Wethington, E., Farell, T. J., and Bisogni, C. A. (2006). "A lot of sacrifices:" work–family spillover and the food choice coping strategies of low-wage employed parents. *Social science medicine (1982)*, 63(10):2591–2603.

Economic Research Service (ERS), U.S. Department of Agriculture (USDA). (2019). SNAP policy database, SNAP Policy Data Sets. Retrieved from https://www.ers.usda.gov/data-products/snap-policy-data-sets/.

Falbe, J., Thompson, H. R., Becker, C. M., Rojas, N., McCulloch, C. E., and Madsen, K. A. (2016). Impact of the berkeley excise tax on sugar-sweetened beverage consumption. *American journal of public health*, 106(10):1865–71.

Feng, J., Fan, L., and Jaenicke, E. C. (2023). The distributional impact of SNAP on dietary quality. *Agricultural Economics*.

Finifter, D. H. and Prell, M. A. (2013). *Participation in the supplemental nutrition assistance program (SNAP) and unemployment insurance : how tight are the strands of the recessionary safety net?* . United States Department of Agriculture, Economic Research Service.

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

Firpo, S. P., Fortin, N. M., and Lemieux, T. (2018). Decomposing wage distributions using recentered influence function regressions. *Econometrics*, 6(2):28.

Fletcher, J. M., Frisvold, D., and Tefft, N. (2010). Can soft drink taxes reduce population weight? *Contemporary Economic Policy*, 28(1):23–35.

French, S. A., Tangney, C. C., Crane, M. M., Wang, Y., and Appelhans, B. M. (2019). Nutrition quality of food purchases varies by household income: the shopper study. *BMC public health*, 19(1):231–231.

French, S. A., Wall, M., Mitchell, N. R., Shimotsu, S. T., and Welsh, E. (2009). Annotated receipts capture household food purchases from a broad range of sources. *The international journal of behavioral nutrition and physical activity*, 6(1):37–37.

Fulgoni, V. L., Keast, D. R., and Drewnowski, A. (2009). Development and validation of the nutrient-rich foods index: a tool to measure nutritional quality of foods 1–3. *The Journal of Nutrition*, 139(8):1549–1554.

Gao, X., Qi, L., Qiao, N., Choi, H. K., Curhan, G., Tucker, K. L., and Ascherio, A. (2007). Intake of added sugar and sugar-sweetened drink and serum uric acid concentration in US men and women. *Hypertension*, 50(2):306–312.

Gleason, S., Wolford, B., Wilkin, M., Hofer, B., Woloshin, D., and Sallack, L. (2018). Analysis of SNAP-Ed data for all states. USDA Food and Nutrition Service. Retrieved from https://fns-prod.azureedge.us/sites/default/files/ops/SNAPED-Data-AllStates.pdf.

Gregory, C., Ver Ploeg, M., Andrews, M., and Coleman-Jensen, A. (2013). Supplemental nutrition assistance program (snap) participation leads to modest changes in diet quality. No. 1477-2017-3991.

Gregory, C. A., Mancino, L., and Coleman-Jensen, A. (2019). Food security and food purchase quality among low-income households: Findings from the national household food acquisition and purchase survey (foodaps). *U.S. Department of Agriculture Economic Research Service*. ERR-269.

Grossman, M. (1972a). *The demand for health : a theoretical and empirical investigation*. National Bureau of Economic Research. Occasional paper ; 119. National Bureau of Economic Research; distributed by Columbia University Press, New York.

Grossman, M. (1972b). On the concept of health capital and the demand for health. *The Journal of political economy*, 80(2):223–255.

Grummon, A. H. and Taillie, L. S. (2018). Supplemental Nutrition Assistance Program participation and racial/ethnic disparities in food and beverage purchases. *Public Health Nutrition*, 21(18):3377–3385.

Guenther, P. M., Casavale, K. O., Reedy, J., Kirkpatrick, S. I., Hiza, H. A., Kuczynski, K. J., Kahle, L. L., and Krebs-Smith, S. M. (2013). Update of the healthy eating index: HEI-2010. *Journal of the Academy of Nutrition and Dietetics*, 113(4):569–580.

Guenther, P. M., Kirkpatrick, S. I., Reedy, J., Krebs-Smith, S. M., Buckman, D. W., Dodd, K. W., Casavale, K. O., and Carroll, R. J. (2014). The Healthy Eating Index-2010 Is a valid and reliable measure of diet quality according to the 2010 dietary guidelines for Americans 1–3. *The Journal of Nutrition*, 144(3):399–407.

Haeck, C., Lawson, N., and Poirier, K. (2022). Estimating consumer preferences for different beverages using the BLP approach. *Economics & Human Biology*, 46:101138.

Hake, M. and Dawes, S. (2024). Food insecurity among seniors and older adults in 2022. Feeding America. Retrieved from https://www.feedingamerica.org/sites/default/files/2024-07/Food

Hanson, K. L. and Connor, L. M. (2014). Food insecurity and dietary quality in US adults and children: a systematic review. *The American Journal of Clinical Nutrition*, 100(2):684–692.

Harmon, B. E., Boushey, C. J., Shvetsov, Y. B., Ettienne, R., Reedy, J., Wilkens, L. R., Le Marchand, L., Henderson, B. E., and Kolonel, L. N. (2015). Associations of key diet-quality indexes with mortality in the multiethnic cohort: the dietary patterns methods project. *The American journal of clinical nutrition*, 101(3):587–597.

Hastings, J., Kessler, R., and 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.

Hastings, J. and Shapiro, J. M. (2018). How are snap benefits spent? evidence from a retail panel. *The American economic review*, 108(12):3493–3540.

Hirsch, S., Tiboldo, G., and Lopez, R. A. (2018). A tale of two italian cities: brand-level milk demand and price competition. *Applied economics*, 50(49):5239–5252.

Hiza, H. A., Casavale, K. O., Guenther, P. M., and Davis, C. A. (2013). Diet quality of Americans differs by age, sex, race/ethnicity, income, and education level. *Journal of the Academy of Nutrition and Dietetics*, 113(2):297–306.

Hudak, K. M., Racine, E. F., and Schulkind, L. (2021). An increase in snap benefits did not impact food security or diet quality in youth. *Journal of the Academy of Nutrition and Dietetics*, 121(3):507–519.e12.

Hwalla, N., El Labban, S., and Bahn, R. A. (2016). Nutrition security is an integral component of food security. *Frontiers in life science*, 9(3):167–172.

Imbens, G. W. and Newey, W. K. (2009). Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica*, 77(5):1481–1512.

Ingram, J. (2020). Nutrition security is more than food security. *Nature food*, 1(1):2–2.

Kim, E. J., Ellison, B., McFadden, B., and Prescott, M. P. (2021). Consumers' decisions to access or avoid added sugars information on the updated nutrition facts label. *PloS one*, 16(3):e0249355–.

Koning, L. d., Malik, V. S., Kellogg, M. D., Rimm, E. B., Willett, W. C., and Hu, F. B. (2012). Sweetened beverage consumption, incident coronary heart disease, and biomarkers of risk in men. *Circulation*, 125(14):1735–1741.

Krueger, A. B. and Mueller, A. (2010). Job search and unemployment insurance: New evidence from time use data. *Journal of public economics*, 94(3):298–307.

Kuka, E. (2020). Quantifying the benefits of social insurance: Unemployment insurance and health. *The review of economics and statistics*, 102(3):490–505.

Leung, C. W., Ding, E. L., Catalano, P. J., Villamor, E., Rimm, E. B., and Willett, W. C. (2012). Dietary intake and dietary quality of low-income adults in the supplemental nutrition assistance program. *The American journal of clinical nutrition*, 96(5):977–988.

Leung, C. W., Epel, E. S., Ritchie, L. D., Crawford, P. B., and Laraia, B. A. (2014). Food Insecurity Is Inversely Associated with Diet Quality of Lower-Income Adults. *Journal of the Academy of Nutrition and Dietetics*, 114(12):1943–1953.e2.

