Distributional Analyses on Diet Quality in the United States

A THESIS
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Travis Alan Smith

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTORATE OF PHILOSOPHY

TIMOTHY K.M. BEATTY

AUGUST, 2014

Acknowledgements

I am grateful for my parents and their support in my desire to pursue a PhD. I am also thankful for my wonder wife and her unyielding support throughout these past four years.

I would also like to thank my advisor, Tim Beatty, for his encouragement to push myself to my full abilities, and the members of my advisory committee – Ezra Golberstein, Rob King and Joe Ritter – for their valuable input.

And finally, a special thanks to my all my friends and fellow classmates.

Dedication

To my wife, Jessica, and my dog, Sammy.

Abstract

This dissertation takes a distributional approach to examining dietary quality in the United States. Diet quality is a direct input to health, is often used as a proxy for well-being, and is an outcome variable for a wide variety of economic interventions. This makes diet quality a particularly important, yet understudied, outcome for program evaluation and describing food bundles that individuals choose.

The first chapter describes the evolution of adult dietary quality in the U.S. over the last two decades. Contrary to popular wisdom, there have been statistically significant improvements at all levels of diet quality. For the population as a whole, we find significant improvements across all levels of diet quality. Further, we find improvements for both low-income and higher-income individuals alike. Counterfactual distributions of dietary quality are constructed to investigate the extent to which observed improvements can be attributed to changes in the nutritional content of foods and to changes in population characteristics. We find that 63% of the improvement for all adults can be attributed to changes in food formulation and demographics. Changes in food formulation account for a substantially larger percentage of the dietary improvement within the lower-income population (19.6%) as compared to their higher-income counterpart (6.4%). The sheer myriad of overlapping policies and public awareness initiatives during this time period make it difficult to pin down the exact causes behind such improvements. This chapter motivates two program evaluation studies in the two chapters that follow.

The second chapter estimates distributional effects of food consumed at school and away from home on child dietary quality. Using a fixed-effects quantile estimator, two non-consecutive days of food intake are used to identify the effect of eating away from home and at school. I find considerable heterogeneity in the estimated impacts. The study finds that food away from home, as compared to home-prepared food, has a negative impact on the distribution of dietary quality except at low quantiles. Main results suggest that school food has both positive and negative impacts across the distribution of dietary quality. I find positive impacts on dietary quality at low quantiles of the outcome distribution, whereas food from school has a negative impact at the upper

end of the distribution of diet quality. While food consumed under the National School Lunch and Breakfast Programs may not benefit every child, especially the average child, it does improve the diets of many children who otherwise would have poorer dietary quality. The implication is that U.S. schools are fertile grounds to improve nutrition skill formation, especially for the most nutritionally disadvantaged.

This final chapter estimates the effect of replacing food assistance benefits, which typically come in the form of a food voucher, for an equal value of cash on the quantity and quality of food consumed in a household. We utilize an experiment in which a portion of beneficiaries were chosen at random to receive their benefits in the form of cash. We take a distributional approach because we believe it is important to analyze low-consuming households separate from high-consuming households. We find some evidence that a cash system would increase kilocalorie consumption in the portion of the distribution below recommended levels of consumption and decrease consumption in the portion of the distribution well above any reasonable threshold. This finding implies that a cash transfer system may both alleviate food insecurity and decrease overconsumption. The cash system appears to have a positive impact on the distribution of dietary quality in quantiles above 40. Virtually all of the improvement in quality comes from a decrease in consumption of less-healthy foods by the cash receiving group. Overall, these findings imply that beneficiaries are no worse off under a cash transfer system and in fact, may be better off.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Diet-related illness and disease not only generate direct individual medical costs but also create negative externalities in the form of higher health insurance premiums, greater public health care expenditures, losses in worker-productivity and lower tax revenues (Cawley, 2004). Nutrition in general is highly correlated with four of ten major causes of death in the United States: coronary heart disease, cancer, stroke, and type 2 diabetes (Jemal et al., 2008). The *quality* of one's diet specifically is associated with increased risks of coronary heart disease, stroke and diabetes (Chiuve et al., 2012), cardiovascular disease (Nicklas, O'Neil and Fulgoni, 2012) and many types of cancer (Bosire et al., 2013; Reedy et al., 2008; Shahril, 2012). Therefore, it is important to monitor and understand how Americans' diet quality changes as we learn more about its importance to health and productivity.

Understanding and improving dietary quality has been a longstanding policy initiative of the U.S. Federal government. In 1977, the Federal Government issued its first official recommendations: the Dietary Goals for Americans (USDA, 2008). These recommendations later became the Dietary Guidelines for Americans (DGA) in 1980 and are now in their seventh incarnation, the 2010 DGA. Most recently, major informational campaigns have included the *Food Guide Pyramid* released in 1992 (subsequently updated in 2005 as the *MyPyramid* and in 2011 as the *MyPlate*) and the 1994 nutrition label mandate.

Given the complexity of the linkages between diet and health, as well as the broad implications diet has on policy formation, it is important to better understand the interplay between diet quality and various nutrition assistance programs. Nutrition assistance programs make up over half of the total Farm bill, with the Supplemental Nutritional Assistance Program (SNAP) garnering roughly half while the school food programs make up about another 10%. These two programs in particular are under constant scrutiny and evaluation. This dissertation will look at (1) how diet quality has changed over the past twenty years and what has contributed to these changes, (2) how SNAP could impact diet quality (and quantity) under a cash transfer system rather than its current voucher system and (3) how the school food programs stack up against home prepared meals.

In all analyses, I use a measure of diet quality developed by the U.S. Department of Agriculture as my main metric – the Healthy Eating Index-2005. This particular measure has be shown to be highly correlated with other measures of diet quality and correlates well with diet-related disease (see Chiuve et al., 2012). I take a distributional approach in each chapter because I believe it is important to analyze low dietary quality individuals separate from high dietary quality individuals. In each case, the distributional approach proves to be a valuable tool in analyzing the effects in the tails of the outcome distribution.

Chapter 2

Is Diet Quality Improving? Distributional Changes in the United States, 1989–2008

2.1 Introduction

Poor nutrition is a contributing factor to four of ten major causes of death in the United States: coronary heart disease, cancer, stroke, and type 2 diabetes (Jemal et al., 2008). Poor diet quality is associated with increased risks of coronary heart disease, stroke and diabetes (Chiuve et al., 2012), cardiovascular disease (Nicklas, O'Neil and Fulgoni, 2012), breast cancer (Shahril, 2012), colorectal cancer (Reedy et al., 2008) and prostate cancer (Bosire et al., 2013). Moreover, diet quality is often used as a measure of well-being in developing countries (Ravaillon, 1996) and developed countries (Strauss and Duncan, 1998). In this paper, we study how the distribution of adult diet quality in the United States has evolved over the last two decades.

Improving dietary quality has long been a focus of government policy because of its direct impact on human health, particularly among the poor. Specific interventions have included increasing the resources available to households to buy food (e.g., Supplemental Nutrition Assistance Program–SNAP)[1] and providing healthy foods directly to individuals (e.g., the School Breakfast Program, School Lunch Program, Special Supplemental Nutrition Program for Women, Infants, and Children–WIC, and Fresh Fruit and Vegetable Program). Policies have also aimed at increasing the information available to individuals about what constitutes a healthy diet: the *Food Guide Pyramid* was released in 1992 and subsequently updated in 2005 as the *MyPyramid* and in 2011 as the *MyPlate*, Federally approved SNAP-Education programs grew from 7 active States in 1992 to 50 in 2004, mandatory nutrition labeling was enacted in 1994 and mandatory calorie postings in restaurants was introduced in 2011. Current policy proposals seek to improve diet quality by restricting the range of foods eligible for purchase under SNAP and change the relative prices of foods via taxes or subsidies.

In this paper, we use stochastic dominance to compare the distribution of dietary quality over time and between income groups. Stochastic dominance is frequently used in the economics literature to analyze the distribution of income or wealth. This empirical approach allows us to completely characterize the nature of the changes in dietary quality over time, paying close attention to low-income individuals, whose diets are of particular concern to policymakers. Stochastic dominance is particularly well suited to

[1]The Food and Nutrition Act of 2008 specifically aims "to provide for improved levels of nutrition among low-income households."

studying diet quality, where exact thresholds between "good" diets and "poor" diets is fuzzy.

Further, we construct counterfactual distributions of dietary quality to investigate the extent to which observed improvements can be attributed to changes in the nutritional content of foods and to changes in demographics. In short, we ask how would the distribution of dietary quality change if food in 1989 were formulated in as it was in 2008? Further, what would have the distribution of dietary quality looked like in 1989 had the demographic landscape of 2008 prevailed?

When comparing the observed distributions of dietary quality, we find a statistically significant and economically meaningful improvement across the entire population over the period 1989–2008. Improvements occur for individuals in households above and below our chosen poverty threshold. Counterfactual estimates indicate that 53.3 percent of the dietary improvement in the U.S. population can be attributed to changes in demographics (i.e., an aging, more educated, and ethnically diverse population) and an additional 10.1 percent of the improvement is attributed to changes in food composition (e.g., decreases in saturated fats, sugars and sodium). The residual 36.6 percent is unexplained by either changes in demographics or food composition.

The paper proceeds as follows. We begin by describing a widely used measure of dietary quality – the Healthy Eating Index (HEI-2005) – that forms the basis of our analysis. We then turn to a description of our primary data sources, the National Health and Nutrition Examination Survey (NHANES) and the earlier Continuing Survey of Food Intakes by Individuals (CSFII); we extend the HEI-2005 to the earlier study period 1989-91. We then motivate our empirical approach by providing a brief overview of stochastic dominance. Following the presentation of results, we discuss the economic and policy implications in the final section.

2.2 Measuring Diet Quality

The healthfulness of an individual's diet depends on two factors: *energy balance* and *dietary quality*. Energy balance is the relationship between calories consumed and energy expended, which results in body weight management (Hall et al., 2012). Dietary quality can be expressed as a per calorie metric that measures the degree to which a diet complies with a set of criteria (here, the Dietary Guidelines for Americans via the Healthy Eating Index). In this paper, we focus on dietary quality, and we note that there is evidence that higher quality diets are associated with decreased obesity rates (Epstein et al. 2001, 2008), i.e. improved energy balance.

Each chapter of this dissertation uses the Healthy Eating Index (HEI) – developed in 1995 in order to measure compliance with the U.S. government's recommendations for healthful eating – as the measure of diet quality. The HEI has been widely used and evaluated as a valid measure of diet quality (Guenther et al., 2008). In the medical literature it has been found to be a significant predictor of medical outcomes, notably of all cause mortality, mortality due to malignant neoplasms (Ford et al., 2011), and overweight and obesity (Guo et al., 2004). Further, the HEI has been extensively used by economists to measure the outcome of policy interventions, for example Welfare Reform (Kramer-LeBlanc, Basiotis and Kennedy, 1997), School Breakfast Program (Bhattacharya, Currie and Haider, 2006), Food Stamps and WIC (Wilde, McNamara and Ranney, 1999), nutrition labeling (Kim, Nayga and Capps, 2001) and unusually cold weather (Bhattacharya et al., 2003). Finally, it is has also been found to be associated with food insecurity (Bhattacharya, Currie and Haider, 2004) and has been proposed as a possible indicator of food deserts (Bitler and Haider, 2011).

Every five years the *Dietary Guidelines for Americans* (DGA) are revised by the U.S. Departments of Agriculture (USDA) and Health and Human Services (HHS), based on the advice of an expert advisory panel. These guidelines are the U.S. Government's official recommendations for healthful eating and form the basis for information provided to consumers. Many of the USDA's food-assistance programs must be in compliance with the DGA. The HEI was updated in 2005 to reflect the 2005 DGA (frequently called the HEI-2005, see Guenther, Reedy and Krebs-Smith, 2008).[2] Because the HEI-2005 was constructed with the 2005 DGA as its basis, one can think of using this index as a consistent measure of dietary quality with 2005 defined as the base period.

The HEI (henceforth, HEI refers to the HEI-2005) is the sum of 12 components based on consumption of various foods or nutrients. Each component assigns a score ranging from 0 to 5 (total fruit, whole fruit, total vegetables, dark green/orange vegetables and

[2]For a comprehensive review of dietary indices see Kant (1996) and Kourlaba and Panagiotakos (2009).

legumes, total grains, whole grains), 0 to 10 (milk, meats and beans, oils, saturated fat, sodium) or 0 to 20 for the percentage of calories from solid fats, alcoholic beverages, and added sugars (SoFAAS) creating a maximum score of 100. Table C.1 provides exact details of the scoring.

Table 2.1: Healthy Eating Index-2005 Standards for Scoring.

	Score				
Component	0	5	8	10	20
Total fruit	0 $\rightarrow$	≥ 0.8 cup eq/1000 kcal			
Whole fruit	0 $\rightarrow$	≥ 0.4 cup eq/1000 kcal			
Total vegetables	0 $\rightarrow$	≥ 1.1 cup eq/1000 kcal			
Dark green/orange veg.	0 $\rightarrow$	≥ 0.4 cup eq/1000 kcal			
Total grains	0 $\rightarrow$	≥ 3.0 cup eq/1000 kcal			
Whole grains	0 $\rightarrow$	≥ 1.5 cup eq/1000 kcal			
Milk	0 $\longrightarrow$			≥ 1.3 cup eq/1000 kcal	
Meats and beans	0 $\longrightarrow$			≥ 2.5 oz eq/1000 kcal	
Oils	0 $\longrightarrow$			≥ 12 g/1000 kcal	
Saturated fat	≥ 15 $\longrightarrow$		10 $\longrightarrow$	$\leq 7\%$ of energy	
Sodium	≥ 2.0 $\longrightarrow$		1.1 $\longrightarrow$	≤ 0.7 g/1000 kcal	
Calories from SoFAAS[a]	≥ 50 $\longrightarrow$				$\leq 20\%$ of energy

Source: Recreated from Guenther et al. (2007).

[a]Solid Fat, Alcohol, and Added Sugar

There is debate among nutritionists about how a given HEI score maps into the notion of "healthy" versus "unhealthy" diet quality. One generally accepted rule of thumb is that total scores of more than 80 are considered "good," scores of 51-80 as "needs improvement," and scores of less than 51 as "poor." Characterizing a diet based on a single cut-off is difficult (analogous to characterizing what it means to be poor based on a poverty line). A key advantage of the stochastic dominance methods used in this research is that they allow general statements about improvements in dietary quality over time or between subpopulations without having to define a specific threshold.

It is worth repeating that the components of the HEI are density based (the ratio of an individual's component intake to their total calorie intake) rather than quantity based. By design, the HEI measures the *relative* quality of foods consumed, independent of total calories (and of energy expenditure). We use the total HEI score as the underlying metric of interest in this study for two reasons. First the HEI score has been extensively validated and tested as a measure of diet quality (Guenther et al., 2008). Second, joint tests of dominance are limited in practice to two or three dimensions, rather than the dozen component scores that make up the HEI.[3]

2.3 Data

Our sample uses nationally representative, repeated cross-sectional, individual food intake data from two surveys: the Continuing Survey of Food Intakes by Individuals (CSFII, 1989-91 and 1994-96), and the continuous waves of the National Health and Nutrition Examination Survey (NHANES, 2001-08). In both surveys, respondents report 24-hour dietary intakes and demographic information including income and household size.[4] Each survey wave is an independently drawn sample, which is representative of the U.S. with the USDA overseeing the food intake component in both surveys. Finally, for consistency across samples, we focus on adults 20 years and older.

The HEI-2005 is calculated by linking the USDA's MyPyramid Equivalents Database (MPED) to food intake surveys. The MPED decomposes individual foods into MyPyramid guideline equivalents so that each HEI component can be computed as shown in table C.1. As noted above, because there is no officially released MPED for the 1989-91 CSFII, the HEI-2005 has not been previously computed for surveys prior to 1994. Of the 3,953 unique foods reported by adults 20 and older on day one in the 1989-91 CSFII, 3,907 (98.8 percent) of these foods are also reported in the 1994-96 CSFII. We therefore use the 1994-96 MPED to calculate the HEI-2005 for individuals in 1989-91.[5]

We classify individuals as low-income if household income falls below 185-percent of the Federal Poverty Guidelines. This is a policy relevant threshold that serves as

[3]Duclos, Sahn and Younger (2006) provide a thorough discussion of multidimensional orderings using stochastic dominance. Alkire and Foster (2011) propose an alternative "counting" method that enables one to examine many dimensions with the caveat of having to choose a threshold *a priori*.

[4]For all surveys but the 2001-02 NHANES, a second day of dietary intake was obtained. In keeping with standard practice, we analyze the first day of intake. One alternative is to average day 1 and 2 intakes where available. Another approach is to estimate models of usual intake (see, Dodd et al., 2006). Assuming that measurement bias and within person variation, if present, is consistent across survey waves, our results are invariant to usual intake methods. As shown in the supplementary appendix online, results are robust to using two days of intake.

[5]The supplementary appendix online contains a description of how to map the MPED for 1994-96 CSFII to the 1989-91 CSFII in greater detail.

an upper bound on the cutoff for many Federal nutrition assistance programs. During our sample period, the cutoff for SNAP is 130 percent and 185 percent for WIC. The Federal Poverty Guidelines are also used as an eligibility criterion for the National School Lunch Program, School Breakfast Program, Child and Adult Care Food Program, and the Expanded Food and Nutrition Education Program.[6] Table 2.2 reports the mean HEI scores for the population as a whole and for individuals above and below 185 percent of the poverty line for each of the periods in our sample.[7]

Table 2.2: Healthy Eating Index–2005 Summary Statistics

Population	1989-91	1994-96	2001-04	2005-08
U.S. population	50.16 (13.97)^{a}	51.10 (13.88)^{a}	51.50 (11.91)	52.46 (12.49)
	[10.09, 96.42]	[10.69, 97.47]	[13.52, 99.46]	[8.78, 95.38]
N	9,498	9,867	8,640	9,258
Low-income	48.96 (19.83)^{a}	49.36 (15.45)^{a}	49.65 (13.29)^{a}	51.37 (14.99)
	[10.09, 90.25]	[10.69, 93.81]	[15.08, 99.46]	[8.78, 94.60]
N	4,965	3,433	3,551	3,857
Higher-income	50.56 (11.19)^{ab}	51.73 (13.16)^{ab}	52.36 (11.09)^{b}	52.92 (11.29)^{b}
	[11.51, 96.42]	[13.63, 97.47]	[13.52, 93.97]	[10.00, 95.38]
N	4,533	6,434	5,089	5,401

Note: Standard deviations in parentheses. Maxima and minima in brackets.

^{a}Within-population mean is significantly lower than 2005-08 at the 5-percent level.

^{b}Within-year higher-income mean is significantly different from low-income at the 5-percent level.

Table 2.2 shows a consistent pattern of increasing dietary quality across groups over time. Comparing the most recent period 2005-08 to the earlier periods, we see a significant increase (at the 5-percent level) for the population at large over 1989-91 and 1994-96. Low-income individuals appear to have a stagnant HEI score over 1989–2004, and then a significant increase in 2005-08. We also compare low and higher-income

[6]Federal Poverty Guidelines are updated each year to reflect changes in the Consumer Price Index for Urban consumers (CPI-U) and are a function of household income and size (U.S. Department of Health and Human Services, 2013).

[7]There are various ways to calculate the HEI score for a population of interest (see Freedman et al. (2008, 2010) for in-depth discussions). Because we are interested in the number and depth of *individuals* below a particular HEI score, we use the *mean score* of individuals instead of an alternative measure *score of the population ratio*. The mean score is computed by calculating each individual's HEI score and then averaging over the population, whereas the score of the population ratio is calculated as the population's total component intake over total calorie intake and then calculating each score from this population ratio.

individuals within year and find that higher-income individuals have significantly higher mean HEI scores for all years in the data, though in the final year of the data the mean HEI gap between groups is smallest.

2.4 Stochastic Dominance

We have seen that mean HEI scores have increased for all groups over the interval 1989–2008. But does the mean HEI obscure variation in dietary quality across individuals? For example, is the increase in diet quality due to general improvements across the population at a steady rate or due to larger improvements amongst those with the lowest (or highest) diet quality? To address these possibilities, we study the entire *distribution* of dietary quality for groups of interest using an approach common in the study of income and well-being, stochastic dominance.[8]

2.4.1 Definitions

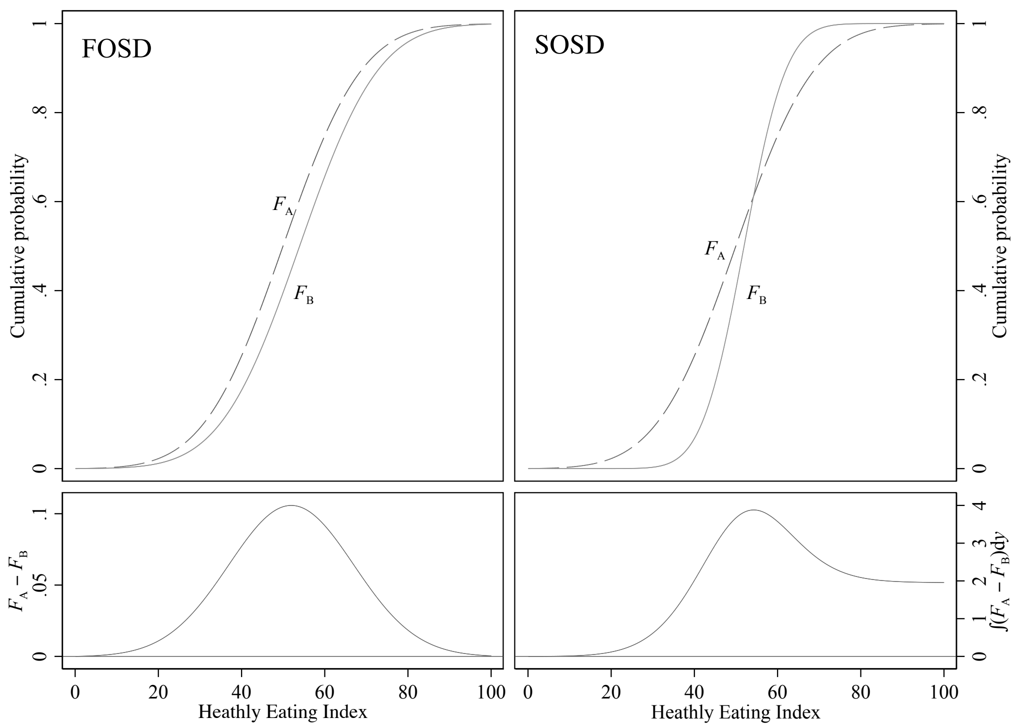
Consider two distributions of HEI scores with cumulative distribution functions $F_A(z)$ and $F_B(z)$, for a population of interest in two distinct time periods, or alternatively for two mutually exclusive subpopulations within a single time period. We say that distribution B *first-order* stochastically dominates (FOSD) distribution A if

$$F_B(z) \leq F_A(z) \; \forall \, z$$

with strict inequality for some z. In other words, no matter where the threshold for "healthy" is set, a greater share of the population characterized by distribution B have a "healthy" diet. This relationship is illustrated in the left panel of Figure 2.1.

[8]Stochastic dominance approaches have also been used to study changes in body mass index (Madden, 2011) and environmental quality (Maasoumi and Millimet, 2005), and extended to qualitative health measures (Allison and Foster, 2004).

Figure 2.1: First and second order stochastic dominance



Distributional studies of well-being often look to higher orders of stochastic dominance, notably second-order stochastic dominance (SOSD). While FOSD counts the number of individuals falling below a given 'healthy diet threshold' (which would in turn determine the "headcount ratio"), SOSD captures the depth, or severity of inadequate diets. SOSD is sensitive to the extent to which diets fall in the lower tails of the distribution.

To formally define SOSD, let $D_A^1(z) = F_A(z)$, and likewise for B, so that FOSD of B over A can be written as $D_B^1(z) \leq D_A^1(z)$. F_B will *second order* stochastically dominate F_A if

$$\int_0^z \left[D_B^1(y) - D_A^1(y)\right] dy \leq 0 \;\forall\, z$$

with a strict inequality for some value of z. This relationship is illustrated in the right panel of Figure 2.1 which shows that the CDFs cross thereby ruling out FOSD over the entire range of HEI. The integrated difference between F_A and F_B, shown in the subpanel, is strictly positive, and thus F_B second-order stochastically dominates F_A. More generally, dominance at order s of B over A is then defined as $D_B^s(z) \leq D_A^s(z)$ where,

$$D_j^s(z) = \int_0^z D_j^{s-1}(y)dy \;\forall\, z \quad \text{for } j = A, B$$

with a strict inequality for some value of z.

Stochastic dominance maps into social welfare under fairly standard assumptions about the utility derived from a healthy diet (Deaton, 1997). For example, if B FOSD A then for any social welfare function $\mathcal{W}$ defined on the distribution of diet quality $F(z)$ such that $\mathcal{W}(F) = \int U(z)dF(z)$ where U is *any* monotonically nondecreasing utility function of z ($U' \geq 0$), it must be true that social welfare derived from distribution B will be at least as high as the welfare derived from A. We can extend the mapping of social welfare to SOSD by requiring U to be nondecreasing and concave in z ($U' \geq 0$, $U'' \leq 0$). Note that because dominance of order s implies dominance of order $s + 1$, it follows that the latter is a less stringent condition. Thus, welfare implications are the strongest in the first-order case. Finally, we also make the standard assumption of anonymity, so that each individual is weighted equally in the social welfare function.

2.4.2 Estimation

A useful expression for $D_j^s(z)$ in empirical analyses is (Davidson and Duclos, 2000):

$$D_j^s(z) = \frac{1}{(s-1)!}\int_0^z (z-y)^{s-1} dF_j(y). \tag{2.1}$$

Integrating the empirical analogue of (2.1) by parts leads to a natural estimator of $D_j^s(z)$

$$\hat{D}_j^s(z) = \frac{1}{\hat{N}_j(s-1)!}\sum_{i=1}^{n_j} \theta_i (z-y_i)^{s-1} I(y_i \leq z) \tag{2.2}$$

where we account for complex survey design (e.g., CFSII and NHANES) by letting θ_i be an individual's sample weight, $\hat{N}_j = \sum_{i=1}^{n_j} \theta_i$ is the population size in distribution j (with corresponding sample size n_j), and $I(\cdot)$ is the indicator function. The first-order case leads to the empirical CDF

$$\hat{D}_j^1(z) = \hat{F}_j(z) = \frac{1}{\hat{N}_j}\sum_{i=1}^{n_j} \theta_i I(y_i \leq z) \tag{2.3}$$

and the statistic for the second order case follows directly.

2.4.3 Inference

We are interested in testing the hypothesis that the distribution of dietary quality in one time period dominates the distribution in another time period. For example, allowing distribution F_B be the more recent time period, the null hypothesis of an increase in dietary quality at order $s \in \{1, 2\}$ is,

$$H_0^{s+}: D_B^s(z) \leq D_A^s(z) \quad \forall\, z \quad \text{vs.} \quad H_a^{s+}: D_B^s(z) > D_A^s(z) \quad \text{for at least one } z$$

where the positive superscript denotes the hypothesis of dietary improvement. Whereas in testing the null hypothesis that dietary quality has decreased (denoted by H_0^{s-}), the signs would be reversed. One could also posit a null of equality but notice that rejection of both H_0^{s+} and H_0^{s-} implies rejecting equality.

Bishop, Formby and Thistle (1989, hereafter BFT) propose a multiple testing procedure by hypothesizing dominance in both directions. That is, testing the null of H_0^{s+}

and H_0^{s-} versus their respective alternatives and drawing inferences from the combined acceptance/rejection. A variety of approaches to drawing inferences based on the BFT procedure have been proposed, such as multiple comparison tests (Anderson, 1996; Davidson and Duclos, 2000) or Kolmogorov–Smirnov (KS) type tests (McFadden, 1989; Barrett and Donald, 2003; Linton, Massoumi, and Whang, 2005; Bennett, 2013). Multiple comparison approaches are based on arbitrarily chosen ordinates, which can lead to test inconsistency (Davidson and Duclos, 2000; Barrett and Donald, 2003). Therefore, in this study we use a KS type statistic that compares all objects within the support of the two distributions.

Let $\mathcal{Z}$ be defined as the union of the supports of A and B. Define the following functionals for each order s

$$\hat{d}_s^+ = \sup_{z \in \mathcal{Z}} [\hat{D}_B^s(z) - \hat{D}_A^s(z)] \tag{2.4}$$

$$\hat{d}_s^- = \sup_{z \in \mathcal{Z}} [\hat{D}_A^s(z) - \hat{D}_B^s(z)]. \tag{2.5}$$

Notice that that the null hypotheses can be rewritten in terms of these functionals. That is, the null of increased diet quality (F_B dominating F_A) at order s is simply $H_0^{s+}\colon \hat{d}_s^+ \leq 0$, and similarly for decreased diet quality (F_A dominating F_B) using $\hat{d}_s^-$. When the distributions are mutually independent, KS-type tests based on $\hat{d}_s^\pm$ are consistent (McFadden, 1989).[9] Test statistics are calculated using,

$$\hat{T}_s^\pm = \sqrt{\frac{n_A n_B}{n_A + n_B}} \hat{d}_s^\pm. \tag{2.6}$$

Because there are infinitely many $F_A(z)$ satisfying the null such that $F_B(z) \leq F_A(z)$, the limiting null distribution is not uniquely defined and depends on the underlying unknown distributions of F_A and F_B. We follow Barrett and Donald (2003) and use the *least favorable configuration* (LFC) to construct the null distribution. The LFC is the point in the null distribution that is least favorable to the alternative hypothesis

[9]The independence assumption seems reasonable given that our data are repeated cross-sections in which sampling units are independently drawn in each survey (see Bhattacharya (2007) for a more detailed discussion) and that some surveys were separated by nearly 20 years in time. In section 6, we relax this assumption to ignorability (also called, conditional independence or unconfoundedness) to construct aggregate counterfactual decompositions.

(i.e., $F_A = F_B$). As a result, the test is conservative; rejection of the null under the LFC implies rejection at *any* point in the null distribution. We construct a bootstrap distribution of $\hat{T}_s^{\pm}$ to simulate the p-values.

We use a recentering bootstrap approach, which has been shown to perform well against alternative methods (see, Barrett and Donald, 2003; Linton, Maasoumi and Whang, 2005). Let $\hat{D}_j^{s*}(z)$ be defined as above from (2.2) but computed on a random bootstrap sample drawn with replacement from distribution j.[10] The statistic is recentered by the observed values so that we have $\hat{D}_{jc}^{s*}(z) = \hat{D}_j^{s*}(z) - \hat{D}_j^{s}(z)$. We can then define recentered bootstrap functionals $\hat{d}_s^{*\pm}$ by replacing $\hat{D}_j^{s}(z)$ with $\hat{D}_{jc}^{s*}(z)$ in (2.4) and (2.5). The recentered bootstrap t-statistics are

$$\hat{T}_s^{*\pm} = \sqrt{\frac{n_A n_B}{n_A + n_B}} \hat{d}_s^{*\pm}.$$

We approximate p-values from the distribution of bootstrapped test statistics by

$$\hat{p}_s^{\pm} \simeq \frac{1}{B} \sum_{i=1}^{B} I(\hat{T}_s^{*\pm} > \hat{T}_s^{\pm}). \tag{2.7}$$

The p-values allow for a test of stochastic dominances at order s based on the rule "reject $H_0^{s\pm}$ if $\hat{p}_s^{\pm} < \alpha$" where α represents the conventional levels of statistical significance. Thus under the BFT procedure, rejection of the null H_0^{s-} in favor of H_a^{s-} coupled with a failure to reject H_0^{s+} is viewed as statistical evidence in favor of F_B dominated F_A at order s.

2.4.4 Robustness Check for First Order Stochastic Dominance

To determine the stochastic rankings of two distributions, we must distinguish between four possible true states of nature: the distributions are equal, A lies above B, A lies

[10] Our samples are constructed using multi-stage stratification where each stratum is clustered by two primary sampling units (PSUs). Test statistics based on a simple random bootstrap samples drawn with replacement would be biased and inconsistent. Under the CSFII and NHANES survey design, Rao, Wu, and Yue (1992) show that bootstrap replicate weights can be obtained by randomly picking one PSU within each stratum and internally rescaling the sample weights. We use the user written Stata package `bsweights` (Kolenikov, 2010) to automate the rescaling process to create $B = 1,000$ balanced replicate weights θ_i^* for each sample individual. These weights are used in equation (2.2) to create the bootstrap distribution of $\hat{T}_s^{*\pm}$.

below B, or the curves cross. The BFT procedure described above distinguishes between these four states by conducting two one-sided tests. The result is lower power in detecting a crossing of the CDFs, which could lead to over-classification of dominance (Dardanoni and Forcian, 1999; Gastwirth and Nayak, 1999). This is at least partially due to the fact that rejection of H_0^{1+} or H_0^{1-} by itself is consistent with both FOSD *and* a crossing; hence, the use of two one-sided tests to rule out the crossing under the BFT procedure.

A second drawback to the BFT procedure is how the total error probability α is apportioned to each one-sided test (Dardanoni and Forcian, 1999). As is typical with standard hypothesis testing, the one-sided critical value $c(\alpha)$ is based on ensuring that the probability of committing a Type I error (i.e., rejecting H_0^{1+}, H_0^{1-}, or both when they are true) is less than the nominal level α. But as noted by Dardanoni and Forcian (1999), the BFT procedure does not allow one to control how the total error probability α is allocated to each classification (equality, dominance in one direction, dominance in the opposite direction, or a crossing).

Bennett (2013) improves on the BFT procedure by writing it as a two-stage test that allows one to test for a crossing while giving the researcher flexibility in allocating the total error rate to each stage. Let α and β denote a pair of pre-specified significance levels for the first and second stage respectively. The first stage is to posit a null of equality ($F_A = F_B$) and determine rejection or acceptance based on the critical value $a(\alpha)$. If we accept the null, then we infer that the distributions are indistinguishable. Upon rejection, however, the second stage determines the state of nature among the three alternatives (A dominates B, B dominates A, or they cross) using the critical value $b(\alpha, \beta)$. This allows β to be the portion of the total error probability α allocated to a crossing (i.e., $\alpha\beta$) and the remaining $\alpha(1-\beta)$ is split evenly between dominance in either direction.

Bennett (2013) tabulates asymptotic critical values of $a(\alpha)$ and $b(\alpha, \beta)$ for frequently used significant levels. In the applications below, we wish to calculate the asymptotic p-values. To do so, we need to pre-set the total error rate α, the level at which we are controlling for falsely rejecting equality. We use two levels of significance (10 and 1 percent) so that the second stage is robust to our choice of α. The associated $a(\alpha)$ critical values are $a(0.1) = 1.2239$ and $a(0.01) = 1.6277$ (see, table 1 in Bennett, 2013).

The maximum of the one-sided test statistics found in (2.6) is used in the first stage and the minimum is used in the second stage. To simplify notation, let these statistics be $K_{max} = \max\{\hat{T}_1^+, \hat{T}_1^-\}$ and $K_{min} = \min\{\hat{T}_1^+, \hat{T}_1^-\}$, respectively. We are interested in the distribution of K_{min} conditional on rejecting equality. In other words, if K_{min} is "large enough" (i.e., larger than the second stage critical value $b(\alpha, \beta)$) conditional on $K_{max} > a(\alpha)$, then we reject the null in favor of a crossing. Asymptotically, as shown in proposition 2.6 in Bennett (2013), if $F_A = F_B$ and $b < a$ then

$$\lim_{n_A, n_B \to \infty} \mathbf{P}[K_{min} \leq b | K_{max} > a] \to \frac{2[G_1(b) - G(a, b)]}{1 - G_2(a)} \tag{2.8}$$

where[11]

$$\begin{aligned}
G_1(b) &= 1 - e^{-2b^2} \\
G_2(a) &= \frac{\sqrt{2\pi}}{a} \sum_{k=1}^{\infty} e^{-(2k-1)^2 \pi^2 / (8a)^2} \\
G(a, b) &= \sum_{k=-\infty}^{\infty} e^{-2(k(a+b))^2} - \sum_{k=-\infty}^{\infty} e^{-2(a+k(a+b))^2}.
\end{aligned}$$

The two-stage p-values (denoted $p_{1,\alpha}^{2S}$) are calculated from (2.8) where we use two levels of α. Thus, a $p_{1,\alpha}^{2S}$ value below conventional levels of significance is evidence that the distributions cross. Put differently, larger p-values are consistent with the null hypothesis that the distributions do not cross.

2.5 Results

Our main results are summarized in tables 2.3 and 2.4, and depicted in figures 2.2–2.4. Tables report the bootstrapped p-values for tests of increases and decreases in diet quality (Barrett and Donald, 2003), as well as the asymptotic two-stage p-values (Bennett, 2013). The final column summarizes the inferred ranking of distributions based on these tests. In short, we find that there has been a statistically significant and

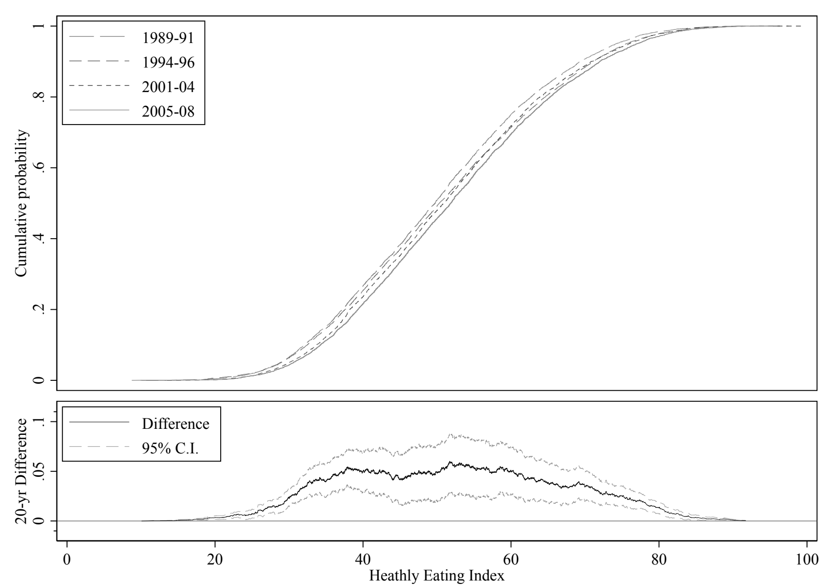
[11]The $G(\cdot)$ functions where taken from an earlier version of the two-stage test (Bennett, 2010). $G_1(b)$ is derived from the K-S distribution. We present $G_2(a)$ as a numerical approximation to the K-S cumulative distribution. As referenced in Bennett (2013), Billingsley (1968, p. 85) shows the closed-form expression of $G(a, b)$.

economically important improvement in the HEI scores over the period under study; Americans at all ranges of dietary quality are consuming a higher quality diet in 2005–2008 than they were in 1989–1991. However, there are differences between income groups with regards to when and where the improvements occurred (table 2.5).

2.5.1 Between Periods

Figure 2.2 shows the empirical CDFs for the U.S. adult population in each period. Distributions shift systematically to the right over time, in other words towards a more nutritious diet. Because the shifts are relatively small, in this and subsequent figures, we present the estimated difference between the earliest period (1989-91) and the latest period (2005-08) in a sub-panel. The area under the difference curve in the sub-panel is equal to the area between the distributions. We can see the twenty-year improvement was positive and pointwise statistically significant for the empirically relevant range of HEI scores.

Figure 2.2: Distribution of adult HEI-2005 scores in the U.S. population



Notes: Confidence intervals are calculated pointwise by bootstrapping. See online appendix for this figure in color.

From the first three rows of table 2.3, we see that in comparing 1989-91 to all subsequent periods the null of decreasing dietary quality is strongly rejected and in no case do we reject the null of an increase in diet quality. In comparing 1994-96 and 2001-04, we are unable to order the distributions in either the first or second order case. We do find strong evidence that 2005-08 FOSD 1994-96, but the results are fairly weak with regards to an ordering of 2005-08 and 2001-04.

Table 2.3: Tests of Stochastic Dominance among U.S. Adults

Distribution		Bootstrap Tests				Two-stage		Inferred
A	B	$\hat{p}_1^-$	$\hat{p}_2^-$	$\hat{p}_1^+$	$\hat{p}_2^+$	$p_{1,0.1}^{2S}$	$p_{1,0.01}^{2S}$	Ranking
1989-91	1994-96	0.007	0.010	0.900	0.660	0.507	0.362	$A \prec_1 B^{***}$
	2001-04	0.028	0.002	1.000	0.937	0.999	0.999	$A \prec_1 B^{**}$
	2005-08	0.002	0.000	1.000	0.877	0.981	0.966	$A \prec_1 B^{***}$
1994-96	2001-04	0.129	0.149	0.383	0.925	0.003	0.001	ND
	2005-08	0.010	0.003	0.999	0.863	0.981	0.965	$A \prec_1 B^{***}$
2001-04	2005-08	0.133	0.051	0.991	0.790	0.972	0.950	$A \prec_2 B^{*}$

Notes: The $\hat{p}_s^{\pm}$ values refer to one-sided tests of the null hypothesis $H_s^{\pm}$ using equation (2.7). The asymptotic $p_{1,\alpha}^{2S}$ values are calculated from (2.8), where $\alpha = 0.1, 0.01$. The notation $A \prec_s B$ reads "Distribution B dominates distribution A at order s," while ND indicates no dominance at order 1 or 2. Inferred ranking is based on statistical significance levels of $^{***}1$, $^{**}5$, and $^{*}10\%$.

Some care is required in interpreting the last two columns of table 2.3, as they report results from Bennet's two-stage test described in section 2.4.4. As noted above, these p-values are for the null hypothesis that the CDFs *do not* cross, as determined by both K_{max} and K_{max} being statistically large. Loosely speaking, these can be interpreted as the (conditional) probability of rejecting the hypothesis of *no* crossing. Thus, a $p_{1,\alpha}^{2S}$ value below conventional levels of significance can be interpreted as evidence that the distributions cross. Bennet's two-stage test supports the main findings above in that there is no statistical evidence that the 1989-91 distribution crosses any of the later years.

2.5.2 Between Income Groups

We now turn our attention to direct comparisons of individuals above and below 185% of the poverty guideline. As noted above, we choose 185% of the poverty line as our cut-off because it is an upper limit on the threshold for many federal nutrition assistance programs.[12] Panel (*a*) of figure 2.3 presents the empirical CDFs and the difference between 1989-91 and 2005-08 for low-income individuals; panel (*b*) likewise for higher-income individuals. Table 2.4 presents results from statistical tests of dominance by income group. For both groups, we find strong evidence that the distribution of dietary quality in 2005–2008 first-order stochastically dominates the distribution in the earliest period, with no evidence of a crossing.[13]

Results support the observation in table 2.2 that a significant portion of dietary improvement among low-income individuals occurred over the period 2001-08. For example, in comparing 1989-91 to 1994-96 we find no evidence of a partial ordering according to the bootstrap results, and the two-stage test confirms this by finding significant evidence of a crossing. In comparing 1989-91 to 2001-04, again the bootstrap results are silent on the ordering, as is the asymptotic test, indicating no dominance at orders 1 or 2. However, in comparing the most recent time period 2005-08 to any of the earlier distributions, all tests show a statistically significant, first-order, improvement in dietary quality, with no evidence of a crossing.

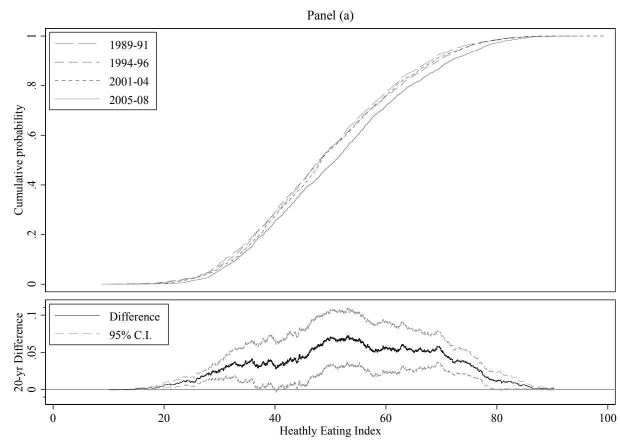
Comparing distributions among higher-income individuals, we can see that 1989-91 is first-order dominated by each subsequent period with no evidence a crossing. Comparing 1994-96 to 2001-04, the bootstrap results indicate a weak rejection of H_s^-, which could lead one to infer a partial ordering. However, when consulting the asymptotic two-stage test, we find significant evidence of a crossing, thereby ruling out first-order dominance. We do see that 2005-08 FOSD 1994-96, but we cannot rank the two most recent time periods.

[12] As pointed out by a referee, the health-education gradient is also of considerable interest. Although our main focus here is on income, we present dominance results by education in the supplementary appendix online for interested readers.

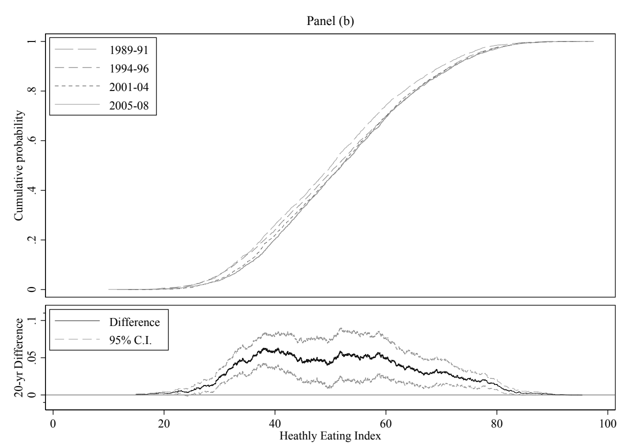
[13] As a sensitivity check, we also considered poverty thresholds of 75 to 250% of the Federal Poverty Guidelines in 25% increments. In all cases, both the low- and higher-income group, as defined by the various thresholds, exhibited a first-order dietary improvement over the 20 year period at less than 5-percent significance levels.

Figure 2.3: Distribution of adult HEI-2005 scores by income group

(a) Low-income



(b) Higher-income



Notes: Confidence intervals are calculated pointwise by bootstrapping. See online appendix for this figure in color.

We can compare the total twenty year improvements in each income group by examining sub-panels (a) and (b) in figure 2.3. We see that low-income individuals experienced relatively smaller increases over the bottom tail of relevant range of HEI as compared to their higher-income counterparts. We can more formally investigate this finding by taking the difference (between above and below 185% of the poverty line)

Table 2.4: Tests of Stochastic Dominance among U.S. Adults by Income Group

Distribution		Bootstrap Tests				Two-stage		Inferred
A	B	$\hat{p}_1^-$	$\hat{p}_2^-$	$\hat{p}_1^+$	$\hat{p}_2^+$	$p_{1,0.1}^{2S}$	$p_{1,0.01}^{2S}$	Ranking
Low-income								
1989-91	1994-96	0.218	0.257	0.570	0.654	0.027	0.009	ND
	2001-04	0.290	0.140	0.977	0.952	0.927	0.878	ND
	2005-08	0.008	0.000	1.000	0.882	0.998	0.996	$A \prec_1 B^{***}$
1994-96	2001-04	0.332	0.300	0.575	0.889	0.034	0.012	ND
	2005-08	0.006	0.003	0.991	0.855	0.984	0.971	$A \prec_1 B^{***}$
2001-04	2005-08	0.031	0.033	0.997	0.818	0.998	0.996	$A \prec_1 B^{**}$
Higher-income								
1989-91	1994-96	0.007	0.006	0.880	0.684	0.429	0.289	$A \prec_1 B^{***}$
	2001-04	0.004	0.000	0.999	0.897	0.999	0.999	$A \prec_1 B^{***}$
	2005-08	0.002	0.000	1.000	0.906	0.985	0.973	$A \prec_1 B^{***}$
1994-96	2001-04	0.083	0.106	0.553	0.932	0.032	0.011	ND
	2005-08	0.007	0.010	0.991	0.886	0.957	0.926	$A \prec_1 B^{***}$
2001-04	2005-08	0.135	0.137	0.913	0.697	0.555	0.410	ND

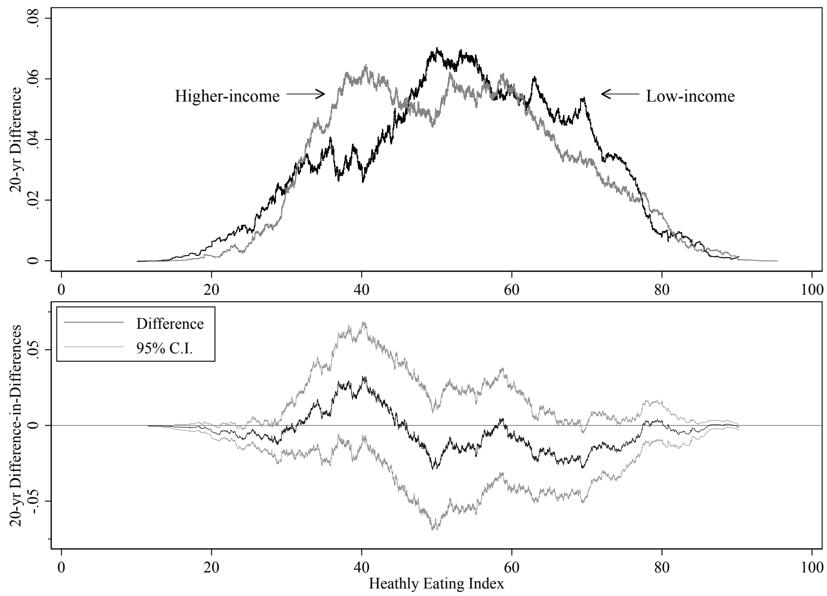
Notes: The $\hat{p}_s^{\pm}$ values refer to one-sided tests of the null hypothesis $H_s^{\pm}$ using equation (2.7). The asymptotic $p_{1,\alpha}^{2S}$ values are calculated from (2.8), where $\alpha = 0.1, 0.01$. The notation $A \prec_s B$ reads "Distribution B dominates distribution A at order s," while ND indicates no dominance at order 1 or 2. Inferred ranking is based on statistical significance levels of $^{***}1$, $^{**}5$, and $^{*}10\%$.

in the differences (between the earlier and later periods). Figure 2.4 superimposes the subfigures in panels (a) and (b) of figure 2.3 in the top panel and then plots the difference between the two in the bottom panel. That is, in the subpanel of figure 2.4 we plot:

$$DD = \left[\hat{D}^1_{high,89}(z) - \hat{D}^1_{high,08}(z)\right] - \left[\hat{D}^1_{low,89}(z) - \hat{D}^1_{low,08}(z)\right].$$

As shown in figure 2.4, considering lower levels of dietary quality below a HEI of 45 we find higher-income individuals experienced a greater improvement over 1989–2008 than low-income individuals. Whereas at higher levels of the HEI distribution, low-income individuals experienced greater increases in dietary quality. In other words, we find some evidence that within the poor dietary quality population, low-income individuals experienced less improvement over the 20-year period as compared to the higher-income individuals.

Figure 2.4: Differences in dietary improvements amongst low- and higher-income populations



Notes: Confidence intervals are calculated pointwise by bootstrapping. See online appendix for this figure in color.

2.5.3 Rate and Location of Change

Given the differential gains in dietary quality noted above, we now investigate *when* in time and *where* in the distribution of dietary quality these improvements took place. For consistency and cross-sample/population comparisons, we focus on fixed portions of the distribution of dietary quality. An obvious choice is to use quartiles, which are all roughly segmented by HEI scores of 40, 50, 60.[14] Table 2.5 measures the amount of dietary improvement occurring in a particular quartile between two time periods as the percentage of total improvement ($\hat{F}_{1989-91}$ - $\hat{F}_{2005-08}$). That is, we measure the area bounded by the two empirical CDFs within each quartile range of the HEI scores. For example, the percentage of improvement in the U.S. over the 20-year period that occurred in the bottom quartile (< 40) between 1989-91 and 1994-96 was 2.8 percent.

[14]Quartile estimates for the U.S., low-, and higher-income populations when samples are pooled across the 20-year period reveal cutoffs of (40.4, 50.7, 61.6), (39.0, 48.8, 60.0), and (40.9, 51.4, 62.2), respectively.

The last column of table 2.5 measures the overall improvements over 1989–2008 within each quartile of the distribution of dietary quality.

Table 2.5: Location and Time-path of Dietary Improvement

	Between period			Total
HEI range	$\hat{F}_{89} - \hat{F}_{94}$	$\hat{F}_{94} - \hat{F}_{01}$	$\hat{F}_{01} - \hat{F}_{08}$	$\hat{F}_{89} - \hat{F}_{08}$
All adults				
0 - 40	2.80	12.61	6.21	21.64
40 - 50	5.65	6.83	9.54	22.01
50 - 60	12.41	1.86	10.25	24.52
60 - 100	20.38	-3.64	15.10	31.83
0 - 100	41.23	17.66	41.09	100.00
Low-income				
0 - 40	2.70	8.49	8.33	19.41
40 - 50	-1.91	6.99	14.73	19.85
50 - 60	4.40	0.81	20.57	25.80
60 - 100	11.35	-4.18	27.84	34.94
0 - 100	16.54	12.11	71.47	100.00
Higher-income				
0 - 40	2.98	15.14	4.49	22.70
40 - 50	8.58	8.46	6.04	23.07
50 - 60	15.24	4.36	4.40	24.00
60 - 100	23.27	-0.99	7.92	30.23
0 - 100	50.07	26.97	22.85	100.00

Note: Numbers represent the percentage of the 20-year improvement coming from the area bounded by the HEI range and the two distributions. $\hat{F}_{89} = \hat{F}_{1989-91}$, $\hat{F}_{94} = \hat{F}_{1994-96}$, $\hat{F}_{01} = \hat{F}_{2001-04}$ and $\hat{F}_{08} = \hat{F}_{2005-08}$.

For the U.S. adult population, improvements below the median (HEI < 50) occurred steadily over the period 1989–2008. Individuals in the upper range of dietary quality (HEI above 50) experienced virtually all of their gains over the periods 1989–1996 and 2001–2008. Overall, there were slightly higher gains in the upper quartiles compared to the lower quartiles for the U.S. population.

Comparing the between period improvements by income group, we see that 71.5% of the total improvement in the diets of the low-income population occurred more recently over 2001–08. This is in contrast to the higher-income population which saw the majority of their improvements occurring over 1989–2001 (77.1%). Improvements in the lower quartiles for the higher-income population have been relatively steady over the 20-year period, whereas most of the improvement in low-income diets within the lower quartiles occurred more recently over 1994-2008. In other words, at the lower end of the distribution of dietary quality, low-income individuals have seen comparatively limited or lagging improvements.

Table 2.5 emphasizes the reasons for targeting the most vulnerable group at risk of poor diets – the low-income, low dietary quality population. This is best seen by examining the last column of table 2.5, which measures the total gains over the 20-year period within each quartile. The higher-income population has had almost proportional gains across all levels of HEI, whereas the low-income population has seen less improvement in the lower quartiles of diet quality.

2.6 Counterfactual Analysis

We now explore whether factors that evolve gradually over time within the population can help explain observed improvements in the distribution of HEI scores between 1989–2008. We focus on two factors in particular: changes in food formulation and changes in the demographic landscape.[15] In the figures below, we focus on the differences between the observed 2005-08 distribution and the 1989-91 counterfactual distributions.

2.6.1 Food Reformulation

The composition of the food supply has changed considerably over the last twenty years in response to changes in policy, regulation, technology and consumer tastes. For example, Vesper et al. (2012) find that levels of transfats in the population declined after new labeling requirements were put in place in 2003. We now investigate how much of the improvement in dietary quality can be attributed to changes in food composition.

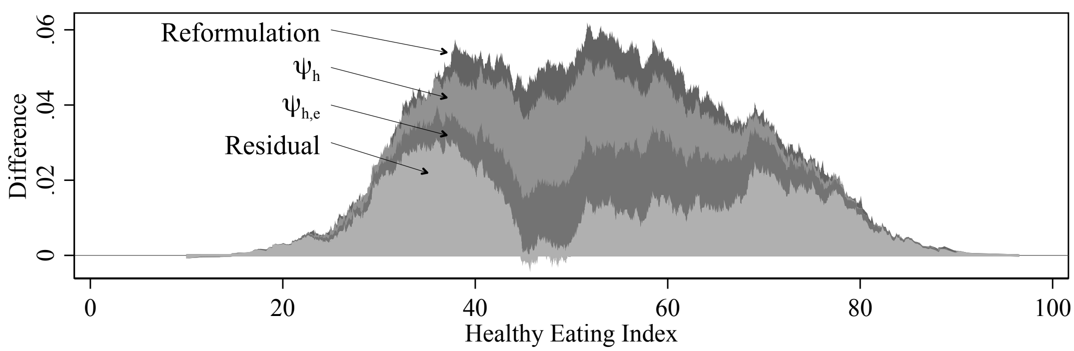
In order to identify foods and food mixtures that have undergone food reformulation (e.g., changes in the type of fat used in processed foods), we use the USDA Food and Nutrient Database for Dietary Studies (FNDDS). FNDDS consists of a series of databases

[15] Educational attainment is missing for 121 individuals in 1989-91 CSFII (61 low-income and 60 higher-income) and 1 higher-income person in 2005-08. These individuals are dropped from all counterfactual analyses. The preceding analysis is robust to their exclusion.

updated every two years in conjunction with the continuous waves of NHANES to reflect the current state of food formulation and packaging. All together we combine the FNDDS to cover 1994–2008. We briefly explain the method here with more details in the supplementary appendix online.

To construct the distribution of dietary quality in 1989–1991 as if food were formulated in 2005–2008, we first identify all foods coded as reformulated in the 1994–2008 FNDDS. We then replace the nutrient values for these food items in the 1989–1991 CSFII with the reformulated values found in the FDNNS. We also replace the MPED values of the 1989–91 reformulated foods with their 2005–08 values. We then construct a new HEI-2005 score based on updated nutrient and MPED values for each respondent in the 1989–1991 sample. Figure 2.5 displays the results from the reformulation counterfactual, as well as results from the next section.

Figure 2.5: Differences between the 1989-91 counterfactual distributions and observed 2005-08 distribution



Note: "Reformulation" is defined in the subsection *Food Reformulation.* ψ_h is a reweighting function (see subsection *Demographic Changes*) which includes dummies for female and race/ethnicity fully interacted with age groups, and $\psi_{h,e}$ additionally includes education dummies, all as defined in table 2.6.

The distribution of HEI that accounts for reformulation lies everywhere to the right of the original 1989–1991 distribution over the relevant range of the HEI. The implication is that, holding food choices constant, changes in food composition could be a contributing factor to dietary improvement. In figure 2.5, the indicated shaded area represents the change in the empirical CDFs attributed to reformulation. The ratio of this area to the total area provides a scalar measure of change. Here, improvements attributed to reformulation represent about 10.1% of the total difference between 1989–91 and 2005–08.

An important caveat is that this exercise captures partial equilibrium effects and some care must be taken interpreting these results. Our counterfactual analysis cannot account for the fact that individuals in 1989–1991 might have chosen different foods had their foods been formulated as they were in 2005–2008. Nevertheless, it shows how food reformulation, all else equal, can play an important role in changing dietary quality.

2.6.2 Demographic Changes

The United States of 2005–2008 is an older, more diverse, and better educated country than the United States of 1989–1991. To the extent that these factors are correlated with healthy eating, they may explain some of the improvements in dietary quality. Table 2.6 illustrates demographic changes using data from our sample and from the U.S. Census. There is a clear decrease in the 30 to 44 year old population and a concomitant rise in the 45 to 64 year cohort. The decrease in the non-hispanic white population has come from an increase in the Hispanic and other race/ethnicity groups. Finally, the overall educational attainment in the population has also increased.

To investigate the effect of evolving population characteristics, we construct counterfactual distributions of HEI scores following an approach proposed by DiNardo, Fortin, and Lemieux (DFL 1996). We ask, "What would the distribution of HEI scores look like had the demographic landscape of 2005–08 prevailed in 1989–91?" We focus on age, race/ethnicity, and educational attainment, all of which have been found to be correlated with diet healthfulness (Popkin, Siega-Riz and Haines, 1996). The intuition is to adjust each individual's sampling weight in the base period 1989-91 conditional on a set of demographics such that it captures the relative probability that the individual would be represented in the more recent 2005–08 sample.

To briefly describe the DFL methodology, let each individual observation be a vector (y, h, t), where y is HEI, h is vector of demographic characteristics, and t is time. Thus, all individuals belong to the joint distribution $F(y, h, t)$. The static joint distribution of HEI and demographics in time t is $F(y, h|t)$. The density of HEI at any point in time $f_t(y)$ can be written as the integral of the HEI density conditional on a set of

Table 2.6: U.S. Population Characteristics, Adults 20 and Older

Demographic	1989-91 CSFII	2005-08 NHANES	1990 Census[a]	2005-07 Census[b]
Age 20 – 29	21.7	19.4	22.7	19.1
Age 30 – 44	35.9	28.3	33.5	29.1
Age 45 – 64	26.2	35.4	26.1	35.0
Age 65+	16.3	16.8	17.6	16.8
Non-Hispanic white	78.8	71.9	78.4	69.5
Non-Hispanic black	10.8	11.3	10.6	11.3
Hispanic	7.7	11.6	7.6	12.8
Other race/ethnicity	2.7	5.2	3.4	6.4
Did not attend high school	8.5	6.0	9.6	6.1
High school, no college	46.2	37.8	44.5	39.7
Attended college	45.2	56.2	45.9	54.2
N	9,377	9,257		

[a]U.S. Census Bureau, General Population Characteristics (CP-1, 3-4).

[b]U.S. Census Bureau, 2005-07 Annual Community Survey 3-year sample.

demographics $f(y|h, t_y)$ at a specific date t_y, over the distribution of demographics $F(h|t_h)$ at date t_h

$$\begin{aligned} f_t(y) &= \int_{h\in\Omega_h} dF(y, h|t_{y,h} = t) \\ &= \int_{h\in\Omega_h} f(y|h, t_y = t)dF(h|t_h = t) \\ &= f(y; t_y = t, t_h = t) \end{aligned}$$

where Ω_h is the domain of individual demographics. Therefore, our question posed earlier can be written with the above notation as the density of HEI scores in 1989-91 had the 2005-08 demographic landscape prevailed: $f(y; t_y = 89, t_h = 08)$. This density is written as

$$\begin{aligned} f(y; t_y = 89, t_h = 08) &= \int f(y|h, t_y = 89)dF(h|t_h = 08) \\ &= \int f(y|h, t_y = 89)\psi(h)dF(h|t_h = 89) \end{aligned}$$

where $\psi(h)$ is a reweighting function defined as $\psi(h) = dF(h|t_h = 08)/dF(h|t_h = 89)$. Applying Bayes' rule to the function we can rewrite $\psi(h)$ as

$$\psi(h) = \frac{\Pr(t_h = 08|h)}{\Pr(t_h = 89|h)} \cdot \frac{\Pr(t_h = 89)}{\Pr(t_h = 08)}.$$

To obtain an estimate $\hat{\psi}(h)$, notice the conditional probabilities $\Pr(t_h = t|h)$ can be estimated using a probit model by pooling the data and estimating the probability an individual is observed in time t conditional on a set of characteristics. As we only compare two dates, the unconditional probabilities $\Pr(t_h = t)$ are simply the weighted sums of individuals in period t_h over the weighted sums of individuals in both periods. Because we are interested in applying the above methodology to tests of stochastic dominance, we replace an individual's sampling weight θ_i with $\omega_i = \theta_i \hat{\psi}_i(h)$ in equation (2.2).