Leung, C. W., Musicus, A. A., Willett, W. C., and Rimm, E. B. (2017). Improving the nutritional impact of the supplemental nutrition assistance program:: perspectives from the participants. *American Journal of Preventive Medicine*, 52(2):S193–S198.

Leung, C. W. and Tester, J. M. (2019). The association between food insecurity and diet quality varies by race/ethnicity: an analysis of national health and nutrition examination survey 2011-2014 results. *Journal of the Academy of Nutrition and Dietetics*, 119(10):1676–1686.

Liu, J., Rehm, C. D., Onopa, J., and Mozaffarian, D. (2020). Trends in diet quality among youth in the united states, 1999-2016. *JAMA : the journal of the American Medical Association*, 323(12):1161–1174.

Lopez, E. and Lopez, R. A. (2009). Demand for differentiated milk products: implications for price competition. *Agribusiness*, 25(4):453–465.

Lopez, R. A. and Fantuzzi, K. L. (2012). Demand for carbonated soft drinks: implications for obesity policy. *Applied Economics*, 44(22):2859–2865.

Lopez, R. A., Liu, Y., and Zhu, C. (2015). TV advertising spillovers and demand for private labels: the case of carbonated soft drinks. *Applied Economics*, 47(25):2563–2576.

Lundeen, E. A., Park, S., Pan, L., and Blanck, H. M. (2018). Daily intake of sugar-sweetened beverages among us adults in 9 states, by state and sociodemographic and behavioral characteristics, 2016. *Preventing chronic disease*, 15:E154–E154.

Mabli, J., Castner, L., Ohls, J., Fox, M., Crepinsek, M., and Condon, E. (2010). Food expenditures and diet quality among low-income households and individuals. Mathematica Policy Research.

Malik, V. S. and Hu, F. B. (2012). Sweeteners and risk of obesity and type 2 diabetes: The role of sugar-sweetened beverages. *Current diabetes reports*, 12(2):195–203.

Malik, V. S., Popkin, B. M., Bray, G. A., Després, J. P., and Hu, F. B. (2010). Sugar-sweetened beverages, obesity, type 2 diabetes mellitus, and cardiovascular disease risk. *Circulation*, 121(11):1356–1364.

Mancino, L. and Gregory, C. A. (2020). Food-insecure households score lower on diet quality compared to food-secure households. *Amber waves*, March 2020(2):1–6.

Massey, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78.

Mozaffarian, D., Fleischhacker, S., and Andrés, J. R. (2021). Prioritizing nutrition security in the us. *Jama*, 325(16):1605–1606.

Nevo, A. (2001). Measuring market power in the ready-to-eat cereal industry. *Econometrica*, 69(2):307–342.

Nguyen, B. T. and Powell, L. M. (2014). The impact of restaurant consumption among us adults: effects on energy and nutrient intakes. *Public health nutrition*, 17(11):2445–2452.

Nguyen, B. T., Shuval, K., Bertmann, F., and Yaroch, A. L. (2015). The supplemental nutrition assistance program, food insecurity, dietary quality, and obesity among us adults. *American journal of public health*, 105(7):1453–1459.

Nguyen, B. T., Shuval, K., Njike, V. Y., and Katz, D. L. (2014). The supplemental nutrition assistance program and dietary quality among US adults: findings from a nationally representative survey. *Mayo Clinic Proceedings*, 89(9):1211–1219.

Nicklas, T. A., Jahns, L., Bogle, M. L., Chester, D. N., Giovanni, M., Klurfeld, D. M., Laugero, K., Liu, Y., Lopez, S., and Tucker, K. L. (2013). Barriers and facilitators for consumer adherence to the dietary guidelines for americans: the health study. *Journal of the Academy of Nutrition and Dietetics*, 113(10):1317–1331.

Niles, M. T., Bertmann, F., Belarmino, E. H., Wentworth, T., Biehl, E., and Neff, R. (2020). The early food insecurity impacts of covid-19. *Nutrients*, 12(7):2096.

Onvani, S., Haghighatdoost, F., Surkan, P. J., Larijani, B., and Azadbakht, L. (2017). Adherence to the healthy eating index and alternative healthy eating index dietary patterns and mortality from all causes, cardiovascular disease and cancer: a meta-analysis of observational studies. *Journal of human nutrition and dietetics*, 30(2):216–226.

Pinstrup-Andersen, P. (2009). Food security: definition and measurement. *Food security*, 1(1):5–7.

Ratcliffe, C., McKernan, S.-m., and Zhang, S. (2011). How much does the supplemental nutrition assistance program reduce food insecurity? *American journal of agricultural economics*, 93(4):1082–1098.

Rehm, C. D., Peñalvo, J. L., Afshin, A., and Mozaffarian, D. (2016). Dietary intake among US adults, 1999-2012. *JAMA*, 315(23):2542–2553.

Reich, M. and West, R. (2015). The effects of minimum wages on food stamp enrollment and expenditures. *Industrial relations (Berkeley)*, 54(4):668–694.

Rennhoff, A. (2008). Paying for shelf space: an investigation of merchandising allowances in the grocery industry. *Journal of Agricultural & Food Industrial Organization*, 6(1).

Reynaert, M. and Verboven, F. (2014). Improving the performance of random coefficients demand models: The role of optimal instruments. *Journal of Econometrics*, 179(1):83–98.

Rios-Avila, F. (2020). Recentered influence functions (RIFs) in Stata: RIF regression and RIF decomposition. *The Stata Journal*, 20(1):51–94.

Rojas, C. and Jaenicke, E. (2024). Introduction to Agribusiness special issue: Economics of food reformulation. *Agribusiness*, 40(3):537–549.

Rothe, C. (2010a). Identification of unconditional partial effects in nonseparable models. *Economics letters*, 109(3):171–174.

Rothe, C. (2010b). Nonparametric estimation of distributional policy effects. *Journal of econometrics*, 155(1):56–70.

Schmeiser, M. D. (2012). The impact of long-term participation in the supplemental nutrition assistance program on child obesity. *Health economics*, 21(4):386–404.

Seligman, H. K., Levi, R., Adebiyi, V. O., Coleman-Jensen, A., Guthrie, J. F., and Frongillo, E. A. (2023). Assessing and monitoring nutrition security to promote healthy dietary intake and outcomes in the united states. *Annual review of nutrition*, 43(1):409–429.

Semmes, C. (1996). *Racism, health, and post-industrialism: A theory of African-American health*. Praeger, Westport, CT.

Shreay, S., Chouinard, H. H., and McCluskey, J. J. (2015). Product differentiation by package size: product differentiation by package size. *Agribusiness*, 32(1):3–15.

Silva, A., Caro, J. C., and Magaña-Lemus, D. (2016). Household food security: Perceptions, behavior and nutritional quality of food purchases. *Journal of Economic Psychology*, 55:139–148.

Singleton, C. R., Young, S. K., Kessee, N., Springfield, S. E., and Sen, B. P. (2020). Examining disparities in diet quality between SNAP participants and non-participants using Oaxaca-Blinder decomposition analysis. *Preventive Medicine Reports*, 19:101134.

Smed, S., Jensen, J. D., and Denver, S. (2007). Socio-economic characteristics and the effect of taxation as a health policy instrument. *Food policy*, 32(5-6):624–639.

Smith, T. A., Berning, J. P., Yang, X., Colson, G., and Dorfman, J. H. (2016). The effects of benefit timing and income fungibility on food purchasing decisions among supplemental nutrition assistance program households. *American journal of agricultural economics*, 98(2):564–580.

Smith, T. A. and Valizadeh, P. (2024). Aging out of wic and child nutrition: Evidence from a regression discontinuity design. *American journal of agricultural economics*, 106(2):904–924.

Smith, T. A., Valizadeh, P., Lin, B.-H., and Coats, E. (2019). What is driving increases in dietary quality in the United States? *Food Policy*, 86:101720.

Taylor, C. A., Spees, C. K., Markwordt, A. M., Watowicz, R. P., Clark, J. K., and Hooker, N. H. (2017). Differences in US adult dietary patterns by food security status. *Journal of Consumer Affairs*, 51(3):549–565.

Ulpiano J. Vázquez-Martínez, P., Morales-Mediano, J., and Antonio L. Leal-Rodríguez, P. (2021). The impact of the covid-19 crisis on consumer purchasing motivation and behavior the impact of the covid-19 crisis on consumer purchasing motivation and behavior. *European research on management and business economics*, 27(3):100166–.

U.S. Department of Agriculture, Economic Research Service. (2024a). Sugar and sweeteners yearbook tables. Retrieved from https://www.ers.usda.gov/data-products/sugar-and-sweeteners-yearbook-tables/.

U.S. Department of Agriculture, Economic Research Service. (2024b). Household food security in the United States in 2023. Retrieved from https://www.ers.usda.gov/publications/pub-details?pubid=109895#:~:text=Related%20Content-,Overview,life%20for%20all%20household%20members.

U.S. Department of Agriculture, Economic Research Service. (2025). Food security in the U.S. Retrieved from https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us.

U.S. Department of Agriculture, Food and Nutrition Service. (2022). USDA actions on nutrition security. Retrieved from https://www.usda.gov/sites/default/files/documents/usda-actions-nutrition-security.pdf.

U.S. Department of Labor. (2012). Comparison of state unemployment insurance laws. Retrieved from https://oui.doleta.gov/unemploy/pdf/uilawcompar/2012/monetary.pdf.

U.S. Department of Labor. (2013). Comparison of state unemployment insurance laws. Retrieved from https://oui.doleta.gov/unemploy/pdf/uilawcompar/2013/monetary.pdf.

U.S. Department of Labor, Bureau of labor statistics: Quarterly census of employment and wages (2020). Retrieved from https://www.cdc.gov/obesity/data/adult.html.

U.S. Food and Drug Administration (2016). Changes to the nutrition facts label. Retrieved from https://www.fda.gov/food/nutrition-food-labeling-and-education/changestonutritionfactslabel.

Valizadeh, P. and Ng, S. W. (2021). Would A national sugar-sweetened beverage tax in the United States be well targeted? *American Journal of Agricultural Economics*, 103(3):961–986.

Valizadeh, P., Popkin, B. M., and Ng, S. W. (2020). Distributional changes in U.S. sugar-sweetened beverage purchases, 2002–2014. *American Journal of Preventive Medicine*, 59(2):260–269.

Wang, D. D., Leung, C. W., Li, Y., Ding, E. L., Chiuve, S. E., Hu, F. B., and Willett, W. C. (2014). Trends in dietary quality among adults in the United States, 1999 through 2010. *JAMA Internal Medicine*, 174(10):1587–1595.

Welsh, J. A., Sharma, A. J., Grellinger, L., and Vos, M. B. (2011). Consumption of added sugars is decreasing in the United States. *American Journal of Clinical Nutrition*, 94(3):726–734.

Yen, S. T. (2010). The effects of snap and wic programs on nutrient intakes of children. *Food policy*, 35(6):576–583.

Zhang, F. F., Liu, J., Rehm, C. D., Wilde, P., Mande, J. R., and Mozaffarian, D. (2018). Trends and disparities in diet quality among us adults by supplemental nutrition assistance program participation status. *JAMA network open*, 1(2):e180237–e180237.

Zhang, Q. and Gallardo, R. K. (2022). Measuring consumers' demand for nutrition attributes: an application to ready-to-heat meals. *Agricultural and Resource Economics Review*, pages 1–16.

Zhang, Y. and Palma, M. A. (2021). Revisiting sugar taxes and sugary drink consumption: evidence from the random-coefficient demand model. *Journal of Agricultural and Resource Economics*, 46(1):37–55.

Zhen, C., Brissette, I. F., and Ruff, R. R. (2014). By ounce or by calorie: The differential effects of alternative sugar-sweetened beverage tax strategies. *American Journal of Agricultural Economics*, 96(4):1070–1083.

Appendix

Appendix to Chapter 2

This appendix includes Table A1, which presents the components of the Healthy Eating Index (HEI)-2010 scores and the standards used for scoring each component. Table A2 shows the variance inflation factors (VIFs) for the independent variables included in the analysis, providing a diagnostic check for potential multicollinearity among covariates in the regression models. Table A3 and Table A4 report the OLS and Unconditional Quantile Regression estimates of HEI-2010 for food-insecure and food-secure households, respectively.

Table A1: Components of the Healthy Eating Index (HEI)-2010 Scores and Standards for Scoring Each Component

Component	Maximum score	Standard for maximum score	Standard for minimum score of zero
Adequacy components			
Total Fruit[1]	5	$\geq$ 0.8 cup equiv. per 1,000 kcal	No Fruit
Whole Fruit[2]	5	$\geq$ 0.4 cup equiv. per 1,000 kcal	No Whole Fruit
Total Vegetable[3]	5	$\geq$ 1.1 cup equiv. per 1,000 kcal	No Vegetables
Greens and Beans[3]	5	$\geq$ 0.2 cup equiv. per 1,000 kcal	No Dark Green Vegetables or Beans and Peas
Whole Grains	10	$\geq$ 1.5 oz equiv. per 1,000 kcal	No Whole Grains
Dairy[4]	10	$\geq$ 1.3 cup equiv. per 1,000 kcal	No Dairy
Total Protein Foods[5]	5	$\geq$ 2.5 oz equiv. per 1,000 kcal	No Protein Foods
Seafood and Plant Proteins[5,6]	5	$\geq$ 0.8 oz equiv. per 1,000 kcal	No Seafood or Plant Proteins
Fatty Acids[7]	10	(PUFAs + MUFAs)/SFAs $\geq$ 2.5	(PUFAs + MUFAs)/SFAs $\leq$ 1.2
Moderation components			
Refined Grains	10	$\leq$ 1.8 oz equiv. per 1,000 kcal	$\geq$ 4.3 oz equiv. per 1,000 kcal
Sodium	10	$\leq$ 1.1 gram equiv. per 1,000 kcal	$\geq$ 2.0 gram equiv. per 1,000 kcal
Empty Calories[8]	20	$\leq$ 19% of energy	$\geq$ 50% of energy

Source: Guenther et al. (2013)

Note: 1: Includes 100% fruit juice. 2: Includes all forms except juice. 3: Includes any beans and peas not counted as Total Protein Foods. 4: Includes all milk products, such as fluid milk, yogurt, and cheese, and fortified soy beverages. 5: Beans and peas are included here (and not with vegetables) when the Total Protein Foods standard is otherwise not met. 6: Includes seafood, nuts, seeds, soy products (other than beverages) as well as beans and peas counted as Total Protein Foods. 7: Ratio of poly- and monounsaturated fatty acids (PUFAs and MUFAs) to saturated fatty acids (SFAs). 8: Calories from solid fats, alcohol, and added sugars; threshold for counting alcohol is $\geq$ 13 grams/1000 kcal.

Table A2: Variance Inflation Factor (VIF) for Independent Variables

Variable	**VIF**
Male	1.394
Age	4.822
White Hispanic	1.235
Some college	1.720
College diploma	1.423
Post graduate degree	1.295
Married	2.486
Divorced	1.519
Smoking	1.383
Number of children	1.787
Employment	3.408
Monthly household income ($1,000’s)	2.794
FAH expenditures per person	2.106
FAFH expenditures per person	1.595
Midwest	1.623
Northeast	1.476
West	1.592

Note: VIF = 1: No correlation between the independent variable and the other variables. VIF between 1 and 5: Generally, it indicates a moderate level of multicollinearity. VIF greater than 5: This might be a cause for concern, indicating a high level of multicollinearity. From the table, all VIFs are less than 5, meaning the independent variables are not highly correlated.