While long run demographic changes such as gender, age and race/ethnicity are plausibly exogenous, the claim that education is uncorrelated with omitted factors that affect diet quality is less plausible. However, we are interested in how *changes* in the distribution of education affects *changes* in diet quality, rather than how education affects diet quality. In other words, the conditional independence assumption $E[\varepsilon|h] = 0$ is unnecessary for our decompositional analysis. Rather, we only need the weaker assumption of ignorability (also called unconfoundedness or selection on observables) to compute the aggregate compositional effects of all demographics. Ignorability asserts that the correlation between education (or any variable in h) and the error term is the same in both periods.[16]

Due to the aggregate decompositional nature of DFL (as opposed to a Oaxaca-style decomposition), the reweighting function $\psi(h)$ does not distinguish between individual variables in the vector h. In the interests of transparency, we construct the counterfactual distributions in two stages: First, we construct a counterfactual distribution accounting for purely demographic changes (gender, age and race/ethnicity) and denote this reweighting function by ψ_h. We then construct a counterfactual distribution

[16]If we believe this assumption does not hold, then we can sign the bias. For example, if we believe that more highly educated individuals use their stock of knowledge more efficiently in 2005-08 than in 1989-91, then we have a positive bias. However, there is no *a priori* evidence to suggest a change in the correlation of education and the error term, let alone as to its direction.

accounting for changes in demographics and changes in education levels, denoted by $\psi_{h,e}$.[17] We investigate the effects of the ordering in section 2.6.4.

Figure 2.5 decomposes the change in the distribution of HEI into four main parts: improvements attributed to reformulation (as shown in the previous section), additional improvements attributed to changes in demographics – with and without education – and finally the residual change. As noted above, 10.1% of total improvement can be attributed to changes in food composition. Here we find that roughly equal proportions of the total improvement in HEI scores can be attributed to changes in gender, age and race/ethnicity (26.6%) and education (26.7%) over the twenty year period.[18] This leaves 36.6% of the improvement unexplained by reformulation and demographics (i.e., the residual improvement). The residual improvement encompasses many competing factors such as changes in tastes, relative food prices, scientific discovery, and attitudes towards food in general.

As above, care must be taken in interpreting these results. One important limitation of the partial equilibrium nature of the counterfactual analysis is that food choices in the counterfactual population would not affect the set of foods made available by food manufacturers. While this assumption is economically unappealing, the exercise provides insight into the effects of changing demographics on diet quality via clear and tractable analytical techniques.

2.6.3 Counterfactuals by Income Group

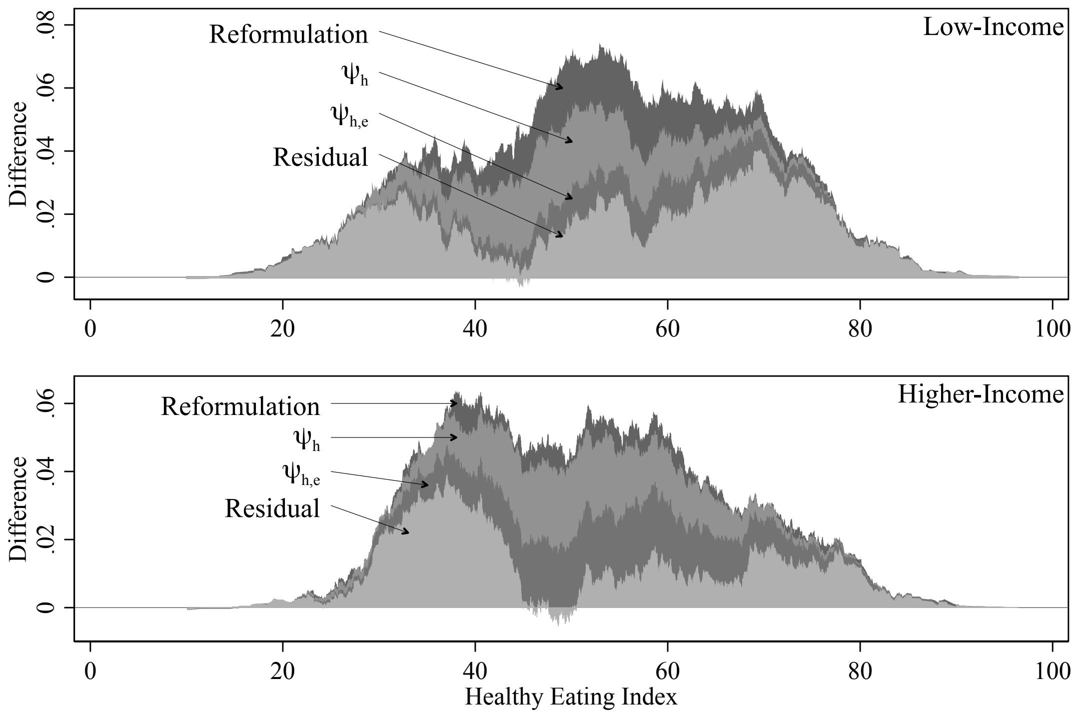
The counterfactual analyses above suggest that an important part of the improvement in dietary quality can be attributed to changes in food composition and demographics. Given that improvements occurred at different rates for different parts of the HEI distribution for lower-income versus higher-income individuals, we now ask whether changes in food composition and demographics account for differing amounts of improvement by income group. Results are presented in figure 2.6.

Changes in food composition account for a substantially larger percentage of the

[17] The conditional probability model includes a dummy for gender, 16 cells of race/ethnicity fully interacted with age dummies, and three education dummies, all as described in table 2.6. Results of the model available available in the supplementary appendix online.

[18] See the supplementary appendix online for dominance results between each counterfactual distribution and the observed 2005-08 distribution.

Figure 2.6: Differences between the 1989-91 counterfactual distributions and observed 2005-08 distribution by income group



Note: "Reformulation" is defined in the subsection *Food Reformulation*. ψ_h is a reweighting function (see subsection *Demographic Changes*) which includes dummies for female and race/ethnicity fully interacted with age groups, and $\psi_{h,e}$ additionally includes education dummies, all as defined in table 2.6.

dietary improvement for lower-income individuals (19.6%) as compared to their higher-income counterparts (6.4%). This is consistent with the observation that low-income individuals eat more processed foods (Drewnowski and Barratt-Fornell, 2004), where much of the reformulation is occurring. Changes in gender, age and race/ethnicity account for a similar share of the improvement for low-income (25.3%) than higher-income individuals (26.8%). For low-income individuals, changes in educational attainment account for half that of higher-income (13.5% versus 27.0%). The remaining residual share of the twenty-year improvement is larger within the low-income population (41.6%) as compared to higher-income (39.8%). This suggests that further research into the determinants of diet quality of low-income individuals may be warranted.

2.6.4 Robustness

The order in which we construct counterfactual distributions using the DFL approach can influence the results. To investigate the robustness of our findings to ordering, we estimate the model using an alternative ordering for each of the three population groups of interest (total population, low-income, higher-income). Note that because reformulation is not estimated, but rather derived from data, it does not matter which order it is considered. Furthermore, the total aggregate effect ($\psi_{h,e}$) remains the same as well. For example, in either case all demographics account for 53.3% of the total improvement within the U.S. population.

Table 2.7 provides estimates for the original order as presented above, as well as an alternative ordering where we first consider educational attainment ψ_e and then use $\psi_{h,e}$ as before. The result places bounds on the magnitude for each set of demographics. For example, the effect of education ranges between 15.6 and 26.7% for the total population, 5.0 and 13.5% for the low-income group and 16.0 and 27.0% for the higher income group. Although point estimates change, relative comparisons remain substantively the same – changes in education appear to account for a larger share of the improvement for the higher-income group relative to the lower-income group. We note that the bounds are relatively large and credibly point-identifying each effect remains a task for future work.

Table 2.7: DiNardo, Fortin and Lemieux Counterfactual Improvements

Order:	U.S. Population	Low-income	Higher-income
Original Order			
1. Reformulation	10.1	19.6	6.4
2. Gender, Age, Race/Ethnicity	26.6	25.3	26.8
3. Education	26.7	13.5	27.0
Total: Reformulation & Demos	63.4	58.4	60.2
Alternative Order			
1. Reformulation	10.1	19.6	6.4
2. Education	15.6	5.0	16.0
3. Gender, Age, Race/Ethnicity	37.7	33.7	37.7
Total: Reformulation & Demos	63.4	58.4	60.2

Note: Numbers represent the percentage of total improvement and may not sum accordingly due to rounding.

2.7 Discussion and Conclusion

Conventional wisdom maintains that the quality of the American diet has been deteriorating for at least the past two decades.[19] In contrast, we document a previously unknown pattern of improvement in U.S. dietary quality. We find statistically significant improvements for all adults over the period 1989–2008, at all levels of dietary quality.

An important caveat is that the HEI measures diet quality on a per calorie basis and does not account for *excess* calorie consumption. To our the best of our knowledge, few studies have examined the quantity-quality isoquant of food in health production, and those that have generally do so within the context of specific foods, in an experimental framework. In a series of dietary intervention experiments, Epstein et al. (2001, 2008) found that increasing healthy food consumption reduced obesity to a greater degree than reducing unhealthy food consumption. Moreover, in Epstein et al. (2008) individuals in the increase-healthy-food group showed no relapse in weight gain in a two year follow up. The implication is that a shift towards a healthier diet could have additional positive impacts on health outcomes driven by quantity, such as obesity. The mechanism is generally thought to be a higher level of satiation, which in turn leads to a reduction in overall calories consumed.

While we find that higher-income individuals consistently have higher dietary quality than low-income individuals, we also find some evidence that the gap is shrinking over the sample period. An important caution is that the diets of low-income individuals in the lowest portion of the diet quality distribution continue to lag.

We also show that most of the improvement in dietary quality can be attributed to changes in food formulation and changes in demographics. Moreover, we find that changes in food formulation help explain considerably more of the improvement in dietary quality for low-income individuals than for higher-income individuals. These findings suggest that the direct and indirect effects of policy on food composition may represent understudied policy levers. How large are these results? In a prospective study that roughly covers our sample period, Chiuve et al. (2012) found significantly lower risks of major chronic diseases across the entire distribution of HEI-2005 scores for

[19] See for example Gregory, Smith and Wendt (2011).

both women (over 1984-2008) and men (1986-2008) who were free of chronic disease at baseline. For example, those in the second quintile were 7 percent less likely of chronic disease incidence than those in the lowest quintile, all else equal. One way to asses the magnitude of changes in HEI over time is to see how many individuals move from low to moderate levels of dietary quality over the period under study. In 1989-91, the 20$^{\text{th}}$ percentile of the HEI distribution was 37.3. In 2005-08, a HEI value of 37.3 represented the 15.4$^{\text{th}}$ percentile of the HEI distribution. In other words 4.6 percent of individuals moved out of this higher-risk category between 1989–2008 due to improvements in diet quality.

Findings of a small but statistically significant increase in dietary quality should not overshadow the fact that there is still considerable room for improvement. Moreover, an important residual share of the change in dietary quality over the period remains unexplained, especially in the tails of the distributions. Because of the sheer number of overlapping and time varying policy initiatives – particularly those that target the poor – credibly identifying effects of specific policies remains a challenging task for future work.

Chapter 3

"Billions and Billions Served" Heterogeneous Effects of School Food and Away Food on Child Dietary Quality

3.1 Introduction

Early decisions in human capital accumulation and skill formation have direct consequences on the productivity of future investments (Cunha et al., 2006). Skills related to health capital, for example, quickly accumulate early on in life and have persistent impacts throughout adolescence and adulthood (McFadden, 2008). Therefore, it is of no surprise that the case for investing early in children, specifically the disadvantaged, is strong (Heckman and Masterov, 2007), and policymakers are particularly interested in programs that target such children. With nutrition in mind, two longstanding Federal programs have gained increasing attention in the United States: the School Breakfast Program (SBP) and the National School Lunch Program (NSLP).[1]

Offered in over 100,000 public and non-profit institutions, the SBP and NSLP serve millions of students every school day.[2] Together, these two Federally subsidized meal programs represent a substantial and repeated exposure to nutrition skill formation, which has strong implications for nutrition capital accumulation. For example, numerous experimental trials have demonstrated that infants and young children have the capability to learn and apply nutrition skills, but the ability to adopt new skills decreases as one matures into adulthood.[3] Outside of school and home, exposure to food-away-from-home (FAFH), such as fast-food and restaurant establishments, has become much more prominent in the daily diet of American children (Poti and Popkin, 2011). While the literature generally agrees that FAFH negatively impacts health, researchers are at odds with respect to the impact of school food. The findings of this paper suggest that the conflicting results may be due to a focus on the average effect of consuming school food, which may mask important effects in the tails of the outcome distribution.

This study adds to the current literature by considering heterogeneous effects of food source across all levels of underlying dietary quality, rather than focusing on average diet quality. I focus on dietary quality because it correlates with body weight (Jennings

[1]Since its inception in 1946, the NSLP has served over 224 billion lunches in the U.S. (FNS-USDA, 2012a). Interestingly, it is estimated that McDonald's has sold over 247 billion hamburgers since its re-opening under its namesake in 1948.

[2]The NSLP is offered in 99% of all public schools and 94% of public and private schools combined (Ralston et al., 2008). Nearly 32 million lunches were served daily in 2011, with roughly two-thirds at a free or reduced-price (FNS-USDA, 2012a). The SBP served 12.1 million students in 2011, with 10.1 million receiving a free or reduced-priced breakfast (FNS-USDA, 2012b).

[3]Benton (2004) and Birch (1999) provide thorough reviews of such studies.

et al., 2011) and academic achievement (Florence, Asbridge and Veugelers, 2008) in children, is a predictor for many chronic diseases in adulthood (Chiuve et al., 2013) and is at the forefront of Federal and State policies aimed at reforming nutritional standards in schools.

Three food sources are considered: food from home (FFH), from away from home (FAFH) and food from school (FFS). I define underlying dietary quality as a child's "proneness" to consume a healthful diet.[4] For example, a child who is prone to a very low quality diet, possibly due to parental or environmental factors, may exhibit large benefits from a school lunch and/or breakfast. On the other hand, a child prone to a high quality diet who consumes a meal from school may experience a decrease in overall dietary quality. Therefore, examining the the average treatment effect (ATE) of participating in school food programs may mask important heterogeneous effects. This study expands on existing literature by estimating the quantile treatment effect (QTE) of food source on the distribution of child dietary quality.

Several studies have investigated the mean effects of food source on various aspects of child health. Many have found that FAFH increases calorie intake (Bowman et al., 2004; Powell and Nguyen, 2013) and reduces diet quality among children (Mancino et al., 2010). When examining the effect of participating in the SBP and the NSLP separately, several authors are in agreement that the former is beneficial (Bhattacharya, Currie and Haider, 2006; Millmet and Tchernis, 2012; Millimet, Tchernis and Husain, 2010; Schanzenbach, 2009) but the latter is not (Schanzenbach, 2009; Millimet et al., 2010). For example, Schanzenbach (2009) found that school lunches increased average daily calorie intake by about 40 kilocalories and that children who consume school lunches have higher rates of obesity by 1 to 2%. Overall, while the consensus appears to be that child health is negatively impacted by FAFH, researchers are at odds when considering the average impact of a school breakfast or lunch.

Several approaches to identifying the effect of food source on dietary outcomes have been used. A fixed effect, or first-differencing approach, is easily implemented by using two days of dietary intake typically found in U.S. nationally representative data sets

[4]The term "proneness" was introduced by Doksum (1974). In the present study, proneness can be thought of as the (fixed) degree to which one consumes a healthful diet. Perhaps a more relatable example from the economics literature is ability, which is defined by how prone one is to more favorable labor market outcomes.

(e.g., Mancino et al., 2010; Powell and Nguyen, 2013). When examining more long-term outcomes, such as body weight, an individual fixed effect approach becomes more problematic (e.g., Schanzenbach, 2009). A second approach is to use instrumental variables (Hinrichs, 2010), which comes with limitations such as the exclusion restriction and access to a credible instrument.[5] Gundersen, Kreider and Pepper (2012) step back from identification and place bounds on the effect of participating in the free and reduced price lunch program. The bounding approach relies on the weaker *monotone instrumental variable* (MIV) assumption.[6] Bhattacharya, Currie and Haider (2006) use variation in the timing of the interview (i.e., if school is in session or not) coupled with SBP availability via difference-in-differences. Finally, Schanzenbach (2009) uses regression discontinuity, which has assumptions similar in spirit to the MIV assumption. A drawback of regression discontinuity is that the effects are estimated only for those near the income eligibility cutoff, again, possibly masking any heterogeneous effects.

All of the aforementioned studies examined average effects. A major limitation of this approach is that it is not very informative about effects in the tails of the outcome distribution. This paper uses a quantile regression technique to determine the effect of food source on dietary quality across the entire distribution. Where this study differs methodologically from those previously published is the identification procedure for examining distributional effects. Identification is accomplished by using within-person variation of dietary intake on two nonconsecutive days but maintains the nonseparable property of the disturbance term, also called the rank variable. In other words, rank is determined by "total proneness," which is function of both the individual fixed effect and random error. This advancement allows the coefficients of interest to be interpreted in the same manner as cross-sectional quantile estimation. Contrast this with location-shift quantile estimators that model the individual fixed effect as a separate additive term (Canay, 2011; Galvao Jr., 2011; Graham et al., 2009; Lamarche, 2010; Ponomareva, 2011). By separating the total disturbance, the coefficient of interest is now interpreted as the effect relative to one's on fixed effect, which is not very informative from a policy

[5]See Millmet and Tchernis, (2012) for an application to the SBP using alternative assumptions to estimate treatment effects without the exclusion restriction.

[6]The MIV assumption relaxes the exclusion restriction (see, Manski and Pepper, 2000). Gundersen et al. (2012) assume participation in the free or reduced price programs is monotonically associated with income to overcome selection into free or reduced-price programs. They find receipt of free or reduced-price lunches improves child health outcomes.

standpoint.

The paper proceeds as follows: the next section more formally defines dietary quality and introduces a widely used measurement, the Healthy Eating Index-2005, which forms the basis of the analysis. After a brief overview of the data, I use summary measures to motivate a more detailed analysis. I then discuss the identification and estimation strategy, followed by the main results. The final section discusses policy implications and conclusions.

3.2 Dietary Quality

The overall healthfulness of a child's diet can be distinguished by two factors: *energy balance* and *dietary quality.* Energy balance is relationship between calories consumed and calories expended, which results in body weight management (Hall et al., 2012). Dietary quality, on the other hand, represents the degree to which a child's diet is meeting a set of criteria, for example, eating the correct proportions of healthy foods while maintaining moderation in less-healthy foods. It is important to note that energy balance and dietary quality are interconnected: experimental studies have shown when children switch to higher-quality diets, as opposed to calorie-restriction diets, sustained weight control is observed (Epstein et al. 2001, 2008).

I quantify dietary quality using the Healthy Eating Index (HEI). The HEI was developed in 1995 to measure compliance to the U.S. Government's official recommendations for healthful eating, the *Dietary Guidelines for Americans* (DGA). Every five years, based on an expert advisory panel, the DGA are revised by the U.S. Departments of Agriculture (USDA) and Health and Human Services (HHS). As such, the HEI has been updated several times to reflect the most current state of nutrition knowledge. This paper uses the HEI-2005 and will henceforth refer to the HEI-2005 as simply HEI.[7]

The HEI is the sum of 12 components based on the consumption of various foods or nutrients. Each component assigns a score ranging from 0 to 5 (total fruit, whole fruit, total vegetables, dark green/orange vegetables and legumes, total grains, whole grains), 0 to 10 (milk, meats and beans, oils, saturated fat, sodium) or 0 to 20 for the percentage

[7]Future work will use the HEI-2010. Note that the two have many similarities (see, National Cancer Institute, 2013).

of calories from solid fats, alcoholic beverages, and added sugars (SoFAAS) creating a maximum score of 100. Appendix table C.1 provides exact details of the scoring (see also, Guenther et al., 2008a).

The HEI has been widely used and evaluated as a valid measure of diet quality (Guenther et al., 2008b). In the medical literature, lower HEI scores are associated with higher risks of coronary heart disease, stroke and diabetes (Chuive et al., 2012), cardiovascular disease (Nicklas et al., 2012), breast cancer (Shahril, 2012), colorectal cancer (Reedy et al., 2008) and prostate cancer (Bosire et al., 2013). Economists have used the HEI as an indicator of well-being to analyze distributional trends (Beatty, Lin and Smith, 2014) and to study the impacts of the Supplemental Nutritional Assistance Program (Gregory et al., 2013). It is important to reiterate that the HEI is a per-calorie measure of dietary quality and does not directly consider excessive calorie intake. Although at first glance this distinction may seem limiting, it is important and necessary to analyze the relative *quality* of foods consumed across various food sources.[8]

3.3 Data

I use data from three waves of the National Health and Nutrition Examination Survey (NHANES) covering 2003-08. Each survey wave is an independently drawn sample, which is representative of the U.S. with the USDA overseeing the food intake component. The NHANES provides rich information on dietary intakes so that HEI scores can be calculated according to Guenther et al. (2008a) (see also, the Appendix table C.1). Each wave was conducted from November in the odd year to October in the even year. For the 2003-08 NHANES, respondents report 24-hour dietary intakes on two nonconsecutive days (Day 1 and Day 2). Day 1 intakes are administered in-person during the medical exam, and Day 2 intakes are conducted 3–10 days later in a follow-up telephone interview. All interviews are conducted by trained dietary interviewers with the aid of three-dimensional measuring instruments.

A primary goal of this research is to understand how school food affects dietary

[8]It is also worth noting that analyzing energy balance is problematic for several reasons: (a) quantifying energy expenditure is difficult (Crouter, Clowers, and Bassett, 2005); (b) calorie needs vary substantially for boys and girls of different ages, making distributional comparisons within the population difficult; (c) most importantly, calorie consumption is not monotonic; one would need to make an assumption about the asymmetric relationship between under- and over-calorie consumption.

quality. As such, I focus on school-aged children (4–19) that report attending kindergarten through high school during the school year and have complete dietary intakes on both days ($n = 7,009$). Thus, children that have dropped out of school or graduated are excluded. I also exclude those attending schools that do not offer a lunch ($n = 379$), as done elsewhere in the literature (Gleason and Suitor, 2003; Gunderson et al., 2012; Millmet et al., 2000; Schanzenbach, 2009). The final sample with complete information consists of $6,630$ children.

Broadly defined, food from home (FFH) are items bought at the grocery store, food from school (FFS) are meals received at school, and food away from home (FAFH) primarily consists of fast-food and full-service restaurant items. Also included in FAFH are items bought in vending machines, received as a gift, and street food. Thus, for example, a candy bar purchased from a vending machine at school is considered FAFH, not FFS. Appendix B.2 contains complete details for mapping the 25 original food source codes into one of the three categories.

3.4 Summary Measures

In specifying regression models, I will use home-prepared food (FFH) as the reference category. This will give FFH a control interpretation, which is reasonable since children in the U.S. eat at home nearly every day (table 3.1). FFS and FAFH will be considered the policy variables of interest (i.e., treatments). Food served in schools and at away-from-home venues are sources of political debate and subject to policy interventions.[9] Variation in FFS and FAFH is considerable as evidenced by table 3.1. Roughly 40% of children ate at school on at least one day in 2003-08, and over three-quarters of children reported FAFH consumption on Day 1, Day 2 or both.

By design, and as noted earlier, HEI scores are bounded between 0 and 100. In table 3.2 we can see that no child is observed on the bounds. In fact, dietary scores for both Day 1 and Day 2 are approximately normally distributed with a central tendency

[9]For example, the NSLP has undergone numerous regulatory changes with respect to minimal nutritional values (see Ralston et al., 2008). The most recent change to nutritional standards for the SBP and NSLP begin to take affect in the 2013 school year under the Healthy Hunger-Free Kids Act of 2010 (Code of Federal Regulations, Title 7, Parts 210 and 220, 2013). In fast-food and restaurant establishments, nutritional facts are now mandatory. Moreover, some states, such as New York, are currently seeking to further regulate away-from-home venues.

Table 3.1: Variation in consumption by food source

	Percentage consuming on...		
Food Source	Neither day	One day only	Both days
Home	0.19	2.20	97.61
School	58.72	29.41	11.88
Away	21.65	43.09	35.27

Source: Children aged 4-19 reporting two complete days of intake in the the 2003-08 National Health and Nutrition Examination Survey (NHANES). The sample includes children that report attending kindergarten through high school during the school year and attend schools that offer a lunch. All calculations use survey weights.

of about 50 (see kernel densities in appendix figure B.1). Participation in the away-from-home market is noticeably lower on Day 2, which is most likely correlated with the observed shift in intake records towards the beginning of the week (table 3.2).[10] The allocation of calories across the three food sources varies substantially, is skewed towards 100% for home food and skewed towards 0% for school and away food. By conditioning on participation, we can get a a more complete picture of calorie consumption when children consume FFS and FAFH. In particular, children consume about one-third of their calories at school and about 40% of their calories away form home conditional on participation.

3.4.1 Mean Measures

Table 3.1 suggests using individual variation to estimate the impact of food served in schools and away from home on dietary quality. By including individual fixed effects and assuming conditional exogeneity, unobservable fixed characteristics associated with selection into the SBP and/or NSLP, as well as food-away-from-home market, are no

[10]Note, however, that the NHANES provide specific two-day dietary sample weights that account for nonresponse of dietary intake on Day 1 (about 6% of the total sample) and for additional nonresponse on Day 2 (an additional 10%). The weights also account for the differential allocation by day of the week and for the proportion of weekend-weekday combinations of Day 1 and Day 2 recalls. Less than 2% of the sample used in this study were surveyed on the same day of the week.

Table 3.2: Summary statistics

Variable	Day 1	Day 2
Outcome: HEI-2005		
Mean (st. dev.)	49.88 (16.22)	51.62 (16.81)
[min, max]	[15.11, 92.24]	[11.62, 92.89]
Food source participation (%)		
Home	98.62	98.79
School	25.53	27.64
Away	64.05	49.58
Allocation of calories (% of total)[a]		
Home	65.49 (29.21)	72.06 (27.95)
School	8.65 (17.11)	9.16 (17.16)
Away	25.86 (27.71)	18.79 (25.35)
Allocation of calories (% of total), conditional on participation		
Home	66.40 (28.30)	72.94 (26.92)
School	33.89 (17.13)	33.13 (19.47)
Away	40.38 (24.96)	37.90 (22.96)
Day of the week (%)		
Sunday	13.56	27.60
Monday	14.69	21.11
Tuesday	12.93	16.28
Wednesday	14.11	15.89
Thursday	13.84	5.08
Friday	14.36	10.08
Saturday	16.51	3.95
Number of children	6,630	6,630

Source: Children aged 4-19 reporting two complete days of intake in the the 2003-08 National Health and Nutrition Examination Survey (NHANES). The sample includes children that report attending kindergarten through high school during the school year and attend schools that offer a lunch. All calculations use survey weights.

[a]Calculated as 100 × (calories from source k on day t)/(total calorie intake on day t).

longer confounding. This suggests a general OLS specification such as

$$HEI_{it} = \alpha_i + D_{it}'\beta + X_{it}'\delta + \varepsilon_{it} \tag{3.1}$$

where HEI_{it} is the measure of diet quality for child i on day t and α_i are individual fixed effects. The term D_{it} represents the two policy variables of interest – the share of daily calories consumed at school and away from home. Time-varying controls, such as dummies for the interview day and each day of the week, are included in X_{it}. Finally, ε_{it} is an additive disturbance term. The underlying assumption is that changes in D_{it} (and X_{it}) are uncorrelated with changes in ε_{it}, so that β is consistently estimated. This assumption seems reasonable in the current context given that the second survey day is administered randomly 3–10 days after the first survey day.[11]

Some care must be taken in defining the policy variable. One approach is to simply define D_{it} as a dummy variable that equals one if child i consumed any food from that particular food source on day t. There are several limitations to this approach. For example, a child consuming two meals away from home would be categorized in the same manner as a child consuming one FAFH meal. Moreover, many children multi-source meals (e.g., consume some FFH and FFS in the same meal). Using lunch as an example, I find that 22% of lunches are multi-sourced on Day 1. Finally, the nutrient density (i.e., nutrient per calorie) varies widely depending on where the food was sourced (Lin and Guthrie, 2012) and should be considered when defining the policy variables.

I consider an alternative definition for the policy variables: the proportion of daily calorie intake from each food source. Under this definition, I am capturing the extent to which a child is "exposed" to each food source. Specifically, estimates using this definition capture the effect of substituting some share of calories from one food source to another. In short, defining D_{it} as the share of calories from school and away food will be my preferred definition, but I report both for comparison.

Column (1) of table 3.3 reports results using equation (3.1) with individual fixed effects only. Panels A and B report results using the two alternative policy variable definitions "dummy" and "share," respectively. Under the dummy variable definition in panel A, a child that consumes any FFS exhibits an average increase of 1.23 points on the HEI scale as compared to a child that does not eat at school. Using the average

[11] Changes in X_{it} are uncorrelated with changes in ε_{it} by survey design. Changes in D_{it} are orthogonal to changes in the disturbance term to the extent that changes in any time-varying omitted factors are uncorrelated with changes in the allocation of calories across the three food sources. More concretely, since food preferences are assumed to be fixed over the ten-day survey period, we can assume a relatively small omitted variable problem after conditioning on the individual fixed effect.

of Day 1 and 2 HEI scores in table 4.1, this equates to an average increase in dietary quality by 2.4%.

Table 3.3: Mean regressions with individual fixed effects

	(1)	(2)	(3)	(4)
Panel A: Dummy				
School food	1.228**	1.185**	0.152	0.099
	(0.448)	(0.434)	(0.583)	(0.567)
Away food	-2.521***	-2.047***	-2.111***	-1.811**
	(0.536)	(0.570)	(0.593)	(0.600)
Panel B: Shares				
School food	2.161	2.296	0.187	0.215
	(1.459)	(1.442)	(1.630)	(1.624)
Away food	-5.617***	-4.791***	-4.741***	-4.237**
	(1.192)	(1.228)	(1.296)	(1.290)
Fixed Effects				
Individual	Yes	Yes	Yes	Yes
Interview day	No	Yes	No	Yes
Day of the week	No	No	Yes	Yes
Observations	13,260	13,260	13,260	13,260
No. of children	6,630	6,630	6,630	6,630
R^2 (Panel A)	0.014	0.022	0.027	0.031
R^2 (Panel B)	0.018	0.026	0.030	0.034

Notes: The dependent variable is the HEI-2005. Standard errors are in parentheses and are calculated accounting for stratification and clustering. Panel A: $D^k = 1$ if food source k was consumed on day t and zero otherwise. Panel B: D^k = share of calories consumed from food source k on day t.

Using the share definition in panel B, it is useful to rescale the coefficients by average calorie allocation, conditional on participation (table 4.1). For example, since the average child consumes roughly one-third of their daily calorie intake at school, conditional on participation, the FFS estimate in column (1) of panel B equates to a 0.72 point increase in HEI (or 1.4%). When children frequent food-away-from-home establishments, an average of about 40% of daily caloric intake is consumed there. Thus, using estimates from panel B, one can infer an average decrease in HEI by 2.25 points (or 4.4%) when consuming FAFH. In comparing panel A and B, estimates using definition the share definition tend to be smaller.