Table A3: OLS and Unconditional Quantile Regression Estimates of Healthy Eating Index-2010 Among Food-insecure Households

	OLS	Quantiles								
Variable		Q15	Q25	Q35	Q45	Q50	Q55	Q65	Q75	Q85
Constant	43.890***	30.954***	34.828***	38.563***	42.583***	43.908***	47.346***	49.738***	52.403***	56.006***
	(1.254)	(1.766)	(1.583)	(1.578)	(1.640)	(1.664)	(1.684)	(1.723)	(1.836)	(2.181)
Male	-0.993	-1.059	-1.128	-0.874	-1.110	-1.374*	-1.098	-1.204	-1.496*	-1.266
	(0.609)	(0.857)	(0.768)	(0.766)	(0.796)	(0.808)	(0.817)	(0.837)	(0.891)	(1.059)
Age	0.040**	0.024	0.040	0.037	0.029	0.028	0.009	0.028	0.050*	0.055*
	(0.019)	(0.027)	(0.024)	(0.024)	(0.025)	(0.026)	(0.026)	(0.026)	(0.028)	(0.034)
White Hispanic	1.008	0.604	0.875	1.551*	1.143	0.738	1.046	1.483	1.436	1.049
	(0.737)	(1.038)	(0.931)	(0.928)	(0.964)	(0.978)	(0.990)	(1.013)	(1.079)	(1.282)
Some college	0.596	0.019	-0.020	0.562	1.183	0.757	0.306	0.724	0.551	1.399
	(0.567)	(0.799)	(0.716)	(0.714)	(0.742)	(0.753)	(0.762)	(0.780)	(0.831)	(0.987)
College diploma	4.141***	2.756**	3.222***	2.823**	2.493**	2.126*	2.196*	3.225**	3.402**	6.717***
	(0.952)	(1.340)	(1.201)	(1.197)	(1.244)	(1.262)	(1.277)	(1.308)	(1.393)	(1.655)
Post graduate degree	2.908*	3.313	5.292***	4.586**	2.162	1.978	1.948	1.017	1.466	4.005
	(1.572)	(2.214)	(1.985)	(1.979)	(2.056)	(2.086)	(2.111)	(2.161)	(2.302)	(2.734)
Married	-0.366	-0.173	-0.258	-0.510	-0.235	0.269	0.000	0.491	0.025	-0.783
	(0.611)	(0.860)	(0.771)	(0.769)	(0.799)	(0.810)	(0.820)	(0.839)	(0.894)	(1.062)
Divorced	-0.097	0.236	-0.039	0.225	-0.292	-0.409	-0.421	-0.777	-0.578	-0.946
	(0.687)	(0.967)	(0.867)	(0.864)	(0.898)	(0.911)	(0.922)	(0.943)	(1.005)	(1.194)
Smoking	-4.456***	-3.626***	-4.230***	-4.459***	-4.731***	-4.477***	-4.505***	-5.071***	-5.502***	-5.491***
	(0.562)	(0.791)	(0.710)	(0.707)	(0.735)	(0.746)	(0.754)	(0.772)	(0.823)	(0.977)
Number of children	-0.025	0.487*	0.599**	0.108	0.075	0.001	-0.203	-0.441	-0.566*	-0.692*
	(0.203)	(0.286)	(0.257)	(0.256)	(0.266)	(0.270)	(0.273)	(0.279)	(0.298)	(0.353)
Employment	0.826	0.821	0.697	0.932	0.844	0.797	0.092	0.179	0.188	0.343
	(0.620)	(0.873)	(0.782)	(0.780)	(0.810)	(0.822)	(0.832)	(0.852)	(0.907)	(1.078)
Monthly household income ($1000's)	0.160	0.398**	0.208	0.258*	0.124	0.101	0.136	0.014	0.217	0.352*
	(0.117)	(0.165)	(0.148)	(0.148)	(0.153)	(0.156)	(0.157)	(0.161)	(0.172)	(0.204)
FAH expenditures per person	0.046***	0.053***	0.050***	0.050***	0.046***	0.044***	0.040***	0.040***	0.043***	0.051***
	(0.007)	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.010)	(0.012)
FAFH expenditures per person	0.004	0.013	0.018	-0.003	-0.002	-0.003	-0.002	0.013	-0.002	0.011
	(0.010)	(0.015)	(0.013)	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.015)	(0.018)
F	13.90	5.737	8.411	8.061	7.347	7.025	7.259	9.160	9.929	9.348
Prob>F	0	0	0	0	0	0	0	0	0	0
R-squared	0.0961	0.0420	0.0605	0.0581	0.0532	0.0510	0.0526	0.0655	0.0706	0.0667
Adj R-squared	0.0892	0.0347	0.0533	0.0509	0.0460	0.0437	0.0454	0.0583	0.0635	0.0596
Root MSE	11.91	16.77	15.03	14.98	15.57	15.79	15.98	16.36	17.43	20.70
Number of obs	2,240	2,240	2,240	2,240	2,240	2,240	2,240	2,240	2,240	2,240

Note: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors are in parentheses.

Table A4: OLS and Unconditional Quantile Regression Estimates of Healthy Eating Index-2010 Among Food-secure Households

	OLS	Quantiles								
Variable		Q15	Q25	Q35	Q45	Q50	Q55	Q65	Q75	Q85
Constant	45.079***	30.075***	34.467***	35.812***	39.162***	42.827***	45.393***	51.554***	54.733***	60.418***
	(1.371)	(1.886)	(1.843)	(1.925)	(1.922)	(1.880)	(1.885)	(1.999)	(2.138)	(2.273)
Male	-0.713	-1.015	-1.313*	-0.577	-0.463	-0.698	-0.138	-0.898	-0.263	-0.981
	(0.590)	(0.812)	(0.793)	(0.828)	(0.827)	(0.809)	(0.811)	(0.860)	(0.920)	(0.978)
Age	0.026	0.052**	0.060**	0.072***	0.055**	0.039*	0.022	-0.009	-0.020	-0.001
	(0.017)	(0.024)	(0.023)	(0.024)	(0.024)	(0.024)	(0.024)	(0.025)	(0.027)	(0.029)
White hispanic	2.631***	2.880**	1.954	3.813***	3.896***	3.497**	3.170**	2.648*	2.852*	0.237
	(1.020)	(1.403)	(1.370)	(1.431)	(1.429)	(1.398)	(1.402)	(1.486)	(1.590)	(1.690)
Some college	1.820***	1.059	1.401	1.413	2.551***	1.555*	2.056**	1.637*	2.907***	2.685**
	(0.641)	(0.881)	(0.861)	(0.899)	(0.898)	(0.878)	(0.880)	(0.934)	(0.999)	(1.062)
College diploma	5.465***	3.521***	4.112***	5.104***	4.953***	5.429***	5.178***	6.168***	8.076***	6.698***
	(0.751)	(1.033)	(1.009)	(1.054)	(1.053)	(1.030)	(1.033)	(1.095)	(1.171)	(1.245)
Post graduate degree	6.309***	4.402***	5.163***	6.406***	6.477***	7.135***	6.192***	7.003***	8.386***	7.680***
	(0.969)	(1.333)	(1.303)	(1.361)	(1.358)	(1.329)	(1.333)	(1.413)	(1.511)	(1.607)
Married	3.177***	2.792***	3.159***	3.387***	3.413***	3.257***	3.434***	3.730***	4.090***	3.098***
	(0.629)	(0.865)	(0.845)	(0.883)	(0.881)	(0.862)	(0.865)	(0.917)	(0.981)	(1.043)
Divorced	2.319***	2.646**	1.098	0.212	1.255	2.141*	1.759	2.333**	3.928***	2.292*
	(0.798)	(1.098)	(1.073)	(1.120)	(1.118)	(1.094)	(1.097)	(1.163)	(1.244)	(1.323)
Smoking	-4.288***	-4.865***	-4.690***	-4.499***	-5.169***	-5.880***	-5.865***	-5.574***	-4.026***	-3.381***
	(0.702)	(0.966)	(0.944)	(0.986)	(0.984)	(0.963)	(0.966)	(1.024)	(1.095)	(1.164)
Number of children	-0.534**	0.151	-0.072	-0.228	-0.287	-0.571	-0.537	-0.999***	-1.082***	-1.086***
	(0.254)	(0.349)	(0.341)	(0.356)	(0.356)	(0.348)	(0.349)	(0.370)	(0.396)	(0.421)
Employment	-0.798	0.340	0.158	0.024	-0.247	-0.682	-0.786	-1.528	-1.688	-1.034
	(0.677)	(0.931)	(0.909)	(0.950)	(0.948)	(0.928)	(0.930)	(0.986)	(1.055)	(1.122)
Monthly household income ($1000's)	0.098	0.097	0.033	0.133	0.210**	0.165*	0.150	0.126	0.129	0.060
	(0.070)	(0.096)	(0.094)	(0.098)	(0.098)	(0.096)	(0.096)	(0.102)	(0.109)	(0.116)
FAH expenditures per person	0.049***	0.027***	0.034***	0.053***	0.058***	0.058***	0.056***	0.061***	0.059***	0.064***
	(0.006)	(0.008)	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)	(0.009)	(0.010)	(0.010)
FAFH expenditures per person	0.002	0.014	0.012	0.004	0.008	-0.004	-0.001	-0.006	-0.010	-0.006
	(0.008)	(0.011)	(0.011)	(0.012)	(0.012)	(0.011)	(0.011)	(0.012)	(0.013)	(0.014)
F	21.60	7.772	9.813	13.47	15.18	17.19	15.51	14.91	13.67	10.04
Prob>F	0	0	0	0	0	0	0	0	0	0
R-squared	0.130	0.0509	0.0634	0.0850	0.0948	0.106	0.0967	0.0933	0.0862	0.0648
Adj R-squared	0.124	0.0444	0.0570	0.0787	0.0886	0.0999	0.0905	0.0870	0.0799	0.0584
Root MSE	12.87	17.70	17.29	18.06	18.03	17.64	17.69	18.76	20.07	21.33
Number of obs	2,481	2,481	2,481	2,481	2,481	2,481	2,481	2,481	2,481	2,481