Column (2) includes a dummy for the interveiw day (i.e., Day 1 or Day 2). This variable may be necessary due to differences dietary recall on Day 1 (in-person interview) versus Day 2 (telephone interview). The estimated coefficient on FFS changes very little, and the effect of FAFH is less.

Since NHANES surveys individuals on all days of the week, column (3) of table 3.3 includes an additional fixed effect for the day of the week. Perhaps unsurprisingly, estimates for FFS change dramatically and are no longer significant, most likely because children do not typically attend schools on the weekend. Conversely, FAFH is more likely to be consumed on the weekend, and again estimates are slightly smaller as compared to column (1).

Finally, column (4) reports coefficient estimates with individual fixed effects and dummies for the interview day and each day of the week fixed effects. This will be the preferred specification moving forward. In summary, there appears to a positive but insignificant average impact of school food on dietary quality and a robust, negative average impact of FAFH on dietary quality.

3.4.2 Distributional Measures

To motivate a distributional analysis, I present summary measures by selected quantiles. In table 3.4 I compare two-day average HEI scores for those that never select into FFS or FAFH (column 1 in table 3.1) and those that report consuming FFS or FAFH on at least one day of intake (columns 2 and 3 in table 3.1).[12]

We can see that the positive effect of FFS drops as we move across the distribution of HEI scores, implying much larger positive impacts in the bottom half of the distribution. The effect of FAFH has the smallest impact at the very bottom of the diet-quality distribution. This result is most likely due to the presumption that home-prepared food is more similar to FAFH at low levels of overall dietary quality. Beyond the bottom quartile, the effect of FAFH is relatively constant and thus more closely reflects the mean regressions.

[12]Dividing the sample in this manner is similar to the dummy variable definition used in panel A of table 3.3.

Table 3.4: Differences in the distribution of two-day average HEI-2005 scores by food source

	Quantile					
Food Source	5	25	50	75	95	N
Food from school						
At least one day	36.13	44.27	51.27	58.30	67.72	3,298
Neither day	33.45	42.80	50.25	57.27	68.07	3,332
Difference	2.68	1.47	1.02	1.04	-0.36	
Food away from home						
At least one day	33.94	42.76	50.02	56.82	66.34	5,163
Neither day	35.19	46.26	53.47	60.78	71.15	1,467
Difference	-1.25	-3.50	-3.45	-3.96	-4.81	

Note: A Kolmogorov-Smirnov type test of stochastic dominance (Barrett and Donald, 2003) indicates first-order dominance of FFS (at least one day) over FFS (neither day) with a simulated p-value of 0.041. Bennett's (2013) asymptotic two-stage test, however, indicates the FFS curves cross, leading to the conclusion of second-order dominance ($p = 0.015$). FAFH (neither day) first-order dominates FAFH (at least one day) with a simulated p-value of 0.000.

Of course, results from table 3.4 are confounded by both observable and unobservable individual characteristics. Moreover, given the results of table 3.3, we should also expect the day of the week to play an important role in identifying a more causal interpretation. In the next section, I use an estimator developed in Powell (2014) to estimate quantile treatment effects while controlling for individual fixed characteristics.

3.5 Quantile Estimation with Individual Fixed Effects

In this paper an important departure from the previous literature is how I estimate and identify the impact of food source on dietary quality. Given the summary results of the previous section, it is likely that the impact of food source is heterogenous across the distribution of diet quality. Much of this heterogeneity is most likely due to both observable and unobservable individual characteristics.

Typically, panel data can be used to control for individual heterogeneity via fixed effects. With mean regressions, an additive term α_i is included in the specification, and the estimated coefficients on the treatment (policy) vector D can be interpreted as the impact on both the conditional and unconditional mean (see Fripro). With quantile estimation, an additive fixed effect alters the interpretation of the coefficient of interest. The intuition behind this result is rather straightforward: the τ^{th} quantile of $Y_{it}|D_{it}, \alpha_i$ is in most cases not equal to the τ^{th} quantile of $Y_{it}|D_{it}$. For example, a high-quantile child in the distribution of $Y_{it}|D_{it}$ may become a low-quantile child after *conditioning* on his or her individual fixed effect.

The estimator used in this study accounts for individual heterogeneity without specifying or even estimating an individual "fixed-effect" parameter. Rather, unobservable individual heterogeneity is incorporated into the model by using within-person variation for identification but maintains the nonseparable property of the total disturbance $U^*_{it} = f(\alpha_i, U_{it})$. This allows the parameter vector $\beta(\tau)$ to be interpreted in the same manner as the τ^{th} "cross-sectional" quantile treatment effect. If we are interested in knowing how school and away affects low dietary quality children separately from high dietary quality children, this is precisely what we want to estimate.

3.5.1 Specification

Consider a cross-sectional quantile regression (QR) specification

$$Y_i = D_i'\beta(U_i^*), \quad U_i^* \sim U(0,1) \tag{3.2}$$

where for child i, Y_i is dietary quality, D_i is a vector of the proportion of calorie intake from each food source and U_i^* is "total proneness" to consume a healthy diet.[13] Total proneness is a function of his or her unobservable fixed proneness α_i (e.g., food preferences or environmental factors) and a disturbance term U_i (e.g., day-to-day randomness of food intake). In other words, U_i^* is a rank variable that incorporates heterogeneity into the model by allowing dietary quality to vary across children that have the same

[13]One may also desire to include a set of controls X_i, such as gender, age and race/ethnicity. The addition of "controls" alters the QTE interpretation of the estimates because some of U_i^* becomes observed through X_i. Put differently, if $D_i = (X_i, \tilde{D}_i)$ in equation (3.2) where $\tilde{D}_i$ are the treatments of interest, X_i would also be interpreted as a treatment vector, a distinction that is not necessary in mean regressions. Specifically, estimates in this specification would provide the QTE on the distribution of $Y_i|D_i, X_i$ rather than $Y_i|D_i$. See Powell (2013) for an in-depth discussion and estimation strategy of QTE in the presence of covariates.

observed allocation of calories. It is necessary to assume the relationship between proneness and the outcome to be (weakly) monotonic. That is, children with a higher U_i^* are more prone to a healthier diet for a given allocation of calories across the three food sources.

Clearly, no cross-sectional distinction can be made between α_i and U_i but it is informative to see that the impact of covariates vary according to the total level of proneness U_i^*. Moreover, an estimate of β using equation (3.2) assumes D_i is exogenous. If we believe individual-level fixed characteristics influence a child's allocation of calories across food source, then $U_i^*|D_i \nsim U(0,1)$.

To this end, it is useful to write down the Structural Quantile Function (SQF) introduced by Chernozhukov and Hansen (2005, 2008). The SQF of interest for identifying the QTE can be written as

$$S_y(\tau|d) = d'\beta(\tau), \quad \tau \in (0,1). \tag{3.3}$$

Equation (3.3) defines the τ^{th} quantile of the latent outcome $Y_d = d'\beta(U^*)$ for a fixed allocation of calories and a randomly selected $U^* \sim U(0,1)$.[14] This framework becomes important for describing the various "fixed-effect" quantile estimators and how they relate to the structural equation of interest.

Now consider a panel of students that report dietary intakes on multiple days. In this case, equation (3.2) can be rewritten as

$$Y_{it} = D_{it}'\beta(U_{it}^*), \quad U_{it}^* \sim U(0,1) \tag{3.4}$$

where $U_{it}^* = f(\alpha_i, U_{it})$. Again, α_i is the student's fixed level of proneness and U_{it} is an individual time-varying disturbance term. In this case, conditioning on individual fixed effects can overcome endogeneity concerns if U_{it} is uncorrelated with changes of D_{it}. However, including an additive fixed effect in quantile regression, as done in mean regressions, alters the interpretation of the coefficients. For example, consider the two

[14]Note that capital letters denote random variables (i.e., observed in the data) and lower case letters denote the potential outcome (i.e., the counterfactual outcome to be modeled).

specifications

$$Y_{it} = \alpha_i + D_{it}'\beta(U_{it}) \quad \text{and} \quad Y_{it} = \alpha_i(U_{it}) + D_{it}'\beta(U_{it}) \tag{3.5}$$

of Koenker (2004) and Harding and Lamarche (2009), respectively. The underlying SQF for these two location-shift type specifications found in (3.5) take the form

$$S_Y(\tau|d, \alpha_i) = \alpha_i + d'\tilde{\beta}(\tau), \quad \tau \in (0,1) \tag{3.6}$$

where τ now refers to the τ^{th} quantile of U_{it}, not U_{it}^*. In other words, the quantiles are now defined relative to a the child's fixed level of diet quality. While correct in specification, it is not the ideal interpretation for our primary policy question: what is the effect of food source on low dietary quality children separate from high dietary quality children.

In the estimation strategy laid out below, the policy variables D are allowed unspecified correlation with individual fixed effects, $\alpha_i = h(D_{i1}, \ldots, D_{iT}, \varepsilon_i)$. This arbitrary correlation mirrors mean fixed effects, but α_i is not estimated. The estimator therefore assumes that the unconditional distribution of U_{it}^* is uniform but relaxes the conditional distribution assumption by allowing $U_{it}^*|D_{it}, \alpha_i \nsim U(0,1)$. Specifically, I will estimate a specification that is related to the SQF taking the form

$$S_Y(\tau|d) = \gamma_{ht}(\tau) + d'\beta(\tau), \quad \tau \in (0,1) \tag{3.7}$$

where τ now refers to U_{it}^*, a child's total proneness to consume a healthy diet, which is precisely what we want to estimate. Note that the SQFs in equations (3.3) and (3.7) are the same, except that the implicit constant in (3.3) is now explicitly defined as γ_{ht}. The parameter vector γ_{ht} plays the primary role for identification (sample moment 2 in the next section). The index h can refer to any set of exogenous characteristics that saturate the sample space over time t, or simply time itself. For example, it is likely that a "high-quality diet" on a weekday is much different than a "high-quality diet" on the weekend, if not for the simple fact that children do not attend school on the weekend. Therefore, I construct fixed effects based on the interview day ($t = 1, 2$) and the day of the week in which the survey took place ($h = 1, \ldots, 7$).

3.6 Estimation

The SQF I will estimate is

$$S_{HEI_{it}} = \gamma_{ht}(\tau) + FFS_{it}\beta_1(\tau) + FAFH_{it}\beta_2(\tau), \quad \tau \in (0,1). \tag{3.8}$$

The underlying model corresponding to (3.8) is

$$HEI_{it} = \gamma_{ht}(U_{it}^*) + FFS_{it}\beta_1(U_{it}^*) + FAFH_{it}\beta_2(U_{it}^*) \tag{3.9}$$

where HEI_{it} is the measure of diet quality and γ_{ht} contains the 14 fixed effects as defined by the space $ht = \{h \times t\}$. Estimating equation (3.9) is not straightforward; the function is highly non-convex with many local optima, but it does have a well-pronounce global optimum. For brevity, I list the moment conditions here because they give intuition how estimation proceeds. See Powell (2014) for full details of estimation.

Referring to equation (3.9), let $D \equiv (\gamma_1, \ldots, \gamma_T, X)$ where $X = (FFS, FAFH)$ are the policy variables of interest. To simplify notation of the moment conditions, I will refer to γ_{ht} as simply γ_t but note that the fixed effects still refer to the t^{th} day of intake on the h^{th} day of the week. The (weighted) sample moments are

$$g_i(b) = w_i \frac{1}{T}\sum_{t=1}^{T} X_{it}\left[\mathbf{1}(Y_{it} \le D_{it}'b) - \frac{1}{T}\sum_{s=1}^{T}\mathbf{1}(Y_{is} \le D_{is}'b)\right] \tag{3.10}$$

$$h_t(b) = \frac{1}{N}\sum_{i=1}^{N} w_i \mathbf{1}(Y_{it} \le D_{it}'b) - \tau \quad \text{for all } t \tag{3.11}$$

where w_i is the survey weight supplied by NHANES, which has been normalized to sum to N. The fixed effects force $h_t(b) = 0$ for all t, thus confining all "guesses" of b to the parameter set $\mathcal{B}$,

$$\mathcal{B} \equiv \left\{ b \;\middle|\; \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}(Y_{it} \le D_{it}'b) = \tau \quad \text{for all } t \right\}. \tag{3.12}$$

By letting $\tilde{b}$ be the the coefficient vector on X_{it} we can write $D_{it}'b = \gamma_t + X_{it}'\tilde{b}$. Recalling that we have allowed arbitrary correlation between the fixed effects and the policy variables, we can define $\gamma_t(\tau, \tilde{b})$ as the τ^{th} quantile of the distribution $Y_{it} - X_{it}'\tilde{b}$ for each fixed-effect value t. Therefore, $\hat{\gamma}_t(\tau, \tilde{b})$ solves

$$\frac{1}{N}\sum_{i=1}^{N} w_i \mathbf{1}(Y_{it} - X_{it}'\tilde{b} \leq \hat{\gamma}_t(\tau, \tilde{b})) = \tau \tag{3.13}$$

and it immediately follows that for any guess $\tilde{b}$, $\hat{\gamma}_t(\tau, \tilde{b})$ is known.

Estimation proceeds in a Generalized Method of Moments (GMM) framework:

$$\hat{\beta}(\tau) = \arg\min_{b \in \mathcal{B}} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} g_i(b) \right)' W_n(b) \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} g_i(b) \right). \tag{3.14}$$

For surveys with stratification and clustering, the weighting matrix $W_n(b)$ is defined following Bhattacharya (2005, equation 6). With one or two treatment variables, grid searching is computationally achievable (Chernozhukov and Hansen, 2008) but can still be quite burdensome in practice. If bootstrapping is necessary for inference, the problem is exasperated.

Yu and Moheed (2001) show that parameters in a quantile regression can be estimated via Markov chain Monte Carlo (MCMC) algorithms. Chernozhukov and Hong (2003) generalized this technique into a GMM framework, which is suitable in this instance due to the complex survey design. Moreover, Chernozhukov and Hong (2003) show that inferences can be drawn from the posterior distribution. This key development dramatically reduces computation.

In this paper, I use an adaptive MCMC algorithm using code developed by Powell, Smith, and Baker (2014). Appendix B.4 provides details of the algorithm. In short, estimators and 90-percent confidence intervals are taken from the mean and quantiles of the posterior distribution.

3.7 Results

Figures 3.1 and 3.2 plot coefficient estimates for food from school (FFS) and food away from home (FAFH), respectively, using equation (3.9).[15] As a reminder, the policy variables are defined as the proportion of daily calorie intake from each food source (i.e., $D_{it} \in [0, 1]$). Thus, results can be interpreted as the marginal impact of reallocating calories from FFH to either FFS or FAFH by rescaling estimates using the average conditional allocation of calories found in table 4.1. For example, the coefficient estimate for FFS at the fifth percentile of the HEI distribution is 5.03 and holding calorie consumption constant at 33-percent, results in a 1.68 point increase. School meals are not of higher relative quality across the entire distribution – the coefficient estimate at the 95^{th} percentile is -3.08, implying a 1.03 point decrease in HEI when shifting 33% of calories from home to school.

Although I do not estimate the long-run impacts on dietary quality, numerous experimental trials have shown that simple and repeated exposures to new and healthy foods have lasting impacts on dietary choices (Benton, 2004). For children falling in the lowest quartile of the HEI distribution, figure 3.1 implies a positive daily investment in nutrition skill formation, which could have long-run implications on nutrition capital accumulation.

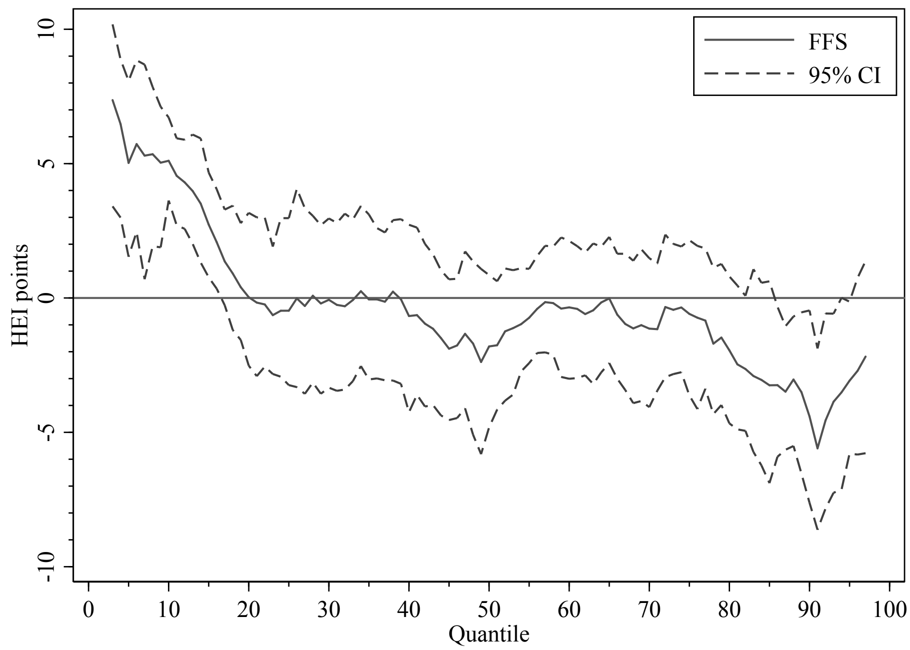
Figure 3.2 clearly shows the negative effects of FAFH on dietary quality. One important finding from figure 3.2 is that home-prepared food is of no higher quality than FAFH in the bottom portion of the HEI distribution. Coupled with the findings in figure 3.1, results suggest that kilocalorie consumption at school increases dietary quality to a larger degree than both home and away-from-home calories for those prone to relatively low quality diets.

3.8 Discussion and Conclusion

Food preferences reflect a complex cognitive structure rooted in early childhood experiences, exposures and environments. The formation of skills related to nutrition, such as

[15]In appendix table B.2, I report coefficient estimates and 95-percent confidence intervals for every fifth quantile. As a basis of comparison, in the appendix I also plot results from two cross-sectional quantile regressions.

Figure 3.1: Impact of Food from School (FFS) on the Distribution of Dietary Quality

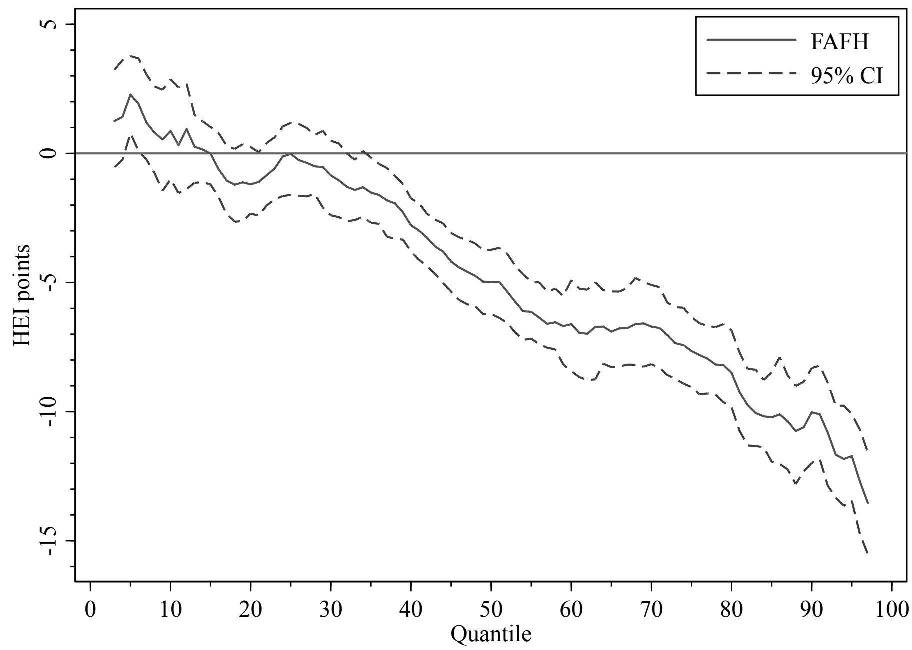


Note: Dietary quality is measured using the HEI-2005. Food from school (FFS) is measured as the share of daily caloric intake from school food.

ability to maintain energy balance and reach satisfactory levels of dietary quality, are learned and applied at early stages in life (Benton, 2004; Birch, 1999). The ability to adopt new skills, however, dissipates as one reaches adulthood (Morales et al., 2002). This insight into nutrition skill formation largely mirrors the expanding body of literature that has emerged in the past decade promoting skill formation and human capital development at early ages (Cunha et al., 2006).

The SBP and NSLP have undergone many reforms since the 1960's (Ralston et al., 2008). Originally aimed at alleviating hunger and malnutrition, these programs now strive to reach a balance between nutritional quality and caloric quantity. The most recent reform came from the Healthy, Hunger-Free Kids Act of 2010. Officially in effect for the 2013 school year, schools now have to meet new caloric and nutritional standards (Code of Federal Regulations, Title 7, Parts 210 and 220, 2013). Early evidence suggests higher standards for school meals improve child health outcomes (Taber et al., 2013).

Figure 3.2: Impact of Food Away from Home (FAFH) on the Distribution of Dietary Quality



Note: Dietary quality is measured using the HEI-2005. Food away from home (FAFH) is measured as the share of daily calories consumed away from home.

Moreover, localized experiments have shown that children are more likely to choose more nutritious meals after such a program is introduced and tend to make progressively healthier food choices the longer the program is in place (Grainger, Senauer and Runge, 2007).

Results of this study suggest there exists a large and meaningful impact of food served under National School Lunch and Breakfast Programs for children that exhibit low underlying dietary quality. These two Federal programs help children from nutrtionally-disadvantaged environments to experience much needed dietary exposure and variety. The daily exposure to a higher quality meal potentially has a lasting and positive impact on nutrition capital accumulation. As policymakers and health advocates look to policy-amendable arenas to improve the American diet, this study suggests the NSLP and SBP are fertile grounds for intervention.

Chapter 4

Cashing Out SNAP: Heterogeneous Impacts on Dietary Quantity and Quality

4.1 Introduction

The Supplemental Nutrition Assistance Program (SNAP) is the United States' principal food and nutrition assistance program, providing $76 billion in food vouchers to over 47 million individuals in 2013 (FNS, 2014); nearly half of all children will receive SNAP at some point before age 20 (Rank and Hirschl, 2009). SNAP has a dual mandate, both to "alleviate hunger" and also to "permit low-income households to obtain a more nutritious diet" (Food, Conservation, and Energy Act of 2008). SNAP attempts to accomplish this goal by providing benefits in the form of an in-kind transfer to ensure that publicly provided benefits are spent on food.

The extent to which SNAP accomplishes its dual mandate, over and above what could be accomplished via an unconditional cash transfer remains an open and important question. Standard economic theory states that as long as the total value of food assistance is less than the household's food budget, the manner in which benefits are transferred should not alter consumption (Southworth, 1945). On the other hand, behavioral economics, specifically the notion of mental accounting or income fungibility – the idea that "money in one mental account is not a perfect substitute for money in another account" (Thaler, 1990) – may provide theoretical support for the hypothesis that the form of the benefit may affect the way in which it is spent.[1]

The empirical evidence on this question is mixed with some studies finding evidence in support of income fungibility (Hymans and Shapiro, 1978; Moffit, 1989; Whitmore, 2002; Hoynes and Schanzenbach, 2009) and others finding evidence against it (Senauer and Young, 1986; Levedhal, 1995; Beatty and Tuttle, 2014) or with mixed results (Breunig and Dasgupta, 2002, 2005). Moreover, differences in expenditures do not necessarily translate into differences in consumption (Aguiar and Hurst, 2005). Previous work that focuses on spending rather than consumption may miss the effect of an in-kind transfer on diet quality and quantity.

We investigate whether receiving food assistance in the form of cash, rather than as a voucher, affects either the quantity and/or the quality of food consumed in the home. This is particularly important to policymakers because the current policy prescription

[1]This type of behavior is sometimes called a 'labeling effect' (e.g, Kooreman, 2000; Beatty et al., 2014). The labeling effect is a behavioral response to the name given to benefit transfer (be it cash or in-kind) that increases the consumption of that good.

implicitly assumes that a voucher dollar is marginally no worse than a cash benefit dollar. With regards to nutrition, the assumption is that beneficiaries' food supplies at home are of equal or superior quality when given food vouchers as compared to an equal transfer of cash. The current program also assumes that the quantity of food used in the home is closer to optimal under the voucher system. We investigate to what extent these assumptions hold and what this implies for SNAP policy. Our identification strategy relies on the cornerstone "cash-out experiment" conducted in 1990 in which a randomized set of the food stamp caseload in San Diego county received their benefits in the form of check rather than food stamps (i.e., the voucher).

This paper advances the literature in many dimensions. First, we study the cash-out effect on consumption rather than expenditure. This aligns our study with the stated goals of the SNAP program. We measure the quantity of food used by a household in terms of kilocalories. Previous work has largely studied the cash out effect in terms of expenditure (Moffit, 1989; Levedhal, 1995; Breunig and Dasgupta, 2002, 2005) although others have investigated kilocalories with mixed results (Ohls et al., 1992; Bishop, Formby and Zeager, 2000; Whitmore, 2002).[2] Second, we measure the quality of food used in the home via a single dietary index. Previous research has taken a one-at-a-time approach – studying individual nutrients and/or food groups (Ohls et al., 1992; Bishop, Formby and Zeager, 2000; Whitmore, 2002). When studied in this manner, drawing a conclusion about the overall effect of receiving cash benefits on diet quality is difficult. We take a holistic view of dietary quality and calculate a single dietary score using a widely-used and validated measure, the Healthy Eating Index.

We also innovate with regards to method. We study the effect of cash-out on the distribution of diet quality and quantity; allowing for heterogeneity between low-consuming households and high-consuming households is important. For example, we may be more concerned about the effect of cash-out on outcomes below or above some reasonable threshold rather than central effects. In contrast to previous studies that have taken a distributional approach, we model *control* variables separate from the *policy* variables. The distinction is subtle but important when interpreting quantile treatment effects (see, for example, Chernozhukov and Hansen, 2005; Powell, 2013). Policy variables

[2]Ohls et al. find no effect. Bishop et al. and Whitmore both find some evidence of overconsumption among voucher recipients, although insignificant.

have a direct impact on the outcome distribution (i.e., a structural treatment effect interpretation), whereas control variables aid in identification and/or possible variance reduction (i.e., an econometric or modeling issue). We discuss these differences and their importance in more detail in the methods sections.

The paper proceeds as follows. We first discuss the related background, followed by a description of our primary dataset, the San Diego cash out experiment. We then show how we construct a measure of dietary quality. We then introduce the structural model of interest. A descriptive analysis is used to motivate our estimation and inference procedures. We concluded with policy implications.

4.2 Background

4.2.1 Mental Accounting and Income Fungibility

We briefly describe two components of mental accounting that relate to the decision making process of food assistance beneficiaries.[3] The first component is how households assign income sources (or flows of resources) to specific accounts based on their usage. For example, a household may have an account to purchase food and a separate account to pay utility bills. The type of resource (e.g., cash, voucher or gifts) can dictate the overall assignment of all resources to each account. If income is not perfectly fungible, then a voucher may be assigned differently than an equal value of cash.

The second component of mental accounting related to food purchases is “transaction utility.” Thaler (1985, 1999) describes the utility derived from a purchase transaction as the value of the “deal,” but we can also think of this as the shadow price of the income source. Indeed, Levedahl (1995) provides a theoretical framework in which food purchased with cash enters the household utility function separately from food purchased with a voucher. In other words, the relative utility of a cash transaction versus a voucher transaction for an otherwise identical food bundle can yield different utilities.[4]

[3] See Thaler (1980, 1985, 1990, 1999) for a more complete treatment of mental accounting.

[4] Specifically, suppose F is food bought with cash, S is food bought with vouchers and X is a composite good where p is the price of food relative to the composite good. The household has total resources Y which is the sum of cash income I and voucher income V. The household maximizes $L = U(F, S, X) + \lambda_1(I - pF - X) + \lambda_2(V - pS)$ where λ_1 and λ_2 are the shadow prices of cash and voucher income, respectively. In equilibrium, the total demand for food will be a function of λ_2/λ_1.

Taken together, these two components of mental accounting describe the theoretical framework in which the way SNAP recipients receive benefits can affect the budgeting process and ultimately the bundle of goods purchased. The next section summarizes the large body of empirical evidence against the notion of income being perfectly substitutable.

4.2.2 Empirical Evidence on Income Fungibility

The question of how food assistance benefits affect expenditure allocations is not a new one (see, Southworth, 1945). Hymans and Shapiro (1976) studied how three income sources – wage, cash transfers and food stamps – affect food expenditures. They find that income-supplement programs "permit a higher standard of food consumption. But they have minimal effects on either the overall marginal propensity to consume food or the income elasticity of food consumption." In effect, they find no difference in expenditures.

Senauer and Young (1986) used data from the 1978-79 Panel Study of Income Dynamics, which straddles an important policy change in the food stamp program – the elimination of the purchase requirement. Despite the program rule changes, they find that in both years the traditional Southworth model is rejected. Interestingly, they (indirectly) cite behavioral mechanisms as one possible explanation. They hypothesize that the in-kind transfer generates a "sense of gratitude or responsibility...to expand their food consumption" (i.e., a labeling effect).

More recent papers have used the San Diego cash out experiment to test income fungibility. Levedahl (1995) develops a theoretical model whereby food bought with cash and stamps enter the utility function as two separate goods. When testing several specifications, Levendahl also finds a higher marginal propensity to spend (MPS) on food when using stamps, also lending evidence against perfect substitutability of resources.

In a series of papers using the San Diego data, Breunig and Dasgupta (2002, 2003, 2005) again find a higher MPS on food purchased with stamps and reject the hypothesis that stigma is the underlying cause. Breunig and Dasgupta (2005) show that households with a single decision maker, as opposed to two or more, exhibit no such cash out effect. Although their main hypothesis is that the intrahousehold allocation process contributes

Cash and voucher income are perfectly fungible only when λ_2/λ_1 is equal to 1.

to the cash out effect, they only test this theory indirectly by calculating the MPS for singles versus multi-headed households. An alternative hypothesis consistent with their finding is that households with one decision maker have a different array of resources as compared to those with more than one decision maker. This would follow from an economies of scale argument.

Finally, Whitmore (2002), using the San Diego data, finds no differences in consumption patterns. On the other hand, she finds that the probability of consuming more than twice the RDA for calories is lower among check recipients.

4.3 Description of Data Sources

Our primary data source is the 1990 San Diego cash out experiment. The purpose of this randomized experiment was to determine how households respond to receiving their food stamp benefits in the form of a check rather than stamps. The experiment began in July 1989, and the survey took place almost a year later from May to mid-August 1990. In September 1990 the remaining caseload and all new cases were cashed out for the next five years.[5]

We utilize in-person surveys of about 600 check and 600 coupon recipients. Selection into these groups were randomly assigned based on the last digit of the household's program case number. The initial in-person interview obtained demographic information from the household head and all household members. The interviewer then notified respondents they would come back seven days later to ask about food used in the household over the seven-day period and about purchases of food at eating establishments.

The main interview occurred seven days later to obtain information about food used in the home over the seven day period. Quantities (in pounds), the price paid, and its source (purchases, WIC voucher, or home-produced/gifted) of all food items used in the home were recorded and grouped into 32 food groups. The main interview also reported

[5]Four cash-out demonstrations were implemented in the early 1990s. The San Diego and Alabama studies were "pure" cash out in the sense that they were randomized experiments (Ohls et al., 1992). However, the Alabama cash-out demonstration was limited to 8 months, the response rate was correlated with treatment status, and some check-receiving households were encouraged to spend their benefits on food (Whitmore, 2002). The Washington State and Alabama ASSETS studies were part of a more comprehensive welfare reform experiment and did not focus solely on cash out. As a result, we focus on the San Diego demonstration as done elsewhere in the literature.

expenditures at eating establishments outside the home over the previous week but did not distinguish these expenditures by food group nor report quantities. In the analysis below, we will use these expenditure categories as a measure of available resources to the household to obtain food. The main interview also asked detailed expenditure information over the previous month for nine other broad categories (housing, utility, medical, transportation, clothing, eduction, dependent care, recreation and personal items). Together, these represent non-food expenditures.

Table 4.1 provides summary statistics for all control variables used in our analysis. As done in previous studies, we drop households that do not report complete information, individuals that are homeless or reside in a group home, and two households that report conflicting information ($n = 151$). As expected, only food bought with cash and voucher income are significantly different between the treatment and control groups (p-value = 0.03).

4.3.1 Measuring Quantity

We use kilocalories as our measure of the quantity of food used in the household. The total amount of calories consumed from the household supply over the seven-day period will depend on household composition, meals served to guests, and meals eaten away from home. In order to compare calorie consumption across all households, we need to rescale calories accordingly (Ohls et al., 1992). In short, this scale is a function of the number of household members, their age/gender, number of meals eaten away from home and the number of meals eaten by guests.

The primary factor is the age/gender composition of household members. For example, as shown in table 4.2, a twelve year old girl's calorie needs are about three-fourths of that of a 30 year old male. Setting the reference person as a 23-50 year old male, the sum of each household member's relative needs generates the household size in adult male equivalents (AME). Thus, the household size for the family of four in table 4.2 in AME is 3.55.[6]

[6]The Recommended Daily Allowance (RDA) for calories given in table 4.2 correspond to 1990's standards. The most current RDA for calories further delineates by three activity levels (sedentary, moderately active and active). The 1990 RDAs given in the data roughly correspond to an active lifestyle. This does not affect our analysis because the relative needs ratios are similar across activity levels. Moreover, the data do not permit us to calculate activity levels.

Table 4.1: Summary statistics for control variables

Variable	Cash	Voucher	p-value
Household Head Characteristics			
Non-Hispanic white	0.34	0.31	0.285
Non-Hispanic black	0.18	0.22	0.103
Hispanic	0.32	0.34	0.617
Other race/ethnicity	0.15	0.13	0.262
Age	32.75 (9.91)	32.16 (9.61)	0.325
Female	0.87	0.89	0.523
Married	0.26	0.23	0.267
Employment and Education			
Less than highschool	0.44	0.42	0.490
Highschool graduate	0.23	0.24	0.694
Some college	0.33	0.34	0.712
At least one member employed	0.19	0.21	0.473
Household Composition			
Infant present (1-2yr)	0.27	0.27	0.915
Toddler present (3-5yr)	0.48	0.50	0.625
Adolescent present (6-11yr)	0.50	0.50	0.806
Teenager present (12-18yr)	0.35	0.34	0.831
Single decision maker[a]	0.60	0.57	0.356
log(household size)	1.12 (0.47)	1.17 (0.46)	0.114
Resources			
Food purchases ($/wk)	65.70 (40.63)	71.31 (44.08)	0.031
WIC purchases ($/wk)	2.61 (6.59)	2.69 (11.56)	0.890
Food gifted/home-grown ($/wk)	1.37 (5.39)	1.11 (3.93)	0.366
Food away from home ($/wk)	9.59 (17.86)	10.40 (17.68)	0.458
Non-food expenditures ($/month)	651.14 (303.81)	651.69 (306.89)	0.977
Number of households	532	530	

Source: 1990 San Diego Cash Out Demonstration.

[a]One non-child adult in the household. See Bruenig and Dasgupta (2002, 2005).

Notes: Standard deviation in parentheses. p-values test the null of equality between cash and voucher.

A second consideration must be given to the number of meals eaten at away establishments versus from household stocks. In the example from table 4.2 we can see that the male head ate two-thirds of his meals at home. Therefore, when calculating the effective household size in Equivalent Nutrient Units (ENU), we scale him down by

Table 4.2: Calculation of household size in AME and ENU

Household member	RDA for energy (kcal)	Relative needs (AME)	Portion of meals at-home	Equivalent Nutrient Units
Male, aged 30	2,900	1	0.67	0.67
Female, aged 30	2,200	0.76	1	0.76
Male, aged 15	3,000	1.03	1	1.03
Female, aged 12	2,200	0.76	1	0.76
Household size		3.55		3.22

Note: AME=adult male equivalent. ENU = Equivalent Nutrient Units. RDA=Recommended Daily Allowance.
Source: Recreated from Ohls et al. (1993).

0.67. In this manner, the example household size in ENU for the survey week is 3.22.

The final consideration is the number of meals served to guests from household stocks. The data provide the number of meals served to guests by age/gender in AME. In the final calculation of effective household size in ENU given in the data, these meals are taken into account. Note that in the example from table 4.2 no meals were served to guests, so that the final size is 3.22. The use of ENU is the preferred equivalence scale, rather than simply AME (Ohls et al., 1993). Moreover, using ENU is consistent with previous studies (Ohls et al., 1992; Bishop, Formby and Zeager, 2000; Whitmore, 2002).

To summarize, we use the total number of reported calories used from household food supplies. We then normalize calories by the household's effect size in ENU. In the descriptive statistics and analyses below, we report calories on a daily basis by dividing by seven days. Since the reported kilocalories represent purchase quantities and not actual consumption, we scale down kilocalories to account for food spoilage, waste and inedible portions. Most current research estimates food losses at the household level in 1990 to be about 37% ± 6% (Hall et al., 2009). Note that these monotonic transformations have no effect on inference, but rather give a more accurate picture of actual consumption on a per-day, per-person basis.[7]

[7] We believe this is important because previous studies (Bishop et al. and Whitmore) have used the ratio of total available calories to the RDA as a measure of food adequacy. This would be an inaccurate measure since many of the available calories are not actually consumed.

4.3.2 Measuring Quality

We construct a *per-calorie* diet quality index based on the quantities of food used from household food supplies. We use the Healthy Eating Index, which was developed in 1995 to measure compliance to the U.S. Government's official recommendations for healthful eating, the *Dietary Guidelines for Americans* (DGA). Every five years, based on an expert advisory panel, the DGA are revised by the U.S. Departments of Agriculture (USDA) and Health and Human Services (HHS). As such, the HEI has been updated several times to reflect the most current state of nutrition knowledge. This paper uses the HEI-2005 and will henceforth refer to the HEI-2005 as simply HEI.[8]

The HEI is the sum of 12 components based on the consumption of various foods or nutrients relative to total consumption. In this way, the HEI captures the relative quality of all foods consumed within the household and is invariant to household composition due to its per-calorie measurement. Each component assigns a score ranging from 0 to 5 (total fruit, whole fruit, total vegetables, dark green/orange vegetables and legumes, total grains, whole grains), 0 to 10 (milk, meats and beans, oils, saturated fat, sodium) or 0 to 20 for the percentage of calories from solid fats, alcoholic beverages, and added sugars (SoFAAS) creating a maximum score of 100. Appendix table C.1 provides exact details of the scoring (see also, Guenther et al., 2008a).

The HEI has been widely used and evaluated as a valid measure of diet quality (Guenther et al., 2008b). In the medical literature, lower HEI scores are associated with higher risks of coronary heart disease, stroke and diabetes (Chuive et al., 2012), cardiovascular disease (Nicklas et al., 2012), and several cancers (Bosire et al., 2013; Reedy et al., 2008; Shahril, 2012). It is important to reiterate that the HEI is a per-calorie measure of dietary quality and does not directly consider excessive calorie intake. This distinction is necessary to analyze the relative quality of foods consumed across households of differing calorie needs, and it allows us to analyze quality separate from quantity.

Like most food recall surveys, respondents report foods, or in this case 32 food groups, and not the individual components of the HEI. For this reason, the USDA has created a dataset that maps some 7,000 individual foods into their HEI components. The

[8]Future work will use the HEI-2010. Note that the two have many similarities (see, National Cancer Institute, 2013).

32 food groups from the San Diego data are not ad hoc, and were chosen based on the Thrifty Food Plan (TFP).[9] In fact, food data collected in the San Diego experiment use the same data entry system, the USDA Human Nutrition Information System (Ohls et al., 1992, p. C.4), used in USDA's 1989-91 Continuing Survey of Food Intakes by Individuals (CFSII). Therefore, the individual foods reported in the CSFII have a rather natural mapping into the 32 TFP food groups.[10]

Using the 1989-91 CSFII, we can calculate the average amounts of each of the 12 HEI components for the 32 food groups on a per-gram basis. For example, the "high-fiber breakfast cereal" group contains 0.03 cups of fruit, 6.1 grams of whole grains and 395 kilocalories per 100 grams (roughly 3 servings).[11] In order to more accurately reflect the mix of foods eaten within each group by beneficiaries in San Diego, we use only CSFII households that reside in the Western U.S. and have income less than 130% of the poverty line.[12]

As a measure of validating our approach we can compare imputed values of kilocalories and saturated fat from the CSFII data to those reported in the San Diego data. Perhaps unsurprisingly, given that similar food recall and data entry methods were used, the imputed measures of these two nutrients are similar using either method (see appendix figures C.1 and C.2).

[9]The TFP is USDA's yardstick for determining benefit levels for many of its nutritional assistance programs, such as the Food Stamp Program in 1990.

[10]See appendix table C.2 for details, which follows the methodology outlined in TFP reports (Carlson et al., 2003, 2007).

[11]Each nutrient value is calculated for 100 grams of edible food. The San Diego data report poundage of food used from stocks, not the amounts eaten. This does not affect our analysis of quality because every HEI component is measured as the ratio of component intake to calorie intake. In other words, the amount of food that is wasted or spoils cancels out in the calculation of each HEI component.

[12]We also considered using the entire sample that report receiving food stamps, as well as everyone in the sample below 130% of the poverty line. However, when we examined the demographic makeups and the types of food eaten across these three groups, we concluded that conditioning on geography and income resulted in a large enough sample and best reflected the San Diego sample.

4.4 Descriptive Measures

4.4.1 Mean Outcomes

The unconditional mean of kilocalorie consumption per day per adult male equivalent after adjusting for waste and spoilage is 2,486 with a standard error of 33.7.[13] The mean HEI-2005 score is 49.93 with a standard error of 0.24.[14] The first column of table 4.3 reports the simple difference in means by regressing the outcome on a dummy for cash households. Columns 2-5 subsequently add in more control variables found in table 4.1. With the exception of the unconditional case for kilocalories, there appears to be no mean effect of receiving benefits in the form of cash or as a voucher.

Table 4.3 does provide us with some additional information. We can see that some of the covariates have a significant impact on households' placement in the outcome distribution. Specifically, resources have a positive impact on a household's central tendency (see appendix tables C.3 and C.4). The R-squared increases substantially when including resources into the regression for kilocalories indicating their value in the inclusion of the analysis. In the next section, we display results for the unconditional distributions.

4.4.2 Distributional Outcomes

We plot the cumulative distribution functions (CDFs) of kilocalorie consumption and HEI scores in figures 4.1 and 4.2, respectively. As noted above, kilocalories have been normalized to adult male equivalents and adjusted for wasted and spoilage. The inverses of the CDFs describe the quantile processes for the cash and voucher groups; the difference between these processes lead to the unconditional QTEs.

Recommended kilocalorie intakes are given in three levels depending on one's level of physical activity. These three levels are represented by the vertical lines and correspond to a 30-year old male. We can see roughly half of all households report consuming food below all three recommendations, while over 20% report consuming levels well above the

[13] The average daily intake of kilocalories in AME in the 1989-91 CSFII sample used to calculate the HEI-2005 in this study is 2,340 with a standard error of 32.8.

[14] Beatty, Lin and Smith (2014) calculate HEI-2005 scores for low-income adults in the 1989-91 CSFII. They find an average of 48.96 for this nationally representative sample, which is reassuring to the imputation method used in this study.

Table 4.3: Mean Effect of Receiving Cash on kilocalories and HEI-2005

	(1)	(2)	(3)	(4)	(5)
Outcome: kilocalories					
Cash	-114.068*	-92.064	-91.755	-107.969	-40.248
	(67.691)	(66.957)	(67.068)	(67.089)	(54.223)
Constant	2544.263***	2214.279***	2271.766***	2480.040***	1922.301***
	(47.910)	(195.593)	(211.894)	(242.139)	(199.770)
R-squared	0.003	0.035	0.035	0.047	0.383
Outcome: HEI-2005					
Cash	0.529	0.280	0.305	0.306	0.412
	(0.473)	(0.455)	(0.453)	(0.453)	(0.449)
Constant	49.650***	49.004***	50.570***	48.359***	47.937***
	(0.335)	(1.328)	(1.432)	(1.635)	(1.655)
R-squared	0.001	0.089	0.098	0.110	0.134
Demographics[a]		*X*	*X*	*X*	*X*
Edu. & emp.[a]			*X*	*X*	*X*
Composition[a]				*X*	*X*
Resources[a]					*X*
Observations	1,062	1,062	1,062	1,062	1,062

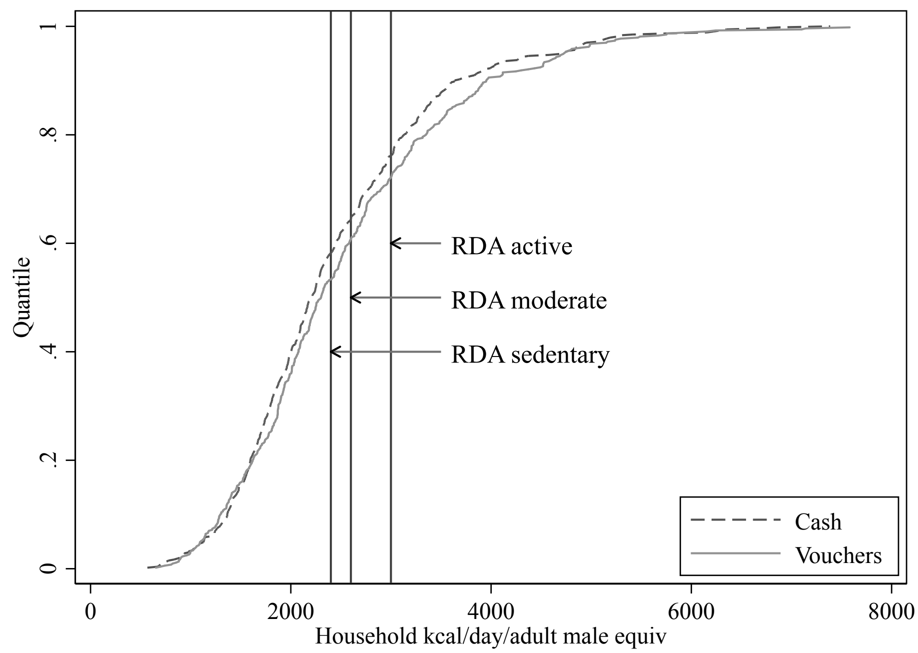
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

[a]See table 4.1 for definitions.

recommendation for an active lifestyle. We acknowledge this is a rather noisy approach to accounting for food spoilage (although it does not affect inference procedures), and that some households may have a mix of sedentary and active individuals. Nevertheless, the point is that our sample consists of a variety of households consuming below, at, or above any reasonable threshold for calorie consumption. Overall, it appears that vouchers encourage food consumption at least above the 20^{th} percentile.

Given figure 4.1, we can see that a stochastic dominance approach for kilocalories, as done in Bishop et al., is inappropriate. This is because one of the underlying assumptions of stochastic dominance is that expected social welfare is monotonic in kilocalorie consumption. Clearly, society would prefer those below subsistence levels, or food insecure households, to consume more calories. On the other hand, society would be better off if over-consuming households consumed less food. Therefore, one can not maintain

Figure 4.1: Distribution of kilocalorie consumption from household food supply



Note: Calories are measured in Equivalent Nutrient Units (ENU), which accounts for the number of household members, their age/gender, number of meals eaten away from home and the number of meals eaten by guests. We also scale the x-axis down by 37% to account for food spoilage and waste. Active, moderate and sedentary refer to physical activity lifestyles and corresponding RDA for kilocalories. See text for explanations.

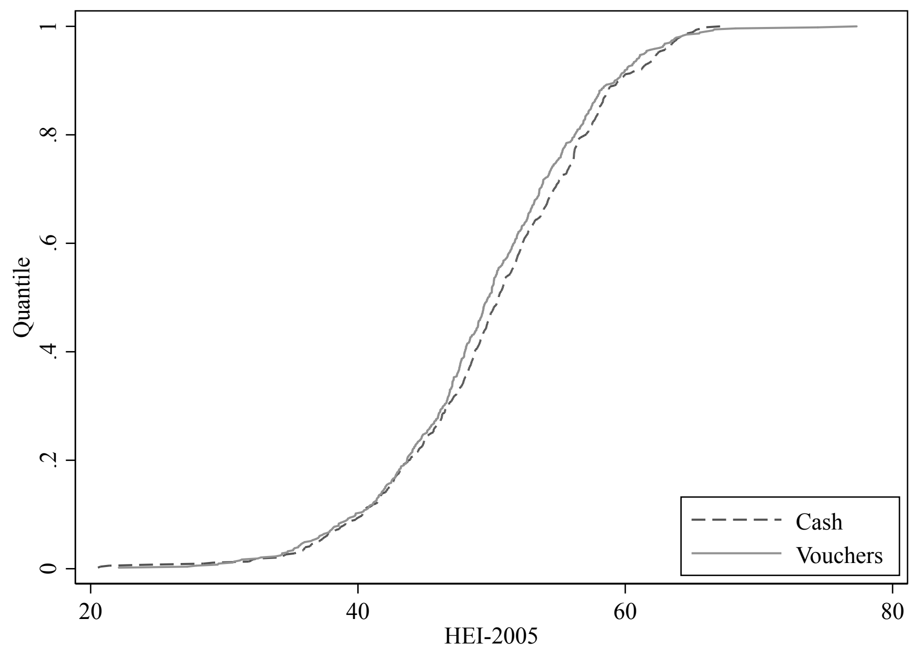
the assumption that social welfare is increasing (or decreasing) at all levels of kilocalorie consumption. The dominance approach however would be appropriate for quality.

The next section develops a framework in which we make no assumptions about social welfare but can test for changes in the distribution of outcomes. We will then estimate a model that controls for covariates while estimating the unconditional QTE.

4.5 The Structural Model

Our model follows a potential outcome framework akin to Heckman and Robb (1986) and extended to quantile treatment effect (QTE) models in Chernozhukov and Hansen (2005). Under this framework, a policymaker has two distinct policies $d \in \{0, 1\}$ from

Figure 4.2: Healthy Eating Index-2005 scores for household food supply



which they may choose to impose on SNAP households. The potential outcome for a household in each treatment state is denoted by Y_d. We refer to these outcomes as potential (or latent) because the true treatment status D is only observed for one component of the potential outcomes, $Y \equiv Y_D$.[15]

This research aims to learn about the distribution of the potential outcomes under a cash versus voucher system. The τ^{th} quantile of the potential outcomes vector $\{Y_d\}$ can be described by the quantile treatment response function $q(d, \tau)$. This function has a structural, or counterfactual interpretation because τ indexes unobserved "proneness," or preference heterogeneity.[16] Under this framework, high values of τ indicate a high preference for food quality or food quantities and vice-versa. Understanding the distributional impacts of the cash out effect is important when we consider the heterogeneity of food preferences.

[15]Following the literature, capital letters denote random variables (i.e., observed in the data) and a lower case letter denotes the potential outcome (i.e., the counterfactual outcome to be modeled).

[16]Doksum (1974) introduced the term proneness, such as "prone to earn high wages" (i.e., ability) or in this case, "prone to consume a high quality diet" or "prone to consume high quantities of food."

The quantile treatment effect (QTE) is the primary object of interest and describes the difference between the two potential outcomes

$$q(d_1, \tau) - q(d_0, \tau) \tag{4.1}$$

where d_1 is the treatment (i.e., cash) and d_0 is the control (i.e., vouchers). Equation (4.1) describes the QTE when cashing out SNAP, which varies by latent food preferences indexed by τ. Note that this is the "unconditional" QTE since we do include any other conditioning variables x and only include the policy variable d.

Under conditional quantile regression (Koenker and Bassett, 1978), the quantile treatment response function is $q(d, x, \tau)$ where control variables directly enter the structural equation. In this case, τ indexes the quantiles of the latent outcome Y_d for a fixed set of exogenous controls $X = x$. This changes the interpretation of the treatment effect to a conditional QTE and equation (4.1) becomes $q(d_1, x, \tau) - q(d_0, x, \tau)$. Estimates using this framework provide the impact of cash benefits for high-consuming households relative to their observed levels of X. Some of these household may actually consume low quantities of food or low quality food in the unconditional distribution.

Returning to the QTE in (4.1), the relationship between Y_d and the quantile response function $q(d, \tau)$ can be represented as (Chernozhukov and Hansen, 2005)

$$Y_d = q(d, U_d^*) \tag{4.2}$$

where U_d^* is a structural error term and is distributed uniformly over $(0, 1)$. U_d^* determines the relative rankings of households depending on treatment state. This interpretation leads to the monotonicity assumption standard in quantile analysis that $q(d, \tau)$ is increasing in τ (Chernozhukov and Hansen, 2005). Thus, U_d^* has a natural mapping to the latent proneness index τ. The rank similarity assumption of the model only requires U_d^* to be identically distributed across treatment states, i.e., $U_{d=1}^* \sim U_{d=0}^*$ (Chernozhukov and Hansen, 2005). $U_d^* \sim U(0, 1)$ is simply a normalization. Concretely, rank similarity implies households with a high preference for food, do so regardless of the treatment state. This assumption is maintained due to the randomization of the experiment.

The error term U_d^* introduces heterogeneity into the model by allowing a household's total proneness, or food preference, to vary within treatment groups. We say *total* proneness because some of this proneness can be observed in the data (e.g., per-capita spending on food or age composition of household members) while the rest is unobserved proneness (e.g., tastes). For example, those with relatively higher spending on food may have a higher taste for food relative to non-food; including this information enables us to better model their total proneness without letting this covariate alter the structural interpretation of equation (4.1). To see this more formally, Powell (2013) defines $U_d^* = f_d(X, U_d)$ where X are observed covariates that explain some of the proneness and U_d is the unobserved component. No functional form restrictions are place on $f_d(\cdot)$, U_d can be multidimensional, and X does not need to be exogenous (Powell, 2013).

By defining $U_d^* = f_d(X, U_d)$, the Powell estimator relaxes the conditional independence assumption of Chernozhukov and Hansen (2005) that $U_d^*|D, X \sim U(0,1)$ to the more flexible assumption that $U_d^*|D, X \sim U_d^*|X$. In other words, the Chernozhukov and Hansen framework treats both D and X as policy variables, whereas the Powell estimator allows X to inform the distribution of U_d^* without treating the additional covariates as policy variables. In either case, we still maintain the assumption that D is exogenous due to the randomization of the cash-out experiment. As pointed out in Powell (2013), if X is empty then all variables are policy variables and we return to the Chernozhukov and Hansen framework since $U_d^*|D \sim U_d^* \sim U(0,1)$.

When the preceding assumptions are met, as we maintain, we can identify an unconditional QTE while still conditioning on the set of covariates X. Specifically, for each $\tau \in (0,1)$ we have a conditional and unconditional quantile restriction,

$$P[Y \leq q(D,\tau)|X, D] = P[Y \leq q(D,\tau)|X] \tag{4.3}$$

$$P[Y \leq q(D,\tau)] = \tau \tag{4.4}$$

respectively. See Powell (2013) for proofs. These two conditions motivate the estimation technique proposed by Powell (2013) and summarized in the next section.

4.5.1 Estimation of Unconditional QTE

Estimation uses a method of moments procedure. The underlying model can be written as

$$Y_i = D_i'\beta(U_i^*) \tag{4.5}$$

where the nonseparable error term U_i^* is distributed uniformly over the unit interval. The first moment condition corresponds to the quantile restriction found in equation (4.3). Using the independence of D_i and $[P(Y_i \leq D_i'\beta(\tau))|X_i, D_i) - P(Y_i \leq D_i'\beta(\tau)|X_i)]$ we can define

$$g_i(b) = D_i\left[\mathbf{1}(Y_i \leq D_i'b) - \hat{\tau}_{X_i}(b)\right] \tag{4.6}$$

where $\hat{\tau}_{X_i}(b) \equiv \hat{P}(Y_i \leq D_i'b|X_i)$. Instead of assuming $\hat{\tau}_{X_i}(b) = \tau$ for all i, as would be the case if X_i were empty, the first moment condition allows us to relax this assumption and let τ be a function of the control variables. Thus, the control variables still inform the distribution of proneness to consume kilocalories or a high quality diet, but we still estimate the unconditional QTE.

The second moment condition uses the unconditional quantile restriction in equation (4.4). This leads to

$$h(b) = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}(Y_i \leq D_i'b) - \tau \tag{4.7}$$

which guarantees that $(100 \times \tau)\%$ of the observations have Y_i below $D_i'b$.

Powell (2013) outlines an estimation procedure where the second moment condition is used to confine all potential values of the coefficients to a compact set of values. First, separate out the constant and define $D_i = (1, \tilde{D}_i)$. Let $\gamma(\tau, \tilde{b})$ be the τ^{th} quantile of $Y_i - \tilde{D}_i'\tilde{b}$ such that

$$\hat{\gamma}(\tau, \tilde{b}) \quad \text{solves} \quad \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}(Y_i - \tilde{D}_i'\tilde{b} \leq \hat{\gamma}(\tau, \tilde{b})) = \tau. \tag{4.8}$$

Thus, for any guess $\tilde{b}$ the constant can be calculated by simply finding the τ^{th} quantile of $Y_i - \tilde{D}_i'\tilde{b}$. With a guess of $\tilde{b}$ in hand and the corresponding calculation of $\hat{\gamma}(\cdot)$ we can now estimate τ_{X_i} as a function of $\tilde{b}$ and $\hat{\gamma}$. This is accomplished with the first moment condition in equation (4.6) such that

$$\hat{\tau}_{X_i}(b) = \hat{P}(Y_i \leq \hat{\gamma}(\tau, \tilde{b}) - \tilde{D}_i'\tilde{b}|X_i). \tag{4.9}$$

One can see this is a probability model where the outcome equals 1 if Y_i is less than or equal to $\hat{\gamma}(\tau, \tilde{b}) - \tilde{D}_i'\tilde{b}$ and zero otherwise. The covariates are X_i. Powell suggests using a logit or probit model, although a linear probability model (LPM) would work just as well. Regardless of the model used to estimate the proneness index τ_{X_i}, we need estimation error and misspecification to be orthogonal to D_i, which again is maintained by the randomization of the experiment.[17]

Estimation proceeds in a Generalized Method of Moments (GMM) framework:

$$\hat{\beta}(\tau) = \arg\min_{b \in \mathcal{B}} \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} g_i(b) \right)' W_n(b) \left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} g_i(b) \right) \tag{4.10}$$

where $W_n(b)$ is an appropriate weight matrix. With one or two treatment variables, grid searching is computationally achievable using the inverse quantile regression approach of Chernozhukov and Hansen (2006). If bootstrapping is necessary for inference, as is the case here, the problem is exasperated.

Yu and Moheed (2001) show that parameters in a quantile regression can be estimated via Markov chain Monte Carlo (MCMC) algorithms. Chernozhukov and Hong (2003) generalized this technique into a GMM framework. Moreover, Chernozhukov and Hong (2003) show that inferences can be drawn from the posterior distribution given the appropriate choice of $W_n(b)$. We use a two-step procedure suggested by Yin (2009) to construct $W_n(b)$ using an adaptive MCMC algorithm outlined in Baker (2013). We first let the algorithm run during a burn-in period (2000 draws). After burn-in, we take 2,000 more draws and construct an efficient weight matrix $W_n(b) = [\frac{1}{n} g(\hat{b})' g(\hat{b})]^{-1}$ where

[17] Powell finds in his simulations and applications that a logit and probit work equally as well. A linear probability model was not investigated but there is no reason to rule it out. In fact, we only need $E[\hat{\tau}_{X_i}(b)] = \tau$ which is satisfied by either of the 3 choices. Moreover, the LPM does not involve the additional maximization procedure needed with a logit or probit and substantially speeds up computation.

$\hat{b}$ is the mean value of the draws. With an efficient weighting matrix, we take 10,000 more draws from which inferences are drawn. We also use this procedure for estimating the "standard" quantile regression where X_i is empty. The result gives robust inference due to the two-step GMM procedure.

4.6 Results

All results presented in this section correspond to the model,

$$Y_i = \gamma(U_i^*) + C_i' \beta(U_i^*) \tag{4.11}$$

where γ is a constant and C_i is a dummy if household i received their benefits in the form of cash. The standard quantile regression of Koenker and Bassett (1978) is labeled as QR in tables 4.4 and 4.5 and figures 4.3 and 4.4. We also estimate 4 specifications of the Generalized Quantile Regression (GQR).[18] The GQR specifications mirror the OLS specifications found in table 4.3. The interpretation of the GQR estimates also mirror the OLS estimates – they both can be interpreted unconditional treatment effects while controlling for a set of covariates. For example, OLS and median GQR both estimate the impact of cash benefits on a measure of central tendency for the outcome. Indeed, we can see results at the 50^{th} quantile across the 5 quantile specifications exhibit a similar pattern to those at the mean.

When examining the impact of cash in the tails of the kilocalorie distribution (table 4.4 and figure 4.3), we can see consistent estimates across QR, GQR-1, GQR-2 and GQR-3. When we include how households spend their resources (GQR-4), we see a positive shift in the QTE at all quantiles – the positive impact at low quantiles becomes significant and the negative impact at higher quantiles is closer to zero. The impact throughout the interquartile range is very flat, indicating virtually no effect in the center of the distribution of kilocalorie consumption. This finding lends evidence to the hypothesis that how benefits are paid impacts the fungibility of income and thus how resources are spent.

[18] A comparison of QR estimates to GQR-1, GQR-2 and GQR-3 can be found in appendix figures C.6-C.5.

Table 4.4: Effect of receiving cash on the distribution of kilocalories at selected quantiles

Outcome: kilocalories	QR	GQR-1	GQR-2	GQR-3	GQR-4
Quantiles					
10	48.64	51.34	48.30	47.93	81.88*
	(55.33)	(56.37)	(56.06)	(55.92)	(50.72)
25	-82.33	-70.58	-76.66	-77.46	-29.02
	(65.86)	(67.43)	(66.07)	(67.23)	(67.22)
50	-100.64*	-81.33*	-84.46*	-91.32*	-33.06
	(73.62)	(66.69)	(68.03)	(68.80)	(61.18)
75	-144.62*	-117.54*	-121.90	-125.68*	-33.95
	(96.90)	(93.00)	(95.39)	(92.25)	(88.53)
90	-330.15**	-281.64*	-282.83*	-288.51*	-111.28
	(218.49)	(206.91)	(207.15)	(222.37)	(182.71)
Demographics[a]		*X*	*X*	*X*	*X*
Edu. & emp.[a]			*X*	*X*	*X*
Composition[a]				*X*	*X*
Resources[a]					*X*
Observations	1,062	1,062	1,062	1,062	1,062

Notes: Robust standard errors in parentheses.

**absolute value of 95% CI is positive. *absolute value of 90% CI is positive.

[a]See table 4.1 for definitions.

The impact of cash on the distribution of dietary quality, as measured by the HEI-2005, is mostly consistent across all specifications indicating the relative weak power of the controls to explain the unobserved proneness to consume a healthful diet. Overall, there appears to be a positive impact across the distribution although only significant above the 40^{th} quantiles.