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors are in parentheses.

Appendix to Chapter 3

This appendix includes Table A5, which presents the FAH HEI-2010 total score and its 12 individual components for Black SNAP participants and Black non-SNAP participants. Table A6 reports the estimated effect of SNAP participation on FAH HEI-2010 scores among Black and White households. Additionally, Tables A7 and A8 display the IVUQR estimates of SNAP's effect on each of the 12 components of the FAH+FAFH HEI-2010 and FAH HEI-2010 scores, respectively, for White households.

Figure A1 displays kernel density distributions of FAH HEI-2010 scores by household race and SNAP participation. Figure A1a focuses on racial differences between White and Black households. For FAH HEI-2010 scores, the density distribution for Black households shifts slightly to the left relative to White households, indicating lower nutritional quality among Black households when only FAH acquisition is considered. Figure A1b narrows the analysis to Black households, distinguishing between SNAP participants and non-SNAP participants. For FAH HEI-2010 scores, the density distributions of Black SNAP and non-SNAP households highly overlap at the lower tails, but non-SNAP households appear more frequently in the 60–90 point range. This indicates that SNAP's effects on FAH nutritional quality may primarily influence the middle and upper parts of the HEI-2010 score distribution rather than the lower end. Figure A1c compares White SNAP and non-SNAP households. Across FAH HEI-2010 scores, White non-SNAP households show a shift toward higher values. This pattern reflects generally better diet quality among White non-SNAP households.

Figures A2 and A3 present unconditional quantile regression estimates of the effect of SNAP participation on HEI-2010 scores for FAH+FAFH and FAH, respectively, among Black and White households. These estimates do not account for the potential endogeneity of SNAP participation. Figure A4 illustrates the distributional effects of SNAP on FAH HEI-2010 scores across the dietary quality distribution for both Black and White households. Figure A5 and Figure A6 present unconditional quantile regression estimates of the effect of SNAP participation on HEI-2010 scores for FAH+FAFH and FAH, respectively, among Black and White non-Hispanic households.

Table A5: FAH HEI-2010 Component Scores for Black Households

Component	Average Component Score for SNAP Participation	Average Component Score for Non-SNAP Participants	Mean difference	
	(1)	(2)	(3)	
FAH HEI-2010	44.26	48.09	-3.83	*
	(0.92)	(1.40)		
Adequacy Components:				
Total Vegetables	1.85	2.41	-0.56	***
	(1.68)	(1.99)		
Greens and Beans	0.80	1.09	-0.29	
	(1.51)	(1.84)		
Total Fruit	1.59	1.84	-0.25	
	(1.80)	(1.91)		
Whole Fruit	1.40	1.94	-0.53	**
	(1.84)	(2.13)		
Whole Grains	1.72	2.48	-0.76	**
	(2.65)	(3.24)		
Dairy	3.77	3.92	-0.15	
	(3.40)	(3.47)		
Total Protein Foods	3.37	3.42	-0.05	
	(1.97)	(1.88)		
Seafood and Plant Proteins	1.40	1.60	-0.20	
	(1.83)	(1.98)		
Fatty Acids	5.48	5.12	-0.36	
	(3.75)	(3.87)		
Moderation Components:				
Sodium	6.22	6.00	-0.22	
	(4.03)	(3.99)		
Refined Grains	6.57	6.66	-0.09	
	(3.53)	(3.98)		
Empty Calories	10.08	11.60	-1.52	**
	(6.63)	(6.88)		

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ indicate significant differences in average FAH HEI-2010 component score between Black SNAP and Black NonSNAP households. Standard deviations are reported in parentheses.

Table A6: The Effects of SNAP on FAH HEI-2010 Scores Among Black and White Households with Monthly Gross Income <= 200% of the Federal Poverty Level

	OLS	IV Regression	Percentiles of FAH HEI-2010 Score								
			10th	20th	30th	40th	50th	60th	70th	80th	90th
SNAP Black	-3.263*	-19.452**	-12.974	-13.590*	-19.491***	-12.234**	-13.484	-14.906**	-5.894	-5.093	-10.625
	(1.766)	(7.511)	(11.051)	(8.166)	(6.581)	(6.009)	(8.397)	(6.484)	(6.997)	(8.433)	(10.575)
SNAP White	-2.163***	-12.635	-16.103*	-21.673***	-23.890***	-13.618**	-14.031*	-15.742***	-5.575	-2.723	-8.382
	(0.872)	(9.986)	(9.381)	(7.908)	(6.471)	(6.022)	(7.829)	(5.921)	(6.669)	(7.962)	(10.358)
First-stage F-statistic		24.155									
First-stage P-value		0.007									
J Test statistic		9.873									
J Test P-value		0.274									

Note: All regressions include controls for household, individual, and state-level characteristics. The instrumental variables are state-level welfare and SNAP administrative policies. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are in parentheses and clustered at the region level.