4.6.1 Are good or bad foods driving the differences?

The HEI allows us to estimate the impact of cash on the distribution of adequacy and moderation scores of the HEI. The adequacy scores range from 0-60 correspond to the first 9 components (total fruit, whole fruit, total vegetables, dark green/orange vegetables and legumes, total grains, whole grains, milk, meats and beans and oils) and the moderation scores are constructed from the last three components (saturated fat, sodium and empty calories) for the last 40 points. Although quantiles in each of these

Table 4.5: Effect of receiving cash on the distribution of HEI-2005 scores for selected quantiles

Outcome: HEI-2005	QR	GQR-1	GQR-2	GQR-3	GQR-4
Quantiles					
10	0.33	-0.17	-0.09	-0.16	0.22
	(1.08)	(1.10)	(1.09)	(1.08)	(1.07)
25	0.38	-0.02	0.03	-0.03	0.23
	(0.80)	(0.81)	(0.78)	(0.79)	(0.75)
50	0.69*	0.53	0.55	0.55	0.64*
	(0.45)	(0.47)	(0.47)	(0.46)	(0.44)
75	1.08*	0.91*	0.95*	0.90*	0.96*
	(0.65)	(0.66)	(0.65)	(0.64)	(0.63)
90	0.57	0.38	0.40	0.36	0.34
	(0.99)	(0.98)	(0.98)	(0.96)	(1.00)
Demographics[a]		*X*	*X*	*X*	*X*
Edu. & emp.[a]			*X*	*X*	*X*
Composition[a]				*X*	*X*
Resources[a]					*X*
Observations	1,062	1,062	1,062	1,062	1,062

Notes: Robust standard errors in parentheses.
**absolute value of 95% CI is positive. *absolute value of 90% CI is positive.
[a]See table 4.1 for definitions.

distributions do not necessarily correspond to the quantiles in the total HEI scores, it still allows us to answer the question: are good or bad foods driving the differences in quality?

The difference in HEI scores across cash and vouchers is mainly driven by households scoring better moderation scores (figure 4.5). As a reminder, a higher moderation score means the household is consuming less “bad” components (i.e., sodium, saturated fat, and empty calories). Coupled with the finding on kilocalories, it is likely that households tend to purchase a food bundle that is less nutrient dense.

Figure 4.3: Cash Out Effect on the Distribution of Kilocalories

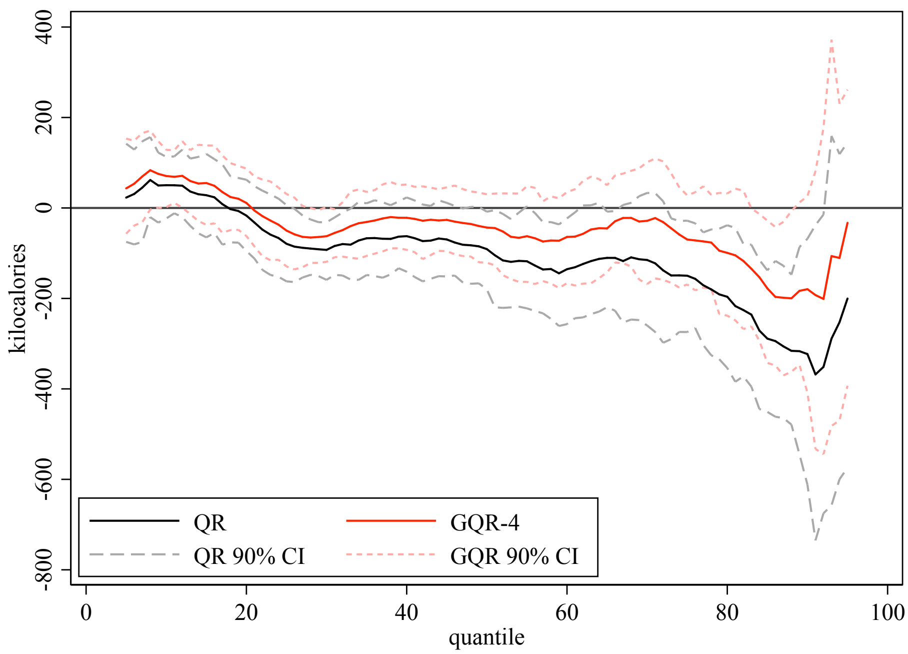


Figure 4.4: Cash Out Effect on the Distribution of Diet Quality

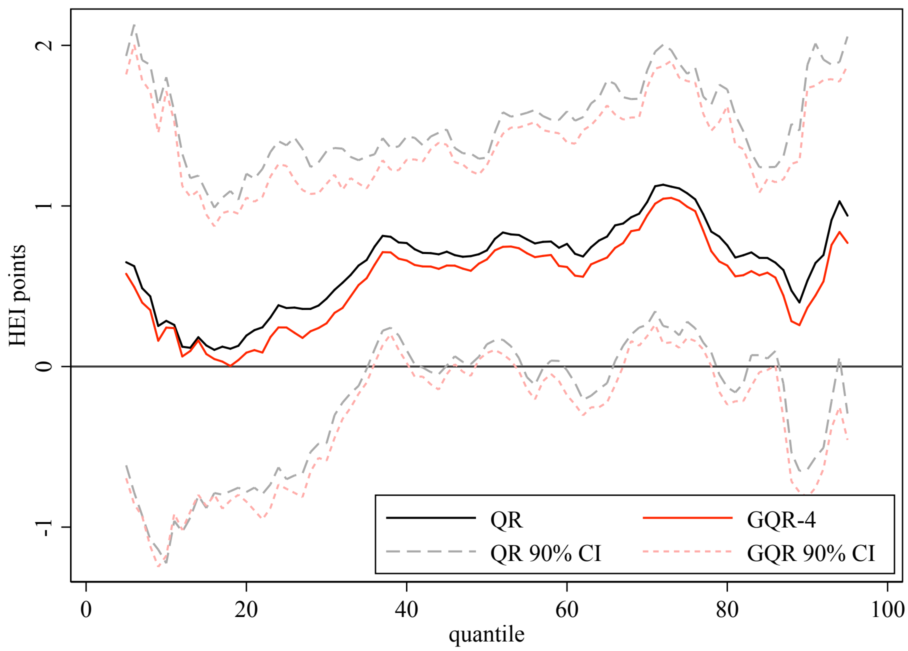
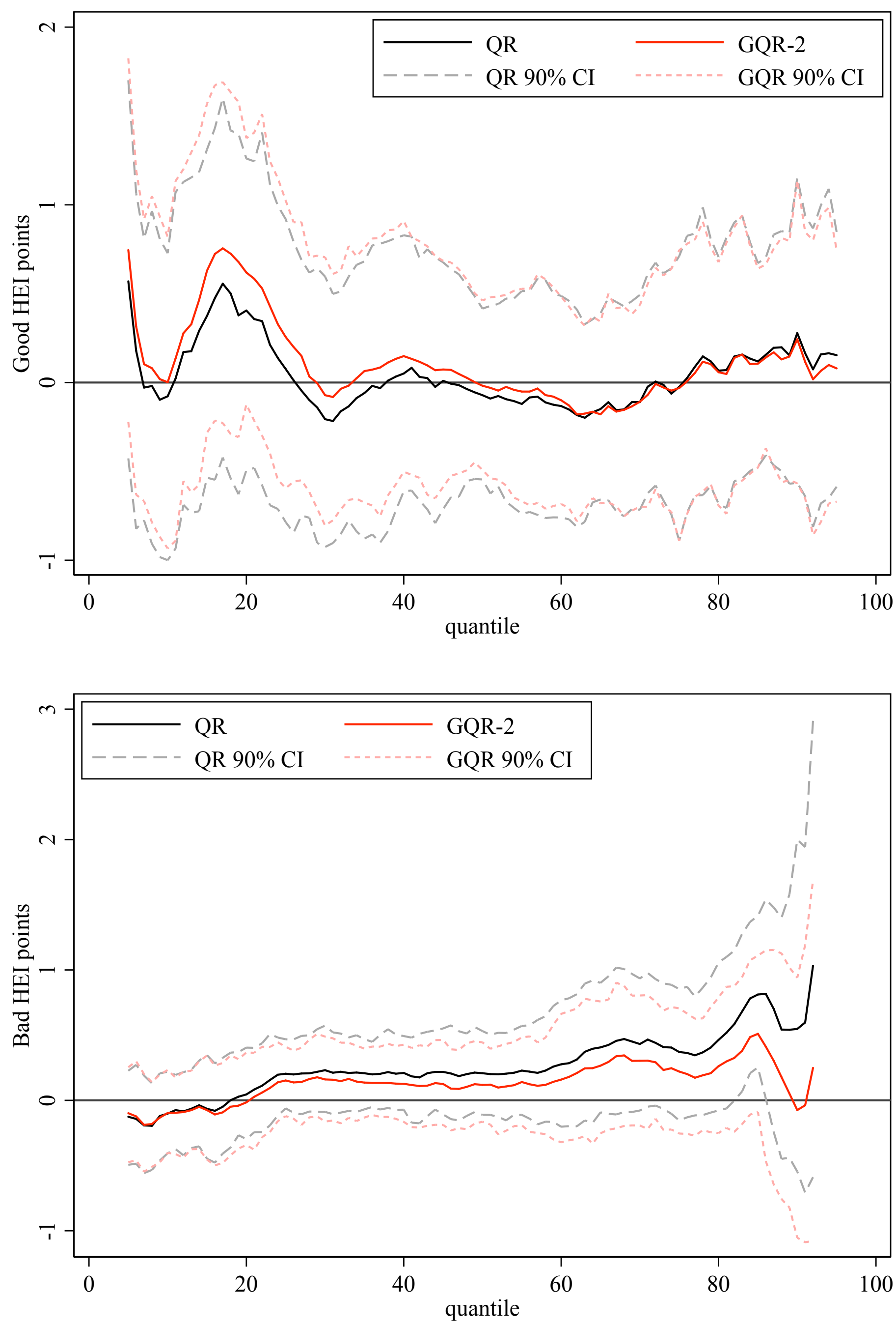


Figure 4.5: Cash Out Effect on the Distribution of Diet Quality by Type



4.7 Discussion and Conclusion

The United State's largest nutritional safety net, the Supplemental Nutrition Assistance Program (SNAP), aims to both alleviate food insecurity and increase the nutritional quality of household food supplies among beneficiaries. One mechanism the current system uses to better reach these goals is to provide benefits in the form of a food voucher, rather than in cash. The voucher system better reaches the goals of SNAP only to the extent that, (1) if benefits were paid in cash, then participants would consume less optimal quantities of food, and (2) cash benefits would be used to purchase lower quality food as compared to in-kind benefits. This study finds evidence to the contrary.

We find that households with a very high proneness for kilocalorie consumption tend to consume *less* calories when given cash rather a voucher. The portion of the kilocalorie distribution that exhibits this effect is well above any reasonable recommendation. Similarly, households with a low proneness for food consumption tend to purchase *more* food under a cash system, and these households are below dietary recommendations. We hypothesize these effects are due to behavioral mechanisms, namely mental accounting. Specifically, the padding of one mental account (i.e., the food budget) alters the types of foods purchased. For example, households that have a high preference for food (i.e., in the upper quantiles of the kilocalorie distribution) tend to buy even more food with a voucher.

When we examine the distribution of dietary quality, we again see more favorable results under a cash system over a voucher system. In the lower portion of the distribution, we find some positive effects of cash but these are insignificant. For quantiles above 40, the effect become larger and significant. Most of this improvement appears to be due households consuming less foods that are calorie-dense.

Chapter 5

Conclusion

In this dissertation I have explored the role of dietary quality in the United States as a means of (1) measuring well-being, (2) analyzing the effectiveness of school food programs and (3) showing how a shift from a voucher food assistance program to a cash transfer program would impact household food consumption.

In the first analysis, I found contrary to popular wisdom that adult dietary quality has been improving in the United States over the last twenty years. Even after accounting for changes in the demographic makeup (i.e., age, race/ethnicity and education) there is still a large a substantial unexplained improvement. Moreover, the role of food reformulation (e.g., the ingredients used in processed foods) has had a small but positive impact on the quality of diets in adult Americans. The latter finding opens the door for future investigation. Although I find small improvements, there is still substantial room for improvement.

In the second analysis, I examined how school food and away food impacts child dietary quality. Although it is almost universally accepted that away food has a negative impact on diet quality, the previous literature has been mixed on the impact of school food. I found that previous studies' focus on mean effects of school food programs have missed tail effects. In particular, children that exhibit low dietary quality at home tend to benefit from eating a school meal whereas those in the upper portion of the distribution do not. The implication is that school food programs not only have room for improvement but already have meaningful impacts on at least some children.

In the third analysis, I took a novel approach to an old question: how should food assistance benefits be paid? I began by focusing on the intended goals of the Supplemental Nutrition Assistance Program (SNAP) – to "alleviate hunger" and also to "permit low-income households to obtain a more nutritious diet." This approach is in contrast to previous studies' focus purely on food expenditures. My approach was able to uncover the impact of a cash system on both the quality and quantity of food supplies in the home. I found that a cash system tends to increase the quality of the diets for those in the upper portion of the distribution. For those in the lower portion of the HEI-2005 distribution, there appears to be positive but insignificant impacts. With regards to the quantity of the diet, one must acknowledge that from a social welfare stance, some households may need more calories (i.e., those that are food insecure) while others perhaps should consume less. With this fact in mind, I find that a cash system increases (decreases) the quantity of the diet for those well below (above) dietary recommendations for kilocalorie intake. Overall, these findings imply a positive impact of a cash system.

Overall, using dietary quality as an outcome measure has proved to be an effective tool. Several previously unknown facts have been uncovered: (1) the quality of the American diet is on the rise, (2) school food programs are effective for at least those that are prone to very low dietary quality at home, and (3) food assistance programs may better reach their intended goals if benefits were paid in cash. In all cases, taking a distributional approach has shown to uncover effects in the tails of the outcome distribution.

References

[1] Aguiar, M. and E. Hurst, (2005). "Consumption vs. Expenditure" *Journal of Political Economy*, (113:5), 919–948.

[2] Alkire, S., and J.E. Foster. 2011. "Counting and Multidimensional Poverty Measurement." *Journal of the Public Economics*, 95(7):476-487.

[3] Allison, R.A., and J.E. Foster. 2004. "Measuring Health Inequality using Qualitative Data." *Journal of Health Economics*, 23(3):505-524.

[4] Anderson, G. 1996. "Nonparametric Tests for Stochastic Dominance in Income Distributions." *Econometrica*, 64(5):1183-1193.

[5] Barrett, G.F., and S.G. Donald. 2003. "Consistent Tests for Stochastic Dominance." *Econometrica*, 71(1): 71-104.

[6] Beatty, T.K.M., L. Blow, T. Crossley and C. O'Dea. "Cash by any other name? Evidence on Labeling From the UK Winter Fuel Payment." *Journal of Public Economics*, forthcoming.

[7] Beatty, T.K.M., B-H Lin, and T.A. Smith, (2014). "Is Diet Quality Improving? Distributional Changes in United States, 1989–2008." *American Journal of Agricultural Economics* 96 (3): 769-789.

[8] Beatty, T.K.M. and C. Tuttle. (2012). "Expenditure response to increases in in-kind transfers: Evidence from the Supplemental Nutrition Assistance Program." Mimeo. 2014.

[9] Bennett, C. 2010. "On Bidirectional Tests for Stochastic Dominance." Presented at the Annual Meetings of the Midwest Econometrics Group, October 1-2.

[10] Bennett, C. 2013. "Inference for Dominance Relations." *International Economic Review*, 54(4):1309-1328.

[11] Benton, D., (2004). "Role of Parents in the Determination of the Food Preferences of Children and the Development of Obesity." *International Journal of Obesity* 28(): 858-869.

[12] Bhattacharya, D., (2005). "Asymptotic Inference from Multi-Stage Samples." *Journal of Econometrics* 126(1): 145-171

[13] Bhattacharya J., T. Deleire, S. Haider, and J. Currie. 2003. "Heat or Eat? Cold-Weather Shocks and Nutrition in Poor American Families." *American Journal of Public Health*, 93(7):1149-1154.

[14] Bhattacharya, J., J. Currie and S. Haider. 2004. "Poverty, Food Insecurity and Nutritional Outcomes in Children and Adults." *Journal of Health Economics*, 23(4):839-862.

[15] Bhattacharya, J., J. Currie and S. Haider. 2006. "Breakfast of Champions? The School Breakfast Program and the Nutrition of Children and Families." *Journal of Human Resources*, 41(3):445-466.

[16] Billingsley, P. 1968. *Convergence of Probability Measures*, New York: John Wiley & Sons.

[17] Birch, L.L., (1999). "Development of Food Preferences." *Annual Review of Nutrition* 19: 41-62.

[18] Bishop, J.A., J.P. Formby, and L.A. Zeager, (2000). "The Effect of Food Stamp Cashout on Undernutrition." *Economic Letters* 67():75-85.

[19] Bishop, J., J. Formby, and P. Thistle. 1989. "Statistical Inference, Income Distributions, and Social Welfare." *Research on Economic Inequality* 1:49-82.

[20] Bitler, M., and S.J. Haider. 2011. "An Economic View of Food Deserts in the United States." *Journal of Policy Analysis and Management*, 30(1):153-176.

[21] Bosire, C., M.J. Stampfer, A.F. Subar, Y. Park, S.I. Kirkpatrick, S.E. Chiuve, A.R. Hollenbeck, and J. Reedy. 2013. "Index-based Dietary Patterns and the Risk of Prostate Cancer in the NIH-AARP Diet and Health Study." *American Journal of Epidemiology* 177(6):504-513.

[22] Bowman, S.A., S. Gortmaker, C. Ebbeling, M. Pereira, D. Ludwig, (2004). "Effects of Fast-Food Consumption on Energy Intake and Diet Quality Among Children in a National Household Survey." *Pediatrics*, 113(1): 112–118.

[23] Breunig, R. and I. Dasgupta, (2002). "A Theoretical and Empirical Evaluation of the Functional Forms Used to Estimate the Food Expenditure Equation of Food Stamp Recipients: Comment." *American Journal of Agricultural Economics*, 84(4):1156-1160.

[24] Breunig, R. and I. Dasgupta, (2003). "Are People Ashamed of Paying with Food Stamps?" *Journal of Agricultural Economics* 54(2): 203-225.

[25] Breunig, R. and I. Dasgupta, (2005). "Do Intra-Household Effects Generate the Food Stamp Cash-out Puzzle?" *American Journal of Agricultural Economics*, 87(3):552-568.

[26] Canay, I.A., (2011). "A Simple Approach to Quantile Regression for Panel Data." *The Econometrics Journal* 14(3): 368–386.

[27] Carlson, A., M. Lino, S.A. Gerrior, and P. Basiotis. (2003). "The Low-Cost, Moderate-Cost, and Liberal Food Plans: 2003 Administrative Report." Center for Nutrition Policy and Promotion, U.S. Department of Agriculture, CNPP-13.

[28] Carlson, A., M. Lino, and T. Fungwe, (2007). "The Low-Cost, Moderate-Cost, and Liberal Food Plans." Center for Nutrition Policy and Promotion, U.S. Department of Agriculture, CNPP-20.

[29] Center for Nutrition Policy and Promotion (CNPP), (2008a). "Diet Quality of Americans in 1994-96 and 2001-02 as Measured by the Health Eating Index-2005." U.S. Department of Agriculture. CNPP-37.

[30] Center for Nutrition Policy and Promotion (CNPP), (2008b). "Diet Quality of Low-Income and Higher Income Americans in 2003-04 as Measured by the Health Eating Index-2005." U.S. Department of Agriculture. CNPP-42.

[31] Chernozhukov, V. and H. Hong, (2003). "An MCMC Approach to Classical Estimation." *Journal of Econometrics* 115(2): 293–346.

[32] Chernozhukov, V. and C. Hansen, (2005). "An IV Model of Quantile Treatment Effects." *Econometrica* 73(1): 245-261.

[33] Chernozhukov, V. and C. Hansen, (2006). "Instrumental Quantile Regression Inference for Structural and Treatment Models." *Journal of Econometrics* 132(2): 491-525.

[34] Chernozhukov, V. and C. Hansen, (2008). "Instrumental Variable Quantile Regression: A Robust Inference Approach." *Journal of Econometrics* 142(1): 379-398.

[35] Chiuve, S.E., T.T. Fung, E.B. Rimm, F.B. Hu, M.L. McCullough, M. Wang, M.J. Stampfer, and W.C. Willett. 2012. "Alternative Dietary Indices Both Strongly Predict Risk of Chronic Disease." *Journal of Nutrition* 142(6):1009-1018.

[36] Code of Federal Regulations, Federal Resister. (2013). "National School Lunch Program and School Breakfast Program: Nutrition Standards for All Foods Sold in School as Required by the Healthy, Hunger-Free Act of 2010, Interim Rule." Title 7, Parts 210 and 220, 78(125):39068-39120. `http://www.gpo.gov/fdsys/pkg/FR-2013-06-28/pdf/2013-15249.pdf`

[37] Crouter, S.E., K.G. Clowers, and D.R. Bassett Jr., (2005) "A Novel Method for using Accelerometer Data to Predict Energy Expenditure." *Journal of Applied Physiology* 100(4): 1324-1331.

[38] Cunha, F., J.J. Heckman, L. Lochner, D.V. Masterov, (2006). "Interpreting the Evidence on Life Cycle Skill Formation." in *Handbook on the Economics of Education* vol. 1, eds. E.A. Hanushek and F. Welch.

[39] Dardanoni V., and A. Forcian. 1999. "Inference for Lorenz Curve Orderings." *The Econometrics Journal* 2(1): 49-75.

[40] Davidson, R. and J.Y. Duclos. 2000. "Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality." *Econometrica*, 68(6):1435-1464.

[41] Deaton, A. 1997. *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* Baltimore: John Hopkins University Press.

[42] DiNardo, J.A., N.M. Fortin and T. Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach." *Econometrica*, 64(5):1001-1044.

[43] Dodd, K.W., P.M. Guenther, L.S. Freedman, A.F. Subar, V. Kipnis, D. Midthune, J.A. Tooze, and S.M. Krebs-Smith. 2006. "Statistical Methods for Estimating Usual Intake of Nutrients and Foods: A Review of the Theory." *Journal of the American Dietetic Association*, 106(10):1640-50.

[44] Doksum, K., (1974). "Empirical Probability Plots and Statistical Inference for Nonlinear Models in the Two-Sample Case." *The Annals of Statistics* 2(2): 267–277.

[45] Drewnowski, A., and A. Barratt-Fornell. 2004. "Do Healthier Diets Cost More?" *Nutrition Today*, 39(4):161-168.

[46] Duclos, J.Y., D.E. Sahn, and S.D. Younger. 2006. "Robust Multidimensional Poverty Comparisons." *Economic Journal*, 116(514): 943-968.

[47] Epstein, L.H., C.C. Gordy, H.A. Raynor, M. Beddome, C.K. Kilanowski, and R. Paluch. 2001. "Increasing Fruit and Vegetable Intake and Decreasing Fat and Sugar Intake in Families at Risk for Childhood Obesity." *Obesity Research*, 9(3):171-178.

[48] Epstein, L.H., R.A. Paluch, M.D. Beecher, and J.N. Roemmich. 2008. "Increasing Healthy Eating vs. Reducing High Energy-dense Foods To Treat Pediatric Obesity." *Obesity*, 61(2): 318-326

[49] Food and Nutrition Service-USDA. (2014). "Supplemental Nutrition Assistance Program Participation and Costs" `http://www.fns.usda.gov/sites/default/files/pd/SNAPsummary.pdf`

[50] Food, Conservation, and Energy Act of 2008

[51] Florence, M.D., M. Asbridge and P.J. Veugelers, (2008). "Diet Quality and Academic Performance." *Journal of School Health* 78(4): 209-215.

[52] Food and Nutrition Act of 2008, 7 U.S.C. Section 2011. 2008. available at `http://www.fns.usda.gov/snap/rules/Legislation/pdfs/PL_110-246.pdf`

[53] Food Nutrition Service-U.S. Department of Agriculture (FNS-USDA). 2012a. "National School Lunch Program Fact Sheet." `http://www.fns.usda.gov/cnd/lunch/aboutlunch/nslpfactsheet.pdf`

[54] Food Nutrition Service-U.S. Department of Agriculture (FNS-USDA). 2012b. "The School Breakfast Program Fact Sheet." `http://www.fns.usda.gov/cnd/breakfast/AboutBFast/SBPFactSheet.pdf`

[55] Ford, E.S., G. Zhao, J. Tsai and C. Li. 2011. "Low-Risk Lifestyle Behaviors and All-Cause Mortality: Findings from the National Health and Nutrition Examination III Mortality Study." *American Journal of Public Health*, 101(10):1922–1929.

[56] Fraker, T.M., (1990). *The Effects of Food Stamps on Food Consumption: A Review of the Literature*, FNS-USDA.

[57] Fraker, T.M., A.P. Martini, and J.C. Ohls, (1995). "The Effect of Food Stamp Cashout on Food Expenditures: An Assessment of the Findings from Four Demonstrations." *Journal of Human Resources* 30(4):633-649.

[58] Freedman, L.S., P.M. Guenther, S.M. Krebs-Smith, and P.S. Kott. 2008. "A Population's Mean Healthy Eating Index-2005 Scores are Best Estimated by the Score of

the Population Ratio when One 24-Hour Recall is Available." *Journal of Nutrition* 138(9):1725-1729.

[59] Freedman, L.S., P.M. Guenther, S.M. Krebs-Smith, K.W. Dodd,and D. Midthune. 2010. "A Population's Distribution of Healthy Eating Index-2005 Component Scores can be Best Estimated When More Than One 24-Hour Recall is Available." *Journal of Nutrition* 140(8):1529-1534.

[60] Galvao Jr., A.F., (2011). "Quantile Regression for Dynamic Panel Data with Fixed Effects." *Journal of Econometrics* 164(1): 142–157.

[61] Gastwirth and Nayak. 1999. "Comments on "Tests of Significance for Lorenz Partial Orders" by J.A. Bishop and J.P. Formby." in J. Silber, ed., *Handbook of Income Inequality Measurement*, chapter 11:336-339.

[62] Gleason, P. and C. Suitor, (2003). "Eating at School: How the National School Lunch Program affects Children's Diets." *American Journal of Agricultural Economics* 85(4): 1047-1061.

[63] Grainger, C., B. Senauer and C.F. Runge (2007). "Nutritional Improvements and Student Food Choices in a School Lunch Program." *Journal of Consumer Affairs* 41(2): 265-284.

[64] Gregory, C., T.A. Smith, and M. Wendt. 2011. "How Americans Rate their Diet Quality: An Increasingly Realistic Perspective." Economic Information Bulletin No. 83, ERS–USDA.

[65] Gregory, C., M. Ver Ploeg, M. Andrews, A. Coleman-Jensen, (2013). "Supplemental Nutrition Assistance Program (SNAP) Participation Leads to Modest Changes in Diet Quality" ERR-147, USDA, Economic Research Service.

[66] Guenther, P.M., J. Reedy, S.M. Krebs-Smith, B.B. Reeve, and P.P. Basiotis. 2007. "Development and Evaluation of the Healthy Eating Index-2005: Technical Report." Center for Nutrition Policy and Promotion, U.S. Department of Agriculture.

[67] Guenther P.M., J. Reedy, and S.M. Krebs-Smith. 2008. "Development of the Healthy Eating Index-2005." *Journal of the American Dietetic Association*, 108(11):1896-1901

[68] Guenther P.M., J. Reedy, S.M. Krebs-Smith, and B.B. Reeve. 2008. "Evaluation of the Healthy Eating Index-2005." *Journal of the American Dietetic Association*, 108(11):1854-1864

[69] Guo, X., B.A. Warden, S. Paeratakul and G.A. Bray. 2004. "Healthy Eating Index and Obesity." *European Journal of Clinical Nutrition*, 58:1580-1586.

[70] Hall KD, Guo J, Dore M, Chow CC, (2009). "The Progressive Increase of Food Waste in America and Its Environmental Impact." *PLoS ONE* 4(11): e7940.

[71] Hall, K.D., S.B. Heymsfield, J.W. Kemnitz, S. Klein, D.A. Schoeller, and J.R. Speakman. 2012. "Energy Balance and its Components: Implications for Body Weight Regulation." *American Journal of Clinical Nutrition*, 95(4):989-994.

[72] Harding, M., and C. Lamarche, (2009). "A Quantile Regression Approach for Estimating Panel Data Models using Instrumental Variables." *Economic Letters* 104(3): 133–135.

[73] Heckman, J.J. and D.V. Masterov, (2007). "The Productivity Argument for Investing in Young Children." *Review of Agricultural Economics* 29(3): 446-493.

[74] Heckman, J. and R. Robb. (1986). "Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes." in *Drawing Inference from Self-Selected Samples*, ed. by H. Wainer, New York: Springer-Verlag, 63-107.

[75] Hinrichs, P., (2010). "The Effects of the National School Lunch Program on Education and Health." *Journal of Policy Analysis and Management* 29(3): 479-505.

[76] Hoynes, H.W. and Schanzenbach, D.W., (2009). "Consumption responses to in-kind transfers: evidence from the introduction of the Food Stamp Program." *American Economic Journal–Applied Economics*, 1(4):109–139.

[77] Hunter, M., (2013). "Adaptive Markov chain Monte Carlo Sampling and Estimation in Mata." *The Stata Journal* `http://arrow.hunter.cuny.edu/research/papers/HunterEconWP440.pdf`

[78] Hymans, S.H. and H.T. Shapiro, (1976). "The allocation of Household Income to Food Consumption." *Journal of Econometrics* 4():167-188.

[79] Jemal, A., R. Siegal, E. Ward, Y. Hao, J. Xu, T. Murray, and M.J. Thun. 2008. "Cancer Statistics, 2008." *CA: A Cancer Journal for Clinicians*, 58(2):71-96.

[80] Jennings, A., A. Welch, E.M.F. van Sluijs, S.J. Griffin, and A.Cassidy, (2011) "Diet Quality is Independently Associated with Weight Status in Children Aged 9-10 Years." *Journal of Nutrition* 141(3): 453-459.

[81] Kant, A. 1996. "Indexes of Overall Diet Quality: A Review." *Journal of the American Dietetic Association*, 96(8):785-791.

[82] Kramer-LeBlanc, C.S., P.P. Basiotis and E.T. Kennedy. 1997. "Maintaining Food and Nutrition Security in the United States with Welfare Reform." *American Journal of Agricultural Economics*, 79(5):1600-1607.

[83] Kim S.Y., R.M. Nayga and O.J. Capps. 2001. "Food Label Use, Self-selectivity, and Diet Quality." *Journal of Consumer Affairs*, 35(2):346-363.

[84] Koenker, R., (2004). "Quantile Regression for Longitudinal Data." *Journal of Multivariate Analysis* 91(1):24–89.

[85] Koenker, R. and G. Bassett, (1978). "Regression Quantiles." *Econometrica* 46(1):33–50.

[86] Kolenikov, S. 2010. "Resampling Variance Estimation for Complex Survey Data." *The Stata Journal*, 10(2):165-199.

[87] Kourlaba, G., and D.B. Panagiotakos. 2009. "Dietary quality indices and human health: A review." *Maturitas*, 62(1):1-8.

[88] Levedahl, J.W., (1995). "A Theoretical and Empirical Evaluation of the Functional Forms Used to Estimate the Food Expenditure Equation of Food Stamp Recipients." *American Journal of Agricultural Economics*, 84(4):1161-1164.

[89] Lin, B-H and J. Guthrie, (2012). "Nutritional Quality of Food Prepared at Home and Away From Home, 1977-2008." EIB-105, USDA-Economic Research Service.

[90] Linton, O., E. Maasoumi, and Y.J. Whang. 2005. "Consistent Testing for Stochastic Dominance under General Sampling Schemes." *Review of Economic Studies* 72(3):735-765.

[91] Maasoumi, E. and Millimet. 2005. "Robust Inference Concerning Recent Trends in Environmental Quality." *Applied Econometrics*, 20(1):55-77.