Table A7: IVUQR Estimates of the Effects of SNAP on 12 FAH+FAFH HEI-2010 Component Scores Among White Households with Monthly Gross Income <= 200% of the Federal Poverty Level

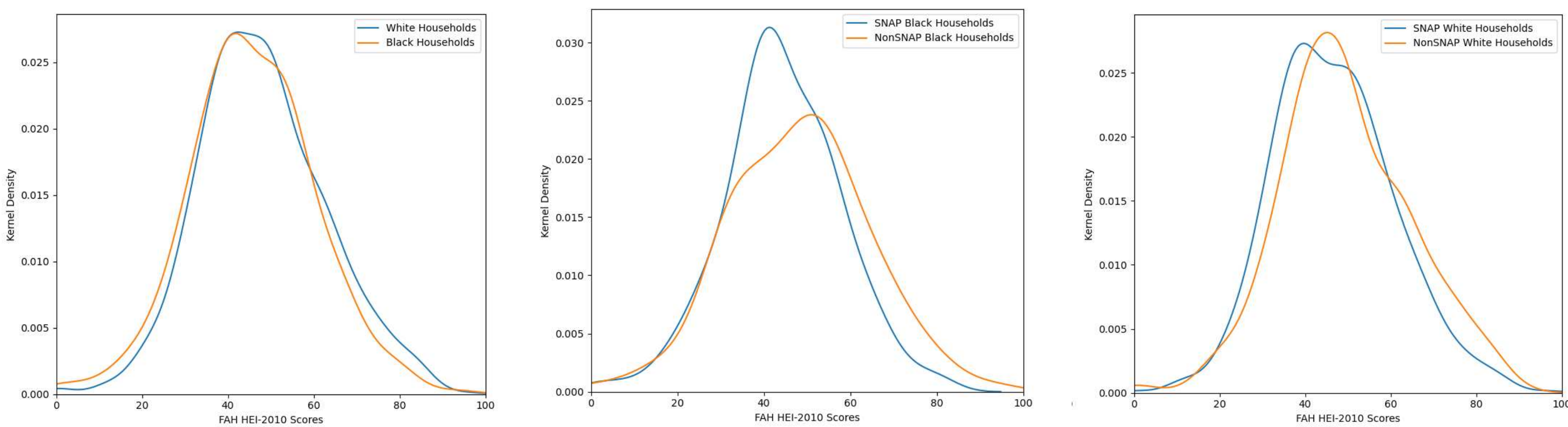
	Percentiles of FAH+FAFH HEI-2010 Component Score								
	10th	20th	30th	40th	50th	60th	70th	80th	90th
Total vegetables	-1.40**	-0.97	-0.79	-0.53	-0.98	-1.71*	-3.27***	-2.24**	0.00
	(0.68)	(0.64)	(0.65)	(0.67)	(0.75)	(0.91)	(1.21)	(1.00)	(0.00)
Greens and beans	-0.19	-0.19	-0.19	-0.31	-1.08*	-2.47**	-4.49***	-6.95**	-0.00
	(0.22)	(0.22)	(0.22)	(0.22)	(0.57)	(1.19)	(1.73)	(3.02)	(0.00)
Total fruits	-0.10	-0.19	-0.54	-0.77	-0.54	0.02	0.17	0.39	0.19
	(0.25)	(0.30)	(0.55)	(0.72)	(0.78)	(0.95)	(1.25)	(1.80)	(0.77)
Whole fruits	-0.29	-0.40	-0.83*	-1.93**	-3.05**	-2.54*	-2.60	-0.65	0.00
	(0.26)	(0.27)	(0.44)	(0.95)	(1.25)	(1.44)	(1.92)	(1.47)	(0.00)
Whole grains	-0.03	-0.03	-0.06	-0.53	-1.06	-1.64	-1.89	-1.30	-7.42*
	(0.24)	(0.24)	(0.24)	(0.46)	(0.74)	(1.01)	(1.51)	(2.14)	(3.93)
Dairy	0.60	0.85	0.36	0.77	0.23	-0.50	0.32	0.62	0.00
	(1.40)	(1.79)	(1.72)	(1.65)	(1.64)	(1.73)	(2.14)	(2.22)	(0.00)
Total protein foods	-1.46	-0.58	-0.19	-0.42	-0.24	0.00	0.00	0.00	0.00
	(1.29)	(1.13)	(1.26)	(1.11)	(0.35)	(0.00)	(0.00)	(0.00)	(0.00)
Seafood and plant proteins	0.25	0.26	0.36	0.13	0.23	0.46	-1.96	-1.86	0.00
	(0.25)	(0.27)	(0.44)	(0.87)	(1.10)	(1.55)	(2.07)	(1.72)	(0.00)
Fatty acids	-0.39	0.52	0.88	-0.33	-1.58	-0.73	-1.31	-1.41	0.00
	(0.49)	(1.93)	(1.64)	(1.56)	(1.88)	(2.34)	(2.80)	(3.31)	(0.00)
Sodium	-0.66	-4.41	-3.94	-3.77*	-1.80	-1.71	-3.87*	0.00	0.00
	(0.56)	(3.25)	(3.10)	(2.24)	(1.87)	(1.96)	(2.34)	(0.10)	(0.00)
Refined grains	0.76	3.87	-1.20	-0.76	0.98	1.82	-0.76	0.00	0.00
	(0.48)	(2.86)	(2.24)	(2.42)	(2.10)	(1.76)	(2.18)	(0.05)	(0.00)
Empty calories	-5.03	-5.68	-6.30**	-6.90**	-6.43**	-4.97*	-6.80**	-6.00*	-2.24
	(6.22)	(3.78)	(3.10)	(3.28)	(2.99)	(2.60)	(2.87)	(3.57)	(3.83)

Note: All regressions include controls for household, individual, and state-level characteristics. The instrumental variables are state-level welfare and SNAP administrative policies. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are in parentheses and clustered at the region level.

Table A8: IVUQR Estimates of the Effects of SNAP on 12 FAH HEI-2010 Component Scores Among White Households with Monthly Gross Income <= 200% of the Federal Poverty Level

	Percentiles of FAH HEI-2010 Component Score								
	10th	20th	30th	40th	50th	60th	70th	80th	90th
Total vegetables	-0.08	-0.41	-1.81**	-0.89	-0.96	-2.20*	-2.81*	-0.92	0.00
	(0.38)	(0.65)	(0.91)	(0.93)	(0.92)	(1.18)	(1.63)	(0.84)	(0.00)
Greens and beans	-0.16	-0.16	-0.16	-0.16	-0.16	-0.16	-3.48	-4.74*	-0.00
	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(0.15)	(2.30)	(2.51)	(0.26)
Total fruits	-0.12	-0.12	-0.21	-0.40	-0.76	-0.84	-1.15	0.17	0.00
	(0.26)	(0.26)	(0.29)	(0.83)	(0.98)	(1.13)	(1.63)	(2.30)	(0.00)
Whole fruits	-0.51**	-0.51**	-0.51**	-1.01*	-2.82*	-4.70**	-4.31*	0.00	0.00
	(0.24)	(0.24)	(0.24)	(0.54)	(1.58)	(2.05)	(2.47)	(0.53)	(0.00)
Whole grains	-0.40**	-0.40**	-0.40**	-0.40**	-2.46**	-3.09**	-2.73	-2.97	-6.60
	(0.20)	(0.20)	(0.20)	(0.20)	(1.05)	(1.50)	(2.07)	(2.80)	(5.52)
Dairy	0.04	-0.31	0.66	0.94	0.70	-0.21	-1.31	-0.20	0.00
	(0.52)	(1.85)	(2.18)	(2.11)	(2.01)	(2.36)	(2.78)	(0.84)	(0.00)
Total protein foods	-0.29	-1.03	-0.84	0.71	0.88	0.00	0.00	0.00	-0.00
	(0.68)	(1.80)	(1.68)	(1.68)	(1.36)	(0.12)	(0.00)	(0.00)	(0.00)
Seafood and plant proteins	-0.04	-0.04	-0.04	-0.11	-0.75	-0.92	-3.60	-1.20	0.00
	(0.22)	(0.22)	(0.22)	(0.26)	(1.23)	(1.92)	(2.59)	(1.77)	(0.00)
Fatty acids	0.11	0.14	0.45	-0.36	-1.41	1.03	0.67	0.00	0.00
	(0.39)	(0.44)	(2.44)	(2.05)	(2.47)	(3.14)	(3.43)	(0.30)	(0.00)
Sodium	-0.48	-6.14	-4.59	-3.57	-2.57	-1.20	0.00	0.00	0.00
	(0.49)	(4.75)	(3.04)	(2.69)	(2.57)	(1.59)	(0.00)	(0.00)	(0.00)
Refined grains	0.73*	5.03	-0.50	1.26	2.55	2.46	0.00	0.00	0.00
	(0.44)	(4.05)	(2.89)	(3.36)	(2.44)	(2.37)	(0.01)	(0.00)	(0.00)
Empty calories	-0.29	-3.31	-5.48	-8.89**	-7.78**	-11.59***	-10.86***	-3.65	0.00
	(0.51)	(4.35)	(4.76)	(4.20)	(3.92)	(4.49)	(3.91)	(3.44)	(0.00)

Note: All regressions include controls for household, individual, and state-level characteristics. The instrumental variables are state-level welfare and SNAP administrative policies. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are in parentheses and clustered at the region level.



(a) White and Black Households

(b) Black Households by SNAP Status

(c) White Households by SNAP Status

Figure A1: Distributions of FAH HEI-2010 Scores by Household Race and SNAP Participation

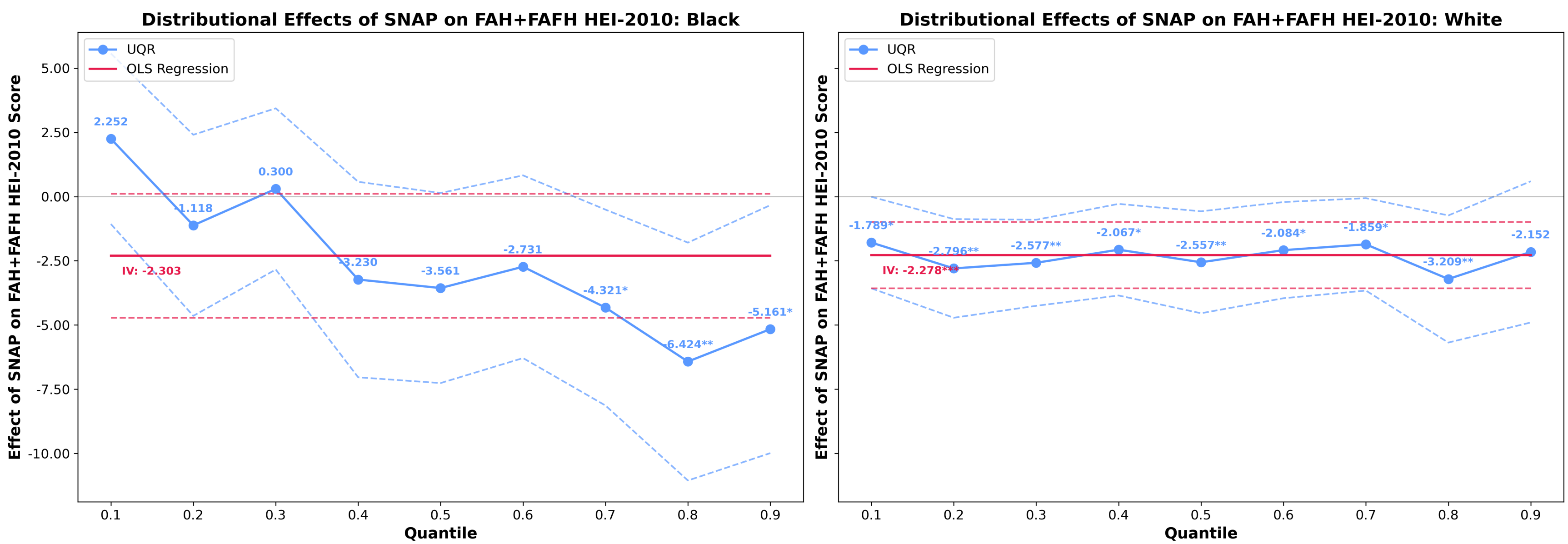


Figure A2: The Distributional Effects of SNAP on FAH+FAFH HEI-2010 Scores Among Black and White Households with Monthly Gross Income <= 200% of Federal Poverty Level, Without Instrumental Variables.

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are bootstrapped and clustered at the regional level. Confidence intervals are based on a 90% confidence level.

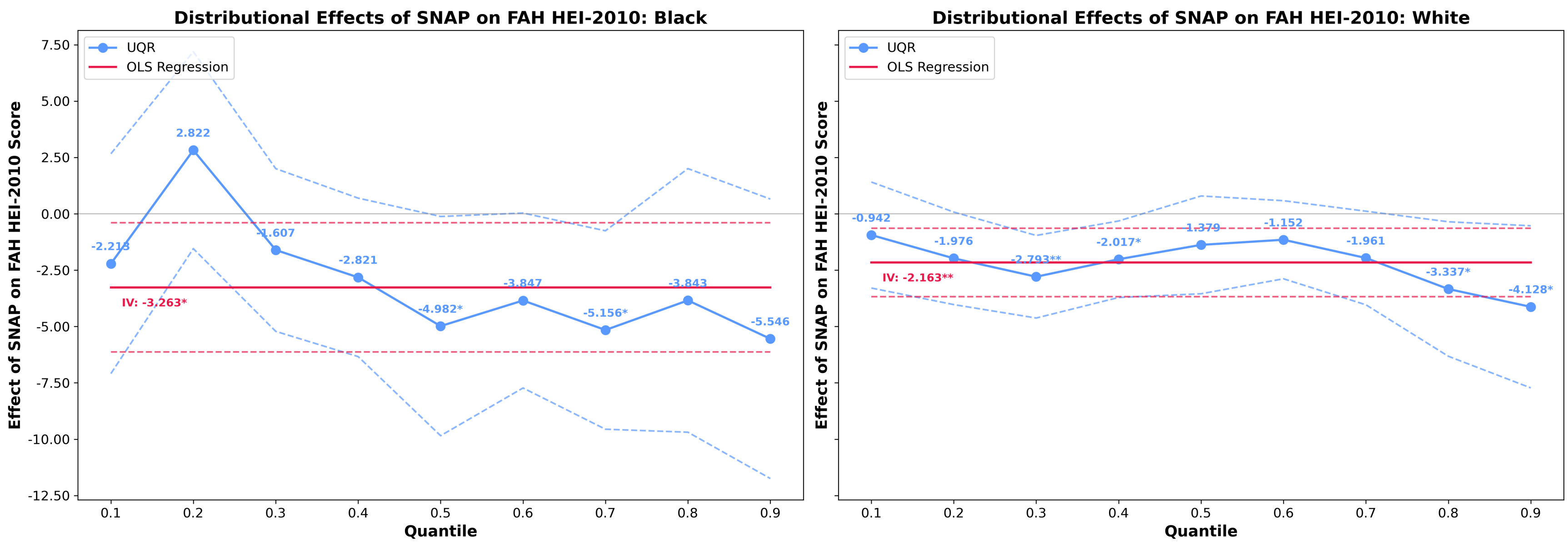


Figure A3: The Distributional Effects of SNAP on FAH HEI-2010 Scores Among Black and White Households with Monthly Gross Income <= 200% of Federal Poverty Level, Without Instrumental Variables

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are bootstrapped and clustered at the regional level. Confidence intervals are based on a 90% confidence level.