[92] Madden, D. 2012. "A Profile of Obesity in Ireland: 2002–2007." *Journal of the Royal Statistical Society: Series A*, 175(4):893-914.

[93] Mancino, L., J. Todd., B-H Lin., (2009). "Separating What We Eat From Where: Measuring the Effect of Food Away From Home on Diet Quality," *Food Policy*, 34(6): 557-562.

[94] Mancino, L., J. Todd., J. Gutherie, B-H Lin., (2010). "How Food Away From Home Affects Children's Diet Quality." U.S. Department of Agriculture-Economic Research Service, ERR 104.

[95] Manski, C.F. and J.V. Pepper, (2000). "Monotone Instrumental Variables: With an Application to the Returns to Schooling." *Econometrica* 68(4): 997-1010

[96] McFadden, D. 1989. "Testing for Stochastic Dominance." in *Studies in the Economics of Uncertainty: In Honor of Josef Hadar*, ed. by T B. Fomby and T K. Seo. New York, Berlin, London, and Tokyo: Springer.

[97] Millimet, D.L., R. Tchernis, M Husain, (2010). "School Nutrition Programs and the Incidence of Childhood Obesity." *The Journal of Human Resources* 45(3): 640–654.

[98] Moffitt, R., (1989). "Estimating the Value of an In-Kind Transfer: The Case of Food Stamps." *Econometrica*, 57(2): 385-409.

[99] Morales M., D.K. Demory-Luce, T.A. Nicklas, T. Baranowski, (2002) "Consistency in food group consumption patterns from childhood to young adulthood: The Bogalusa Heart Study." Meeting of International Society of Behavioral Nutrition and Physical Activity, Seattle, WA.

[100] National Cancer Institute, (2013). "Comparing the HEI-2005 & HEI-2010" `http://riskfactor.cancer.gov/tools/hei/comparing.html`

[101] Nicklas, T.A., C.E. O'Neil, and V.L. Fulgoni III. 2012. "Diet Quality Is Inversely Related to Cardiovascular Risk Factors in Adults." *Journal of Nutrition* 142(12):2112-2118.

[102] Popkin, B.M., A.M. Siega-Riz, and P.S. Haines. 1996. "A Comparison of Dietary Trends among Racial and Socioeconomic Groups in the United States." *New England Journal of Medicine* 335:715-720.

[103] Ponomareva, M., (2010). "Quantile Regression for Panel Data Models with Fixed Effects and Small T: Identification and Estimation." Working paper, Northwestern Economics Department.

[104] Poti, J.M. and B.M. Popkin, (2011). "Trends in Energy Intake among US Children by Eating Location and Food Source, 1977-2006." *Journal of the American Dietetic Association* 111(8) 1156-1164.

[105] Powell, D., (2013). "A New Framework for Estimation of Quantile Treatment Effects: Nonseparable Disturbance in the Presence of Covariates." RAND Working Paper.

[106] Powell, D., (2014). "Did the Economic Stimulus Payments of 2008 Reduce Labor Supply? Evidence from Quantile Panel Data Estimation." RAND Working Paper.

[107] Powell, D., T.A. Smith, M. Baker, (2014). "Generalized Quantile Estimation in Stata." Stata Users Conference, Boston, MA.

[108] Powell, L.M., T.N. Nguyen, (2013). "Fast-Food and Full-Service Restaurant ConsumptionAmong Children and Adolescents." *Journal of the American Medical Association - Pediatrics* 167(1): 14-20.

[109] Ralston, K., C. Newman, A. Clauson, J. Guthrie, and J. Buzby, (2008). "The National School Lunch Program: Background, Trends, and Issues." ERR-61, USDA-Economic Research Service.

[110] Rank, M.R. and T.A. Hirschl, (2009). "Estimating the Risk of Food Stamp Use and Impoverishment During Childhood." *Archives of Pediatric and adolescent Medicine* 163(11): 994-999.

[111] Rao, J.N.K., C.F.J. Wu, and K. Yue. 1992. "Some Recent Work on Resampling Methods for Complex Surveys." *Survey Methodology* 18(2):209-217.

[112] Ravaillon, M. 1996. "Issues in Measuring and Modelling Poverty." *The Economic Journal*, 106(438):1328-1343.

[113] Reedy, J., P.N. Mitrou, S.M. Krebs-Smith, E. Wirfält, A. Flood, V. Kipnis, M. Leitzmann, T. Mouw, A. Hollenbeck, A. Schatzkin, and A.F. Subar. 2008. "Index-based Dietary Patterns and Risk of Colorectal Cancer: The NIH-AARP Diet and Health Study." *American Journal of Epidemiology* 168(1):38-48.

[114] Senauer, B., and N. Young, 1986. "The Impact of Food Stamps on Food Expenditures: Rejection of the Traditional Model." *American Journal of Agricultural Economics*, 68(1):37-43.

[115] Shahril, M.R., S. Sulaiman, S.H. Shaharudin, and S.N. Akmal. 2013. "Healthy Eating Index and Breast Cancer Risk among Malaysian Women." *European Journal of Cancer Prevention*, 22(4):342-347.

[116] Southworth, H.M. (1945). "The Economics of Public Measures to Subsidize Food Consumption." *Journal of Farm Economics*, 27(1):38-66.

[117] Strauss, J., and D. Thomas. 1998. "Health, Nutrition and Economics Development." *Journal of Economic Literature*, 36(2):766-817.

[118] Taber, D.R., J.F. Chriqui, L. Powell, F.J. Chaloupka, (2013). "Association Between State Laws Governing School Meal Nutrition Content and Student Weight Status." *Pediatrics*: doi:10.1001/jamapediatrics.2013.399

[119] Thaler, R. (1980). "Toward a Positive Theory of Consumer Choice." *Journal of Economic Behavior and Organization* 1(1):39-60.

[120] Thaler, R. (1985). "Mental Accounting and Consumer Choice." *Marketing Science* 4(3):199-214.

[121] Thaler, R. (1990). "Anomalies: Saving, Fungibility, and Mental Accounts." *The Journal of Economic Perspectives* 4(1):193-205.

[122] Thaler, R. (1999). "Mental Accounting Matters." *The Journal of Behavioral Decision Making* 12(3):183-206.

[123] U.S. Department of Health and Human Services, 2013. Poverty Guidelines, Research, and Measurement. Available at: http://aspe.hhs.gov/poverty/ (accessed 01-07-2013).

[124] U.S. Department of Agriculture, Agricultural Research Service, Beltsville Human Nutrition Research Center, Food Surveys Research Group (Beltsville, MD). Continuing Survey of Food Intakes by Individuals 1989-91 and Diet and Health Knowledge Survey 1989-91: `http://www.ars.usda.gov/SP2UserFiles/Place/12355000/pdf/csfii8991_documentation.pdf`

[125] Vesper, H.W., H.C. Kuiper, L.B. Mirel, C.L. Johnson, and J.L. Pirkle. 2012. "Levels of Plasma trans-Fatty Acids in Non-Hispanic White Adults in the United States in 2000 and 2009" *Journal of the American Medical Association*, 307(6):562-563.

[126] Whitmore, D. (2002). "What Are Food Stamps Worth?" Princeton University working paper no. 468.

[127] Wilde, P.E., P.E. McNamara and C.K. Ranney. 1999. "The Effect of Income and Food Programs on Dietary Quality: A Seemingly Unrelated Regression Analysis with Error Components." *American Journal of Agricultural Economics*, 91(4):959-971.

[128] Yin, G. (2009). "Bayesian Generalized Method of Moments." *Bayesian Analysis* 4(2):191-208.

Appendix A

Is Diet Quality Improving? Distributional Changes in the United States, 1989–2008

A.1 Calculating the HEI-2005 for the 1989–1991 CSFII

As mentioned in the text, the HEI-2005 is calculated by linking the MyPyramid Equivalents Databases (MPEDs) to food intake surveys via USDA food codes.[1] For the 1994-96 and 2005-08 surveys, recipe modification codes were included in addition to the food codes. These modification codes do not appear in the 1989-91 CSFII.[2] We drop all duplicate food codes that have modifications in the MPEDv1 and retain all unmodified foods. Previous research has shown that these modifications do not have significant impacts on nutrient intakes (Ahuja, Steinfeldt and Perloff, 1999).

Of the 3,953 unique foods reported by adults 20 and older on day one in CSFII 1989-91, 3,907 (98.8 percent) of these foods are in the MPEDv1. A total of 10,439 adults reported complete intakes on day one in 1989–91, and 941 of these adults (9 percent) reported consuming one of the 46 foods not found in MPEDv1. These individuals are dropped from the sample. Alternatively, we could have constructed 'best matches' for each of the 46 foods, as done by the National Cancer Institute when creating an equivalents database for NHANES III (1988-94) for the original HEI (National Cancer Institute website, accessed February, 2012). We tried this for the 46 foods, allowing us to keep all 10,439 adults; results were robust to this approach.

The validity of backdating the food codes has not previously been attempted, and we make no attempts here. However, we do note that the MPEDv2, originally created for 2003-04, was appended with some 800 foods to create the 2005-08 MPED. Thus, if mapping forward is appropriate then mapping backwards seems reasonable.

A.2 Calculating Reformulated Nutrient Values

The FNDDS codes nutrients that have been updated due to reformulation.[3] Reformulation can occur for a whole host of nutrients: water, calories, sodium, fats, etc. We code a food as reformulated if any of these nutrients we reformulated. The logic being

[1] MPEDs can be obtained from the ARS-USDA website `http://www.ars.usda.gov/Services/docs.htm?docid=17558`.

[2] For example, there are modification codes for the type of milk used in a scrambled egg in 1994-2008, whereas no distinction between milk is made in 1989-91.

[3] The FNDDS corresponding to NHANES 2001–08 can be found on the ARS-USDA website `http://www.ars.usda.gov/services/docs.htm?docid=12089`. We obtained a multi-year version dating back to 1994 by submitting a request to ARS-USDA.

that we do not know what or how the food was reformulated (e.g., a change in the type of oil used in vegetable shortening affects more than one nutrient). We replace values for calories, sodium, carbohydrates, alcohol, and saturated fat in the 1989-91 CSFII with the most recent updated value in FNDDS (i.e., if a food was reformulated in 2002 and then again in 2005, we use the 2005 value). As noted in table 1 in the text, these nutrients directly affect calculation of the HEI-2005. We then replace the MPED values in the 1989-91 CSFII with those from 2005-08 if the food was reformulated.

A.3 Results Using Two Days of Intake

In this section of the Online Appendix, we investigate the robustness of our analysis to using two days of dietary intake. Results are in tables A.1 and A.2. We exclude 2001-04 because only one day of intake was reported in 2001-02. It must be noted that the 1989-91 CSFII collected three days of *consecutive* intake data, whereas 1994-96 CSFII and 2005-2008 NHANES collected the second day of intake 3-10 days after day one. For the 1989-91 CSFII, we use the first two days of intake but note that the correlation between the two days could be much higher than in the later surveys. Finally, all day one intakes where reported during in-person interviews. Intakes for day two and three in 19891-91 were collected in a food diary, 1994-96 CSFII was collected via an in-person interview, and 2005-08 NHANES was collected via telephone.

A.4 Results of DFL Probability Model

Results of the DiNardo-Fortin-Lemieux decomposition can be found in table A.3.

A.5 Dominance Results for Counterfactuals and by Education

Dominance results for comparing counterfactual distributions are reporting in table A.4. Results by education level can be found in table A.5.

Table A.1: Mean Health Eating Index–2005 Scores, Two Days of Intake.

Population	1989-91	1994-96	2005-08
U.S. population	52.35 (13.29)[a]	53.16 (12.91)[a]	55.55 (11.72)
	[13.43, 91.72]	[16.65, 94.85]	[11.34, 97.15]
N	7,439	9,323	8,165
low-income	50.76 (18.48)[a]	51.49 (14.62)[a]	54.48 (14.09)
	[13.43, 89.83]	[16.65, 94.85]	[11.34, 97.15]
N	3,860	3,236	3,350
Higher-income	52.88 (10.73)[ab]	53.77 (12.17)[ab]	55.99 (10.61)[b]
	[15.64, 91.72]	[17.84, 93.03]	[16.78, 96.83]
N	3,579	6,087	4,815

Standard deviations in parenthesis. Maximum and minimum in brackets.

[a]Within-population mean is significantly lower than 2005-08 at 5-percent level.

[b]Within-year higher-income is significantly different from low-income at 5-percent level.

Table A.2: Tests of Stochastic Dominance among U.S. Adults, Two Days of Intake

Distribution		Bootstrap Tests				Two-stage		Inferred
A	B	$\hat{p}_1^-$	$\hat{p}_2^-$	$\hat{p}_1^+$	$\hat{p}_2^+$	$p_{1,0.1}^{2S}$	$p_{1,0.01}^{2S}$	Ranking
U.S. Population								
1989-91	1994-96	0.007	0.010	0.900	0.660	0.249	0.141	$A \prec_1 B^{***}$
	2005-08	0.002	0.000	1.000	0.877	1.000	1.000	$A \prec_1 B^{***}$
1994-96	2005-08	0.010	0.003	0.999	0.863	1.000	1.000	$A \prec_1 B^{***}$
Low-income								
1989-91	1994-96	0.218	0.257	0.570	0.654	0.303	0.182	ND
	2005-08	0.008	0.000	1.000	0.882	0.995	0.990	$A \prec_1 B^{***}$
1994-96	2005-08	0.006	0.003	0.991	0.855	0.976	0.957	$A \prec_1 B^{***}$
Higher-income								
1989-91	1994-96	0.007	0.006	0.880	0.684	0.338	0.210	$A \prec_1 B^{***}$
	2005-08	0.002	0.000	1.000	0.906	1.000	1.000	$A \prec_1 B^{***}$
1994-96	2005-08	0.007	0.010	0.991	0.886	0.999	0.998	$A \prec_1 B^{***}$

Notes: The $\hat{p}_s^{\pm}$ values refer to one-sided tests of the null hypothesis $H_s^{\pm}$ using equation (7). The asymptotic $p_{1,\alpha}^{2S}$ values are calculated from (8), where $\alpha = 0.1, 0.01$. The notation $A \prec_s B$ reads "Distribution B dominates distribution A at order s," while ND indicates no dominance at order 1 or 2. Inferred ranking is based on statistical significance levels of $^{***}1$, $^{**}5$, and $^{*}10\%$.

Table A.3: Results of DFL Probability Model

	$\Pr(t=08\|h)$	$\Pr(t=08\|e)$	$\Pr(t=08\|h,e)$
Male	-0.001		-0.001
	(0.020)		(0.020)
Non-Hispanic white, $30-44$	-0.090		-0.104
	(0.090)		(0.092)
Non-Hispanic white, $45-64$	0.307***		0.316***
	(0.080)		(0.084)
Non-Hispanic white, 65+	0.181		0.274**
	(0.092)		(0.096)
Non-Hispanic black, $20-29$	0.168		0.200
	(0.144)		(0.147)
Non-Hispanic black, $30-44$	0.100		0.118
	(0.149)		(0.151)
Non-Hispanic black, $45-64$	0.364*		0.426**
	(0.151)		(0.157)
Non-Hispanic black, 65+	0.045		0.227
	(0.143)		(0.156)
Hispanic, $20-29$	0.385*		0.522**
	(0.148)		(0.157)
Hispanic, $30-44$	0.390*		0.513**
	(0.160)		(0.171)
Hispanic, $45-64$	0.498**		0.658**
	(0.187)		(0.201)
Hispanic, 65+	0.349		0.610**
	(0.206)		(0.209)
Other race/ethnicity, $20-29$	0.386		0.348
	(0.203)		(0.206)
Other race/ethnicity, $30-44$	0.460*		0.442*
	(0.181)		(0.184)
Other race/ethnicity, $45-64$	0.942***		0.987***
	(0.158)		(0.168)
Other race/ethnicity, 65+	0.468		0.616*
	(0.302)		(0.307)
Did not attend high school		-0.353***	-0.568***
		(0.083)	(0.092)
High school, no college		-0.261***	-0.304***
		(0.054)	(0.060)
Constant	0.016	0.311***	0.153
	(0.088)	(0.043)	(0.095)
Obs.	18,634	18,634	18,634

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A.4: Tests of Stochastic Dominance under Reformulation and Counterfactuals

Distribution		Bootstrap Tests				Two-stage		Inferred
A	B	$\hat{p}_1^-$	$\hat{p}_2^-$	$\hat{p}_1^+$	$\hat{p}_2^+$	$p_{1,0.1}^{2S}$	$p_{1,0.01}^{2S}$	Ranking
U.S. Population								
1989-91$_R$	2005-08	0.007	0.000	1.000	0.883	0.981	0.966	$A \prec_1 B^{***}$
1989-91$_{R,E}$	2005-08	0.019	0.001	1.000	0.877	0.985	0.973	$A \prec_1 B^{**}$
1989-91$_{R,H}$	2005-08	0.043	0.004	1.000	0.865	0.981	0.966	$A \prec_1 B^{**}$
1989-91$_{R,H,E}$	2005-08	0.167	0.053	0.996	0.852	0.946	0.907	$A \prec_2 B^{*}$
Low-income								
1989-91$_R$	2005-08	0.048	0.006	1.000	0.881	0.998	0.996	$A \prec_1 B^{**}$
1989-91$_{R,E}$	2005-08	0.070	0.009	0.999	0.879	0.981	0.966	$A \prec_1 B^{*}$
1989-91$_{R,H}$	2005-08	0.124	0.048	0.999	0.893	0.998	0.996	$A \prec_2 B^{**}$
1989-91$_{R,H,E}$	2005-08	0.193	0.113	0.998	0.891	0.998	0.996	ND
Higher-income								
1989-91$_R$	2005-08	0.004	0.000	1.000	0.895	0.985	0.973	$A \prec_1 B^{***}$
1989-91$_{R,E}$	2005-08	0.008	0.001	1.000	0.900	0.998	0.996	$A \prec_1 B^{***}$
1989-91$_{R,H}$	2005-08	0.024	0.003	1.000	0.877	0.985	0.973	$A \prec_1 B^{**}$
1989-91$_{R,H,E}$	2005-08	0.092	0.041	0.988	0.869	0.881	0.809	$A \prec_1 B^{*}$

Notes: The $\hat{p}_s^{\pm}$ values refer to one-sided tests of the null hypothesis $H_s^{\pm}$ using equation (7). The asymptotic $p_{1,\alpha}^{2S}$ values are calculated from (8), where $\alpha = 0.1, 0.01$. The notation $A \prec_s B$ reads "Distribution B dominates distribution A at order s," while ND indicates no dominance at order 1 or 2. Inferred ranking is based on statistical significance levels of ***1, **5, and *10%. R=reformulation; H=(gender, age, race/ethinicty); E=education

Table A.5: Tests of Stochastic Dominance among U.S. Adults by Education

Distribution		Bootstrap Tests				Two-stage		Inferred
A	B	$\hat{p}_1^-$	$\hat{p}_2^-$	$\hat{p}_1^+$	$\hat{p}_2^+$	$p_{1,0.1}^{2S}$	$p_{1,0.01}^{2S}$	Ranking
High school graduate or less								
1989-91	1994-96	0.018	0.113	0.583	0.426	0.029	0.010	ND
	2001-04	0.103	0.067	0.981	0.978	0.900	0.837	$A \prec_2 B^*$
	2005-08	0.009	0.001	0.993	0.855	0.997	0.994	$A \prec_1 B^{***}$
1994-96	2001-04	0.132	0.197	0.465	0.938	0.010	0.003	ND
	2005-08	0.016	0.012	0.997	0.874	0.997	0.994	$A \prec_1 B^{**}$
2001-04	2005-08	0.151	0.059	0.964	0.787	0.959	0.929	$A \prec_2 B^*$
At least one year of college								
1989-91	1994-96	0.105	0.028	1.000	0.869	1.000	1.000	$A \prec_2 B^{**}$
	2001-04	0.154	0.043	1.000	0.888	0.999	0.998	$A \prec_2 B^{**}$
	2005-08	0.016	0.002	1.000	0.869	0.986	0.975	$A \prec_1 B^{**}$
1994-96	2001-04	0.629	0.472	0.258	0.456	0.000	0.000	ND
	2005-08	0.141	0.136	0.777	0.927	0.309	0.186	ND
2001-04	2005-08	0.169	0.085	0.958	0.782	0.818	0.721	$A \prec_2 B^*$

Notes: The $\hat{p}_s^{\pm}$ values refer to one-sided tests of the null hypothesis $H_s^{\pm}$ using equation (7). The asymptotic $p_{1,\alpha}^{2S}$ values are calculated from (8), where $\alpha = 0.1, 0.01$. The notation $A \prec_s B$ reads "Distribution B dominates distribution A at order s," while ND indicates no dominance at order 1 or 2. Inferred ranking is based on statistical significance levels of $^{***}1$, $^{**}5$, and $^{*}10\%$.

Appendix B

"Billions and Billions Served" Heterogeneous Effects of School Food and Away Food on Child Dietary Quality

B.1 HEI-2005 Standards for Scoring

Table B.1: Healthy Eating Index-2005 Standards for Scoring.

	Score						
Component	0		5	8	10		20
Total fruit	0	$\rightarrow$	$\geq$ 0.8 cup eq/1000 kcal				
Whole fruit	0	$\rightarrow$	$\geq$ 0.4 cup eq/1000 kcal				
Total vegetables	0	$\rightarrow$	$\geq$ 1.1 cup eq/1000 kcal				
Dark green/orange veg.	0	$\rightarrow$	$\geq$ 0.4 cup eq/1000 kcal				
Total grains	0	$\rightarrow$	$\geq$ 3.0 cup eq/1000 kcal				
Whole grains	0	$\rightarrow$	$\geq$ 1.5 cup eq/1000 kcal				
Milk	0	$\longrightarrow$			$\geq$ 1.3 cup eq/1000 kcal		
Meats and beans	0	$\longrightarrow$			$\geq$ 2.5 oz eq/1000 kcal		
Oils	0	$\longrightarrow$			$\geq$ 12 g/1000 kcal		
Saturated fat	$\geq$ 15	$\longrightarrow$		10 $\longrightarrow$	$\leq$ 7% of energy		
Sodium	$\geq$ 2.0	$\longrightarrow$		1.1 $\longrightarrow$	$\leq$ 0.7 g/1000 kcal		
Calories from SoFAAS[a]	$\geq$ 50	$\longrightarrow$					$\leq$ 20% of energy

Source: Recreated from Guenther et al. (2007).
[a]Solid Fat, Alcohol, and Added Sugar

B.2 Food Source Coding

In the descriptions that follow, bracketed numbers refer to the code found in the NHANES documentation. *Food at home (FAH)*: store [1], grown or caught by you or someone you know [19], and fish caught by you or someone you know [20]; *food from school (FFS)*: cafeteria at school [7]; *food away from home (FAFH)*: restaurant with waiter/waitress [2], restaurant fast food/pizza [3], bar/tavern/lounge [4], restaurant no additional information [5], cafeteria not at school [6], vending machine [14], common coffee pot or snack tray [15], from someone else/gift [16], mail order purchase [17], residential dining facility [18], sport, recreation, or entertainment facility [24], street vendor, vending truck [25], and fundraiser sales [26].

Contrary to other studies (e.g., Lin and Guthrie, 2013; Mancino et al., 2010), food from child care centers [8] is not included in FFS becuase this venue does not fall under the SBP or NSLP. "Other" food sources were also coded: Community food programs (family/adult day care center [9], soup kitchen/shelter/food pantry [10], Meals on Wheels [11], community food program - other [12], community program no additional information [13]), a catch-all other category (other, specify [91]) and unidentifiable responses (don't know [99]) made up a small proportion of total calorie intake. In preliminary analyses, I considered these items in a fourth "other" category and found them to be of relatively equal quality to FAFH. Therefore, these foods are considered to be FAFH for this research. Please note however, that point estimates are robust to having a fourth category. Computationally, moving from 3 to 4 categories is not trivial, as the curse of dimensionality becomes a formidable problem for inference as discussed in the Estimation section.

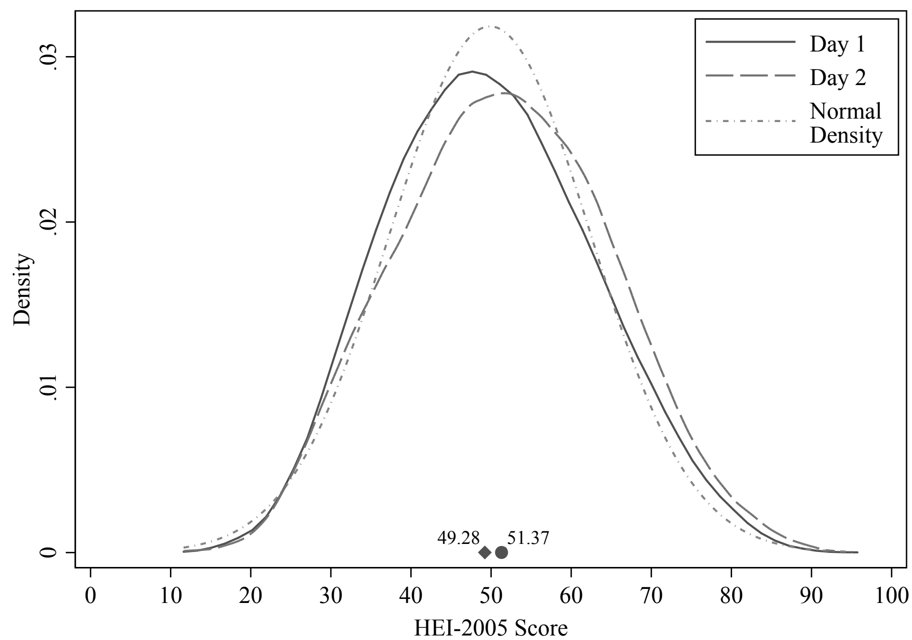
B.3 Kernel density

This section shows kernel density estimates of HEI scores for Day 1 and Day 2. In short, no child is observed scoring a perfect 100 nor the lowest possible score of 0. The densities are close to normally distributed as indicated by the overlay in figure B.1.

B.4 Adaptive Markov chain Monte Carlo (MCMC) Algorithm

The MCMC algorithm used in this paper is a variant of the Metropolis-Hastings algorithm with vanishing adaptation (see Hunter, 2013). I use $T = 2,500$ draws and discard (burn) the first 500. I use a two-step procedure suggested by Yin (2009) to construct the weighting matrix according to Bhattacharya (2005): use draws 251-500 during the burn-in period to construct the mean value $\hat{b}$ and then construct construct $W_n(\hat{b})$. Starting values for the parameters and variance matrix are obtained from equation (3.2), the standard quantile regression of Koenker and Bassett (1978). Using these starting values allow for a smaller burn-in window and quicker adaptation. The proposal distribution is a multivariate Normal density with a targeted acceptance rate of 0.4. Adaptation is achieved through a damping parameter $\delta \in (0, 1)$ which controls how quickly the tuning mechanism decays through $\rho_t = \frac{1}{(1+t)^\delta}$. Here, $\delta = 2/3$. Finally, the scaling parameter

Figure B.1: Kernel Density Estimates of HEI-2005 scores by Day

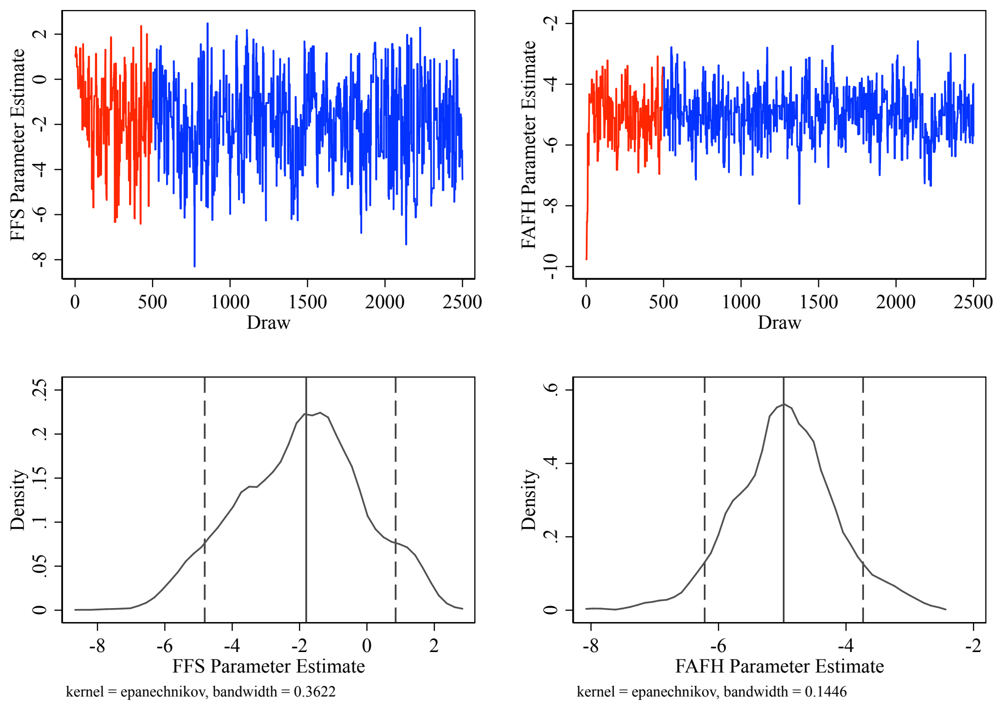


Note: Silverman's rule resulted in a bandwidth selection of approximately 3.5 for both days. The kernel is Epanechnikov. The diamond and circle indicate Day 1 and Day 2 medians, respectively.

is $\lambda = \frac{2.83^2}{d}$ where d is the number of parameters to be estimated.

The figure below reports the Markov chain sequence of draws $(\beta^{(1)}, \ldots, \beta^{(T)})$ for each parameter at $\tau = 0.5$. The kernel densities show the distribution of the posterior distribution from which point estimates (the mean) and inferences (the fifth and ninety-fifth quantiles) are drawn.

Figure B.2: Markov Chain Monte Carlo Performance at the Median



Note: The red lines in the top panels indicate the draws that are burned (discarded). The blue lines represent the draws used to draw inferences and determine the point estimates. The bottom panels plot the densities of the Markov chain sequence, where the solid vertical line is the mean and the dashed vertical lines are the 5th and the 95th percentiles of the quasi-posterior distribution.

B.4.1 Results for Selected Quantiles

Table B.2: Fixed Effects Quantile Regression Results

	FFS		FAFH	
Quantile	Estimate	95% CI	Estimate	95% CI
5	5.026	(1.535, 8.078)	2.278	(0.780, 3.765)
10	5.106	(3.627, 6.717)	0.867	(-0.999, 2.867)
15	2.729	(0.769, 4.665)	-0.009	(-1.208, 1.023)
20	0.028	(-2.520, 3.168)	-1.198	(-2.331, 0.240)
25	-0.474	(-3.242, 2.971)	-0.019	(-1.604, 1.185)
30	-0.068	(-3.336, 2.960)	-0.851	(-2.398, 0.491)
35	-0.054	(-3.032, 3.106)	-1.514	(-2.688, -0.160)
40	-0.672	(-4.251, 2.723)	-2.782	(-3.797, -1.745)
45	-1.888	(-4.542, 0.694)	-4.188	(-5.351, -3.083)
50	-1.803	(-4.816, 0.852)	-4.978	(-6.217, -3.732)
55	-0.736	(-2.424, 1.083)	-6.133	(-7.175, -4.950)
60	-0.350	(-3.007, 2.123)	-6.614	(-8.436, -4.915)
65	-0.016	(-2.421, 2.271)	-6.899	(-8.278, -5.344)
70	-1.140	(-4.063, 1.481)	-6.707	(-8.162, -5.096)
75	-0.591	(-3.624, 2.151)	-7.653	(-9.052, -6.340)
80	-1.949	(-4.667, 0.801)	-8.499	(-9.833, -6.861)
85	-3.248	(-6.888, 0.620)	-10.219	(-11.935, -8.473)
90	-4.407	(-7.619, -0.465)	-10.024	(-11.987, -8.312)
95	-3.077	(-5.789, -0.138)	-11.722	(-13.450, -10.099)

Note: Dependent variable the HEI-2005. Estimates are from equation (3.9) using Powell's (2014) Quantile Regression for Panel Data (QRPD).