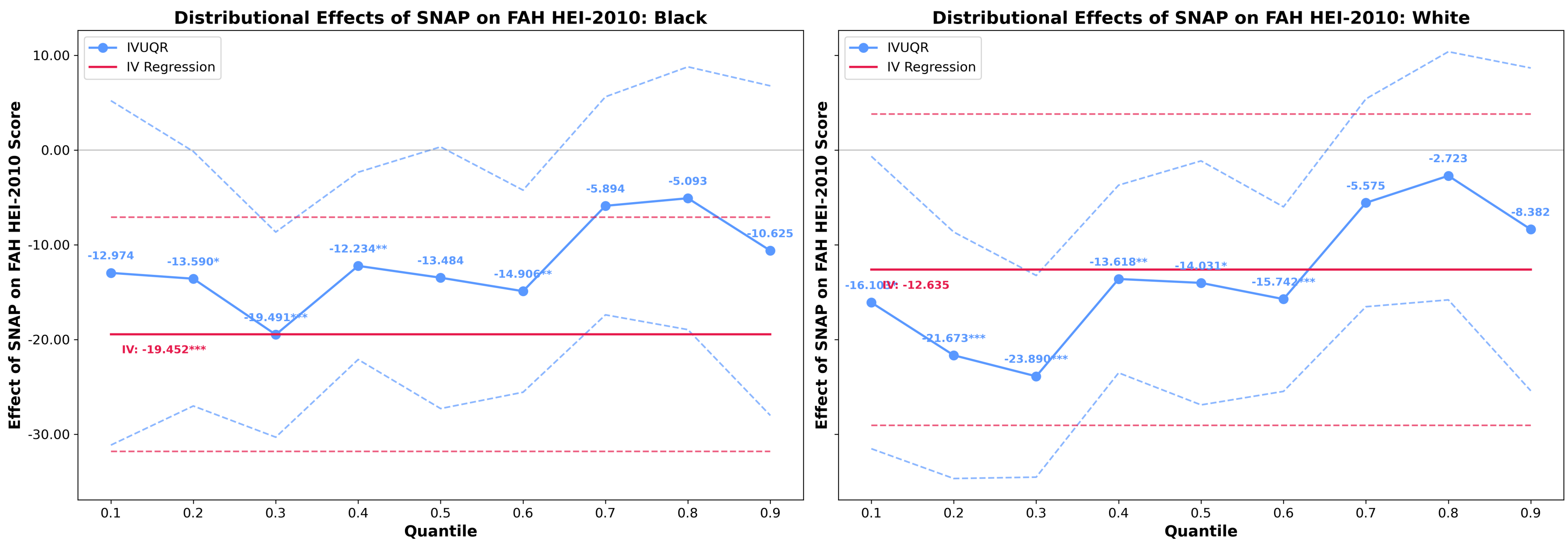


Figure A4: The Distributional Effects of SNAP on FAH HEI-2010 Scores Among Black and White Households with Monthly Gross Income <= 200% of Federal Poverty Level

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are bootstrapped and clustered at the regional level. Confidence intervals are based on a 90% confidence level.

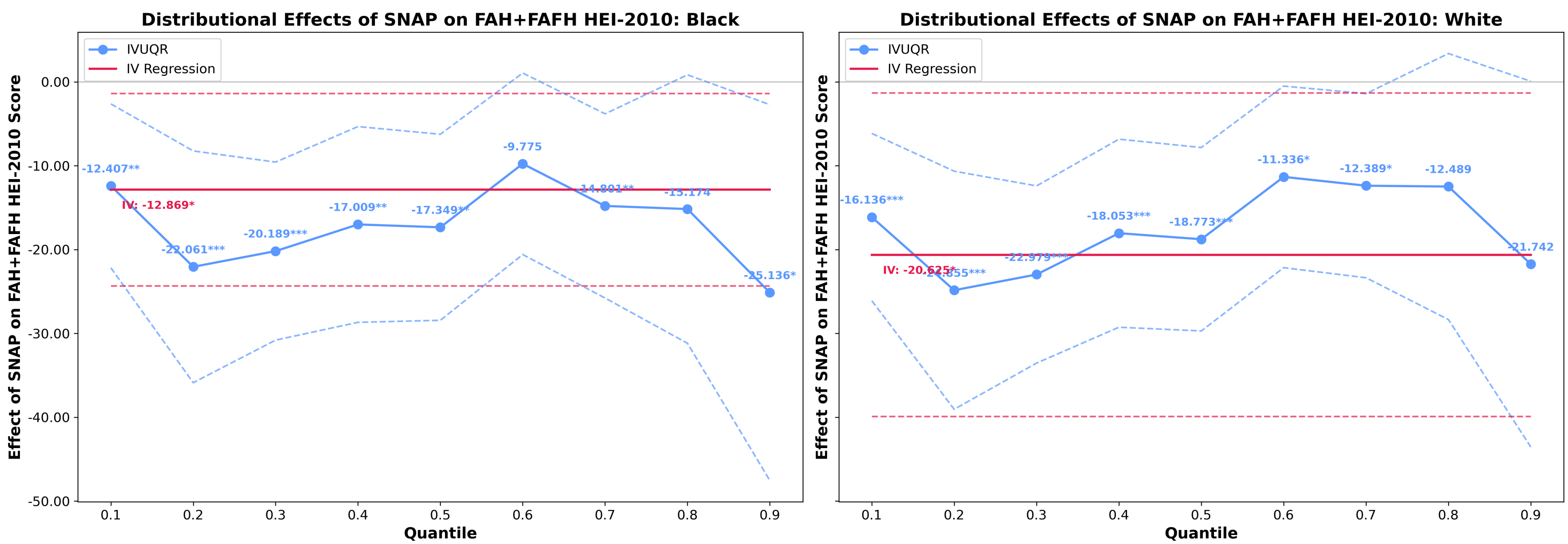


Figure A5: The Distributional Effects of SNAP on FAH+FAFH HEI-2010 Scores Among Black and White non-Hispanic Households with Monthly Gross Income <= 200% of Federal Poverty Level.

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are bootstrapped and clustered at the regional level. Confidence intervals are based on a 90% confidence level.

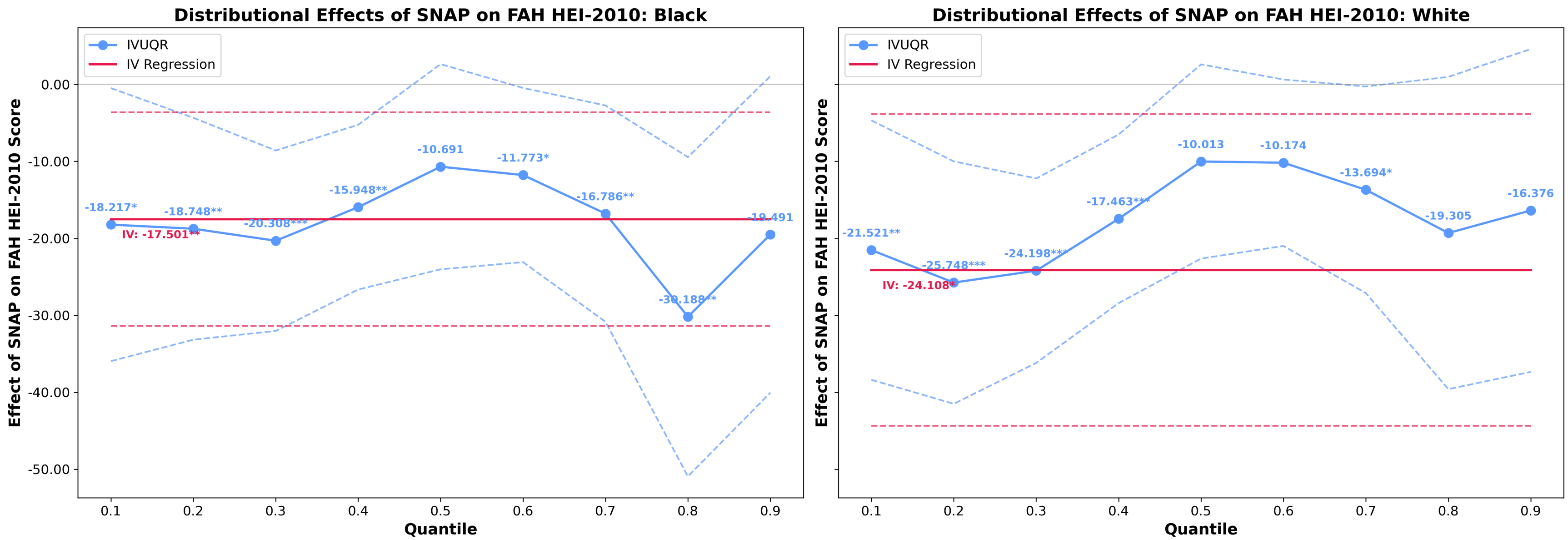


Figure A6: The Distributional Effects of SNAP on FAH HEI-2010 Scores Among Black and White non-Hispanic Households with Monthly Gross Income <= 200% of Federal Poverty Level

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. The standard errors are bootstrapped and clustered at the regional level. Confidence intervals are based on a 90% confidence level.