B.4.2 Comparing QRPD to standard quantile regression

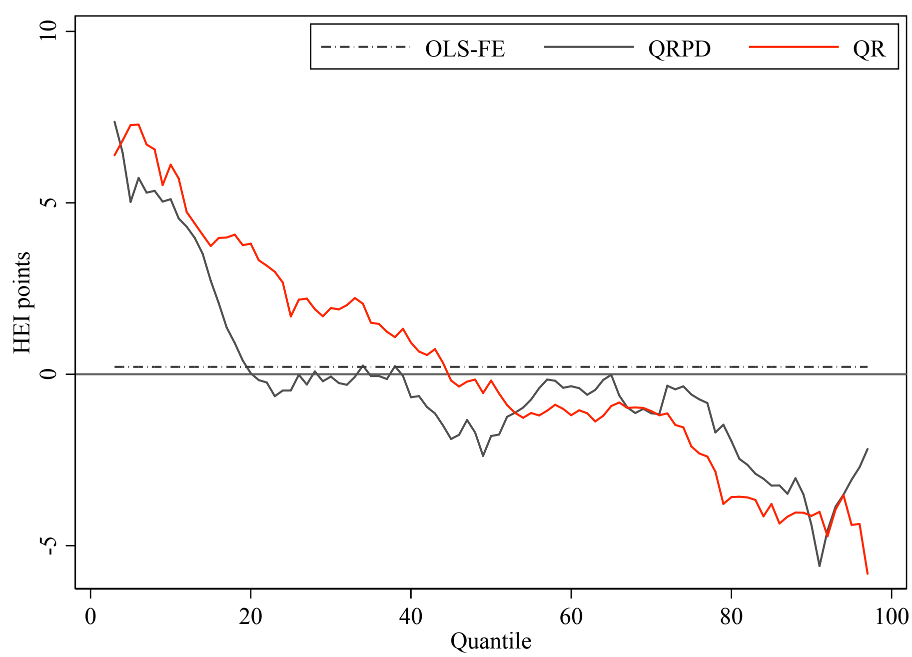
The QRPD estimator uses two moment conditions (equations (4.6) and (3.11)) to estimation the structural quantile function (SQF) found in equation (3.3). The standard quantile regression (QR) of Koenker and Bassett (1978) can also be used to estimate the same SQF using a single moment condition. To see this, define $\frac{1}{T}\sum_{s=1}^{T}\mathbf{1}(Y_{is} \leq D'_{is}b)$ in equation (4.6) as $\tau_i(b)$. This quantity is the fraction of the times individual i's outcome is below $D'_{is}b$. QR sets $\tau_i(b) = \tau$ for all individuals which ignores the fact that we observe each person multiple times. This is analogous to treating each person as a separate cross-sectional observation. The single (weighted) moment condition for QR is therefore

$$g_i(b) = w_i D_{it} \left[\mathbf{1}(Y_{it} \leq D'_{it} b) - \tau \right] \tag{B.1}$$

and estimation again proceeds in the GMM framework. Recall that $D \equiv (\gamma_1, \ldots, \gamma_T, X)$ where γ_t refers to the t^{th} day of intake on the h^{th} day of the week and $X = (FFS, FAFH)$ are the policy variables of interest.

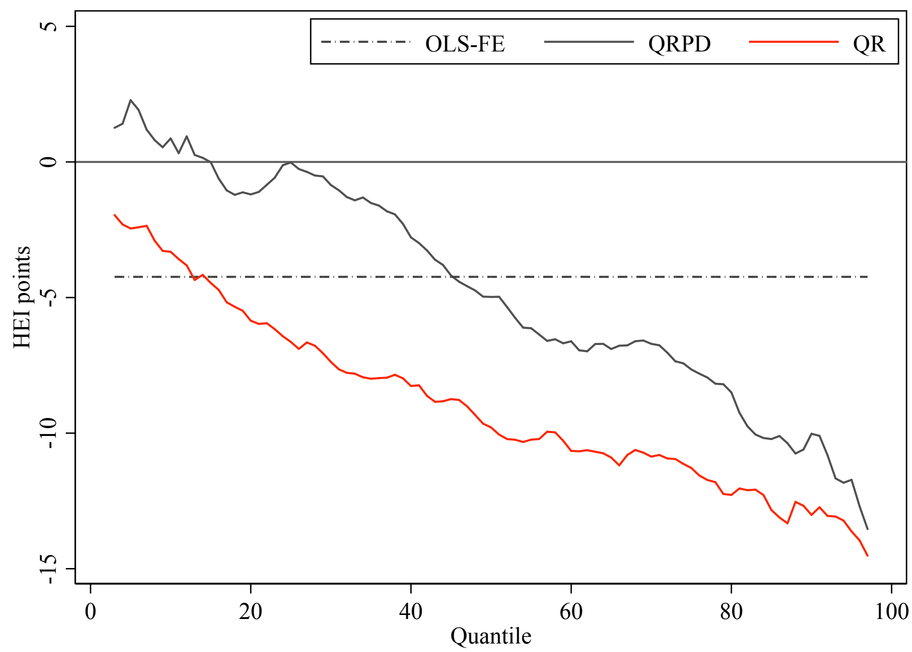
To show how using the additional information for each person affects results, I report estimates from the standard cross-sectional quantile regression (QR) in red in figures B.3 and B.4. The solid black lines are results from Powell's (2014) QRPD, which are replicated from the text. The dashed lines are individual fixed-effect OLS estimates from column (4) of table 3.3.

Figure B.3: Impact of Food From School (FFS) under Alternative Specifications



Note: Dietary quality is measured using the HEI-2005. OLS-FE are fixed effect estimates from column (4) of table 3.3. QR refers to the cross-sectional quantile regression using equation (B.1). QRPD is estimated from equation (3.9).

Figure B.4: Impact of Food Away from Home (FAFH) under Alternative Specifications



Note: Dietary quality is measured using the HEI-2005. OLS-FE are fixed effect estimates from column (4) of table 3.3. QR refers to the cross-sectional quantile regression in equation (B.1). QRPD is estimated from equation (3.9).

Appendix C

Cashing Out SNAP: Heterogeneous Impacts on Dietary Quantity and Quality

C.1 HEI-2005 Standards for Scoring

Table C.1: Healthy Eating Index-2005 Standards for Scoring.

	Score				
Component	0	5	8	10	20
Total fruit	0	$\rightarrow$ $\geq$ 0.8 cup eq/1000 kcal			
Whole fruit	0	$\rightarrow$ $\geq$ 0.4 cup eq/1000 kcal			
Total vegetables	0	$\rightarrow$ $\geq$ 1.1 cup eq/1000 kcal			
Dark green/orange veg.	0	$\rightarrow$ $\geq$ 0.4 cup eq/1000 kcal			
Total grains	0	$\rightarrow$ $\geq$ 3.0 cup eq/1000 kcal			
Whole grains	0	$\rightarrow$ $\geq$ 1.5 cup eq/1000 kcal			
Milk	0	$\longrightarrow$		$\geq$ 1.3 cup eq/1000 kcal	
Meats and beans	0	$\longrightarrow$		$\geq$ 2.5 oz eq/1000 kcal	
Oils	0	$\longrightarrow$		$\geq$ 12 g/1000 kcal	
Saturated fat	$\geq$ 15	$\longrightarrow$	10 $\longrightarrow$	$\leq$ 7% of energy	
Sodium	$\geq$ 2.0	$\longrightarrow$	1.1 $\longrightarrow$	$\leq$ 0.7 g/1000 kcal	
Calories from SoFAAS[a]	$\geq$ 50	$\longrightarrow$			$\leq$ 20% of energy

Source: Recreated from Guenther et al. (2007).
[a]Solid Fat, Alcohol, and Added Sugar

C.2 Calculating Food Pattern Equivalents in the San Diego data

This section describes our methodology for constructing food pattern equivalents for the 32 food groups reported in the San Diego data. Most mappings are self-explanatory and have a direct correspondence. The mapping of grains and meats warrants some discussion.

High-fiber grains were defined based on the definitions found in (Carlson et al., 2003, 2007). In 1990, high-fiber cereals corresponded to 0.8 grams of fiber per ounce, and high-fiber breads corresponded to 1.2 grams of fiber per ounce; no threshold was given for the flour, meal, rice and pasta group. (Carlson et al., 2003). In Carlson et al. (2007), the definition of high-fiber grains for all groups was defined as "more than 50% of ounce equivalents from whole grain." Therefore, we follow Carlson et al., (2003) when defining high-fiber breakfast cereals and breads since these definitions most closely resemble those used in the 1990. We follow Carlson et al., (2007) when defining high-fiber flour, meal, rice and pasta group since no definition of high-fiber was given in either the San Diego data nor Carlson et al., (2003). We note that the distribution of whole grains in the flour, meal, rice and pasta group is extremely bi-modal, either containing 0% whole grains or 75% or more.

Low-costs meats are those in the bottom 33% of the price distribution (Carlson et al., 2003, 2007). We have price data from USDA's successor intake surveys, the 2001-02 and 2003-04 National Health and Nutrition Examination Surveys (NHANES), but not specifically for the 1989-91 CSFII. However, both NHANES datasets roughly yield the same low-cost meats, which seems reasonable if relative prices remain stable. In 2001-02 (2003-04), the real price threshold for low-cost meats was \$3.01/lb (\$3.55/lb). Using the CPI for meats from the BLS, these prices correspond to a low-cost threshold of about \$2.50/lb in 1990. In general, low-cost meats include (in roughly increasing price) wild game, ground and patty meats, canned meats, corned beef, beef brisket, pork and beef roasts, smoked/cured hams, barbecued short-ribs, and broiled/baked pork cutlets.

Table C.2: Mapping of San Diego food groups to CSFII food groups

San Diego food groups	CSFII food groups (codes)[1,2]
Potatoes	white potatoes & Puerto Rican starchy vegetables (71)
High nutrient vegs	dark-green (72) and deep-yellow (73) vegetables
Other vegetables	tomatoes (74) and other vegetables (75)
Mixtures, mostly vegs; condiments	vegetable baby food (76), vegetable with meat (77)
Vitamin C-rich fruit	citrus fruits, juices (61)
Other fruit	dried fruits (62), other fruits (63), juices not citrus (64), fruit baby food (67)
High-fiber breakfast cereals[3]	cereals, not cooked (57)
Other breakfast cereals[3]	cereals, not cooked (57)
High-fiber flour, meal, rice, pasta[4]	flour and dry mixes (50), pastas, cooked cereals, rice (56)
Other flour, meal, rice, pasta[4]	flour and dry mixes (50), pastas, cooked cereals, rice (56)
High-fiber bread[3]	yeast breads, rolls (51), quick breads (52)
Other bread[3]	yeast breads, rolls (51), quick breads (52)
Bakery products, not bread	cakes, cookies, pies, pastries (53), crackers and salty snacks from grain products (54), pancakes, waffles, french toast, other (55)
Grain mixtures	grain mixtures, frozen plate meals and soups mainly grain (58)
Milk, yogurt	Milk and milk drinks - includes yogurt (11)
Cheese	cheeses (14)
Cream; mixtures mostly milk	creams and cream sub. (12), milk desserts, sauces, gravies (13)
Low-cost red meat, variety meat[5]	meat (20), beef (21), pork (22, excludes bacon (226)), lamb, veal, game, other carcass meat (23)
High-cost red meat, variety meats[5]	meat (20), beef (21), pork (22, excludes bacon(226)), lamb, veal, game, other carcass meat (23)
Poultry	poultry (24)
Fish, shellfish	fish and shellfish (26)
Bacon, sausage, lunch meats	bacon (226), organ meats, sausages, lunchmeats, spreads (25)
Eggs	eggs (31), egg mixtures (32), egg sub. (33), egg baby food (34), frozen plate meals with egg as major ingredient (35)
Dry beans, peas, lentils	legumes (41), seeds & mixtures (43), carob products (44)
Mixtures, mostly meat etc.	meat, poultry, fish with nonmeat items (27), frozen plate meals, soups, gravies with meat, poultry, fish gelatins (28)
Nuts, peanut butter	nuts, nut butters, and nut mixtures (42)
Fats, oils	fats (81), oils (82), salad dressings (83)
Sugar, sweets	sugars and sweets (91)
Seasonings	no nutritional value
Soft drinks, punches, ades	soft drinks (924), fruitaids (925), nonfruit bev. (926), powdered mix bev. (927), nonalcoholic bev. (928), bev. concentrates (929)
Coffee, tea	coffee (921), coffee sub. (922), tea (923)
Alcohol	alcoholic beverages (93)

[1]See USDA (1991, page 97–110) for details of food groups and coding.
[2]Food codes in CSFII are 8 digits. Numbers in this table represent the first two (or three) digits of these codes.
[3]Breakfast cereals and breads are defined as high-fiber according to Carlson et al. (2003).
[4]Flour, meal, rice and pasta are defined as high-fiber according to Carlson et al. (2007).
[5]Low- and high-cost meats are based on CNPP price database and definitions used in Carlson et al. (2003).

Table C.3: OLS Results at mean of kilocalories

	(2)	(3)	(4)	(5)
cash	-92.064	-91.755	-107.969	-40.248
	(66.957)	(67.068)	(67.089)	(54.223)
white_nh	87.135	65.593	27.737	178.266*
	(110.068)	(115.566)	(116.311)	(94.332)
black_nh	353.472***	334.928***	334.757***	433.077***
	(120.752)	(124.805)	(124.640)	(100.639)
hispanic	163.487	158.957	155.332	180.401**
	(110.905)	(111.391)	(111.467)	(90.070)
age_head	4.007	3.803	5.423	0.335
	(3.596)	(3.611)	(4.129)	(3.337)
fem_head	123.009	120.201	139.523	280.659***
	(113.957)	(114.543)	(116.576)	(94.227)
married	-303.752***	-299.437***	-194.782**	-184.725**
	(83.001)	(83.741)	(93.186)	(75.463)
less_hs		-57.422	-24.687	40.686
		(83.787)	(84.367)	(68.695)
hs_grad		-38.035	-34.597	45.383
		(90.338)	(90.276)	(73.072)
working		-17.668	24.726	56.513
		(84.246)	(86.400)	(70.817)
infant			74.251	146.455*
			(90.618)	(76.193)
toddler			65.123	122.851*
			(90.788)	(73.346)
adol			95.158	71.841
			(91.879)	(74.144)
teen			-9.797	-3.644
			(93.229)	(75.230)
single			-3.911	89.517
			(81.328)	(65.943)
lnsize			-348.060***	-980.899***
			(129.164)	(109.903)
amtpaid				16.710***
				(0.772)
amtprod_gift				20.605***
				(2.897)
amtwic				10.600*
				(6.048)
amtaway				1.763
				(1.614)
nonfood_exp				-0.256**
				(0.102)
_cons	2214.279***	2271.766***	2480.040***	1922.301***
	(195.593)	(211.894)	(242.139)	(199.770)
Observations	1062	1062	1062	1062
R-squared	0.035	0.035	0.047	0.383

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.4: OLS Results at mean of HEI-2005

	(2)	(3)	(4)	(5)
cash	0.280	0.305	0.306	0.412
	(0.455)	(0.453)	(0.453)	(0.449)
white_nh	-1.226	-1.865**	-1.895**	-1.562**
	(0.747)	(0.781)	(0.785)	(0.781)
black_nh	-5.754***	-6.310***	-6.362***	-6.148***
	(0.820)	(0.843)	(0.842)	(0.834)
hispanic	-0.424	-0.541	-0.675	-0.747
	(0.753)	(0.753)	(0.753)	(0.746)
age_head	0.036	0.030	0.061**	0.054*
	(0.024)	(0.024)	(0.028)	(0.028)
fem_head	1.105	1.107	0.930	1.127
	(0.774)	(0.774)	(0.787)	(0.780)
married	1.416**	1.488***	1.436**	1.284**
	(0.564)	(0.566)	(0.629)	(0.625)
less_hs		-1.794***	-1.828***	-1.929***
		(0.566)	(0.570)	(0.569)
hs_grad		-1.307**	-1.242**	-1.180*
		(0.610)	(0.610)	(0.605)
working		0.194	0.303	0.570
		(0.569)	(0.583)	(0.587)
infant			1.715***	1.629***
			(0.612)	(0.631)
toddler			-0.078	0.000
			(0.613)	(0.607)
adol			-0.079	-0.126
			(0.620)	(0.614)
teen			-0.545	-0.528
			(0.630)	(0.623)
single			0.698	0.757
			(0.549)	(0.546)
lnsize			0.723	-0.087
			(0.872)	(0.910)
amtpaid				0.031***
				(0.006)
amtprod_gift				-0.005
				(0.024)
amtwic				0.088*
				(0.050)
amtaway				-0.030**
				(0.013)
nonfood_exp				-0.001
				(0.001)
_cons	49.004***	50.570***	48.359***	47.937***
	(1.328)	(1.432)	(1.635)	(1.655)
Observations	1062	1062	1062	1062
R-squared	0.089	0.098	0.110	0.134

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.5: Proneness results at 10^{th} quantile of kilocalories

	GQR-1	GQR-2	GQR-3	GQR-4
white_nh	-0.000	-0.009	-0.005	-0.019
	(0.030)	(0.032)	(0.032)	(0.031)
black_nh	-0.011	-0.020	-0.021	-0.021
	(0.033)	(0.034)	(0.035)	(0.033)
hispanic	-0.012	-0.015	-0.012	-0.010
	(0.031)	(0.031)	(0.031)	(0.030)
age_head	0.000	0.000	-0.001	-0.000
	(0.001)	(0.001)	(0.001)	(0.001)
fem_head	-0.057*	-0.057*	-0.058*	-0.073**
	(0.032)	(0.032)	(0.032)	(0.031)
married	0.007	0.010	-0.007	0.003
	(0.023)	(0.023)	(0.026)	(0.025)
less_hs		-0.008	-0.010	-0.021
		(0.023)	(0.023)	(0.023)
hs_grad		0.040	0.039	0.032
		(0.025)	(0.025)	(0.024)
working		0.009	-0.004	0.009
		(0.023)	(0.024)	(0.023)
infant			-0.059**	-0.056**
			(0.025)	(0.025)
toddler			-0.022	-0.023
			(0.025)	(0.024)
adol			-0.022	-0.021
			(0.025)	(0.025)
teen			-0.003	0.005
			(0.026)	(0.025)
single			-0.001	-0.021
			(0.023)	(0.022)
lnsize			0.052	0.126***
			(0.036)	(0.036)
amtpaid				-0.002***
				(0.000)
amtprod_gift				-0.002*
				(0.001)
amtwic				-0.004*
				(0.002)
amtaway				-0.000
				(0.001)
nonfood_exp				-0.000
				(0.000)
_cons	0.151***	0.148**	0.158**	0.254***
	(0.053)	(0.058)	(0.066)	(0.065)
Observations	1062	1062	1062	1062
R-squared	0.005	0.009	0.016	0.091

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.6: Proneness results at 50^{th} quantile of kilocalories

	GQR-1	GQR-2	GQR-3	GQR-4
white_nh	-0.076	-0.083	-0.064	-0.104**
	(0.050)	(0.052)	(0.053)	(0.048)
black_nh	-0.145***	-0.151***	-0.151***	-0.167***
	(0.055)	(0.057)	(0.056)	(0.051)
hispanic	-0.092*	-0.095*	-0.090*	-0.083*
	(0.050)	(0.051)	(0.050)	(0.046)
age_head	-0.002	-0.002	-0.004**	-0.001
	(0.002)	(0.002)	(0.002)	(0.002)
fem_head	-0.121**	-0.124**	-0.135**	-0.167***
	(0.052)	(0.052)	(0.053)	(0.048)
married	0.105***	0.110***	0.062	0.071*
	(0.038)	(0.038)	(0.042)	(0.038)
less_hs		0.001	-0.010	-0.043
		(0.038)	(0.038)	(0.035)
hs_grad		0.044	0.031	0.016
		(0.041)	(0.041)	(0.037)
working		-0.027	-0.040	-0.062*
		(0.038)	(0.039)	(0.036)
infant			-0.087**	-0.088**
			(0.041)	(0.039)
toddler			-0.002	-0.019
			(0.041)	(0.037)
adol			-0.073*	-0.065*
			(0.042)	(0.038)
teen			0.044	0.048
			(0.042)	(0.038)
single			-0.016	-0.044
			(0.037)	(0.034)
lnsize			0.133**	0.352***
			(0.058)	(0.056)
amtpaid				-0.005***
				(0.000)
amtprod_gift				-0.005***
				(0.001)
amtwic				-0.006*
				(0.003)
amtaway				-0.001
				(0.001)
nonfood_exp				0.000*
				(0.000)
_cons	0.726***	0.727***	0.726***	0.818***
	(0.088)	(0.095)	(0.109)	(0.101)
Observations	1062	1062	1062	1062
R-squared	0.029	0.031	0.047	0.221

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.7: Proneness results at 90^{th} quantile of kilocalories

	GQR-1	GQR-2	GQR-3	GQR-4
white_nh	-0.013	-0.006	-0.001	-0.013
	(0.030)	(0.032)	(0.032)	(0.029)
black_nh	-0.064*	-0.058*	-0.055	-0.074**
	(0.033)	(0.034)	(0.034)	(0.031)
hispanic	-0.029	-0.027	-0.028	-0.029
	(0.031)	(0.031)	(0.031)	(0.028)
age_head	0.000	0.000	0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
fem_head	0.014	0.012	-0.002	-0.030
	(0.031)	(0.032)	(0.032)	(0.029)
married	0.046**	0.045*	0.022	0.018
	(0.023)	(0.023)	(0.026)	(0.023)
less_hs		0.008	-0.007	-0.000
		(0.023)	(0.023)	(0.021)
hs_grad		-0.023	-0.025	-0.026
		(0.025)	(0.025)	(0.023)
working		-0.020	-0.026	-0.024
		(0.023)	(0.024)	(0.022)
infant			0.014	0.002
			(0.025)	(0.024)
toddler			0.001	-0.005
			(0.025)	(0.023)
adol			-0.011	-0.004
			(0.025)	(0.023)
teen			-0.006	-0.006
			(0.026)	(0.023)
single			0.021	0.010
			(0.022)	(0.021)
lnsize			0.092***	0.203***
			(0.036)	(0.034)
amtpaid				-0.003***
				(0.000)
amtprod_gift				-0.005***
				(0.001)
amtwic				-0.002
				(0.002)
amtaway				-0.001
				(0.001)
nonfood_exp				0.000**
				(0.000)
_cons	0.898***	0.901***	0.801***	0.873***
	(0.053)	(0.058)	(0.066)	(0.062)
Observations	1062	1062	1062	1062
R-squared	0.010	0.012	0.026	0.194

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.8: Proneness results at 10^{th} quantile of HEI-2005

	GQR-1	GQR-2	GQR-3	GQR-4
white_nh	-0.014	0.016	0.016	0.007
	(0.030)	(0.031)	(0.031)	(0.031)
black_nh	0.177***	0.203***	0.203***	0.195***
	(0.032)	(0.033)	(0.033)	(0.033)
hispanic	-0.024	-0.018	-0.013	-0.016
	(0.030)	(0.030)	(0.030)	(0.030)
age_head	-0.001	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
fem_head	-0.019	-0.017	-0.014	-0.021
	(0.031)	(0.031)	(0.031)	(0.031)
married	-0.032	-0.038*	-0.019	-0.020
	(0.022)	(0.022)	(0.025)	(0.025)
less_hs		0.075***	0.073***	0.071***
		(0.022)	(0.022)	(0.023)
hs_grad		0.035	0.031	0.031
		(0.024)	(0.024)	(0.024)
working		0.007	0.015	-0.000
		(0.022)	(0.023)	(0.023)
infant			-0.036	-0.035
			(0.024)	(0.025)
toddler			0.046*	0.038
			(0.024)	(0.024)
adol			-0.026	-0.032
			(0.024)	(0.024)
teen			0.062**	0.057**
			(0.025)	(0.025)
single			-0.008	-0.014
			(0.022)	(0.022)
lnsize			-0.062*	-0.019
			(0.034)	(0.036)
amtpaid				-0.001***
				(0.000)
amtprod_gift				-0.001
				(0.001)
amtwic				-0.002
				(0.002)
amtaway				0.000
				(0.001)
nonfood_exp				0.000
				(0.000)
_cons	0.119**	0.052	0.115*	0.133**
	(0.052)	(0.056)	(0.064)	(0.065)
Observations	1062	1062	1062	1062
R-squared	0.072	0.081	0.100	0.115

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.9: Proneness results at 50^{th} quantile of HEI-2005

	GQR-1	GQR-2	GQR-3	GQR-4
white_nh	0.060	0.076	0.083	0.065
	(0.049)	(0.052)	(0.052)	(0.052)
black_nh	0.265***	0.279***	0.283***	0.274***
	(0.054)	(0.056)	(0.056)	(0.056)
hispanic	0.000	0.002	0.013	0.019
	(0.050)	(0.050)	(0.050)	(0.050)
age_head	-0.003*	-0.003	-0.005***	-0.005***
	(0.002)	(0.002)	(0.002)	(0.002)
fem_head	-0.057	-0.063	-0.061	-0.075
	(0.051)	(0.051)	(0.052)	(0.052)
married	-0.068*	-0.066*	-0.069*	-0.064
	(0.037)	(0.038)	(0.042)	(0.042)
less_hs		0.052	0.053	0.060
		(0.038)	(0.038)	(0.038)
hs_grad		0.052	0.051	0.049
		(0.040)	(0.040)	(0.040)
working		-0.054	-0.065*	-0.085**
		(0.038)	(0.039)	(0.039)
infant			-0.134***	-0.141***
			(0.041)	(0.042)
toddler			-0.017	-0.029
			(0.041)	(0.040)
adol			0.019	0.020
			(0.041)	(0.041)
teen			0.008	0.004
			(0.042)	(0.042)
single			-0.022	-0.024
			(0.036)	(0.036)
lnsize			0.005	0.060
			(0.058)	(0.061)
amtpaid				-0.002***
				(0.000)
amtprod_gift				0.000
				(0.002)
amtwic				-0.003
				(0.003)
amtaway				0.001
				(0.001)
nonfood_exp				0.000
				(0.000)
_cons	0.583***	0.550***	0.663***	0.697***
	(0.087)	(0.094)	(0.107)	(0.109)
Observations	1062	1062	1062	1062
R-squared	0.051	0.055	0.068	0.083

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table C.10: Proneness results at 90^{th} quantile of HEI-2005

	GQR-1	GQR-2	GQR-3	GQR-4
white_nh	0.038	0.042	0.042	0.041
	(0.030)	(0.032)	(0.032)	(0.032)
black_nh	0.055^{*}	0.058^{*}	0.059^{*}	0.058^{*}
	(0.033)	(0.034)	(0.034)	(0.034)
hispanic	0.046	0.046	0.048	0.049
	(0.030)	(0.030)	(0.031)	(0.031)
age_head	-0.000	-0.000	-0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)
fem_head	0.011	0.010	0.015	0.015
	(0.031)	(0.031)	(0.032)	(0.032)
married	-0.033	-0.031	-0.034	-0.031
	(0.023)	(0.023)	(0.026)	(0.026)
less_hs		0.022	0.023	0.025
		(0.023)	(0.023)	(0.023)
hs_grad		0.039	0.039	0.038
		(0.025)	(0.025)	(0.025)
working		-0.013	-0.018	-0.021
		(0.023)	(0.024)	(0.024)
infant			-0.037	-0.027
			(0.025)	(0.026)
toddler			-0.014	-0.012
			(0.025)	(0.025)
adol			-0.003	-0.002
			(0.025)	(0.025)
teen			-0.000	-0.000
			(0.026)	(0.026)
single			-0.016	-0.015
			(0.022)	(0.023)
lnsize			0.001	0.000
			(0.035)	(0.037)
amtpaid				-0.000
				(0.000)
amtprod_gift				0.000
				(0.001)
amtwic				-0.003
				(0.002)
amtaway				0.001
				(0.001)
nonfood_exp				0.000
				(0.000)
_cons	0.870^{***}	0.851^{***}	0.895^{***}	0.894^{***}
	(0.053)	(0.057)	(0.066)	(0.068)
Observations	1062	1062	1062	1062
R-squared	0.008	0.011	0.014	0.016

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Figure C.1: Measures of calories from two data sources

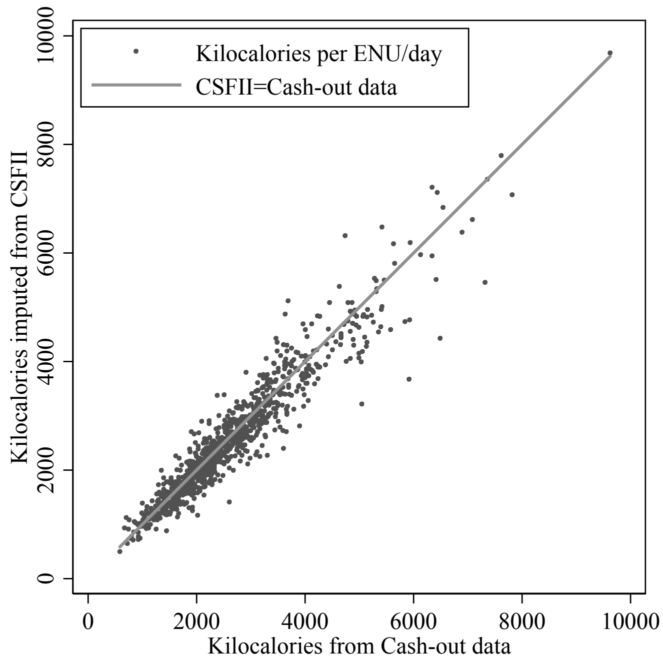


Figure C.2: Measures of saturated fat from two data sources

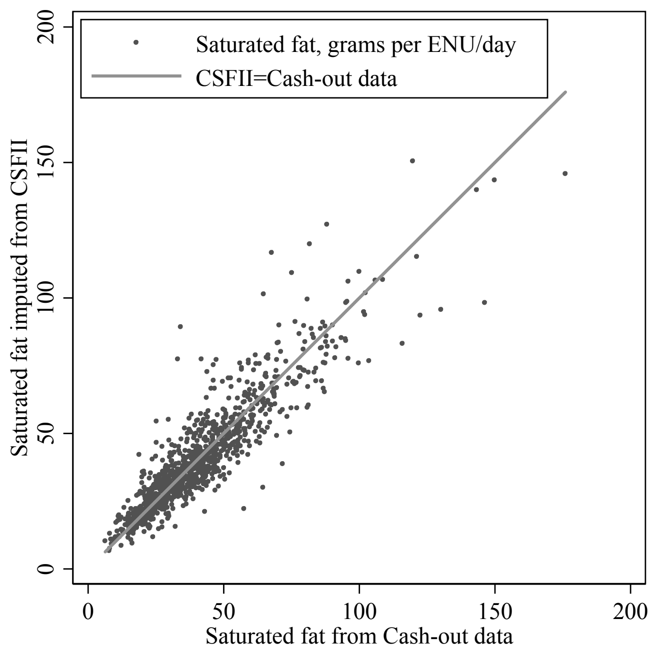


Figure C.3: Cash Out Effect on the Distribution of Kilocalories

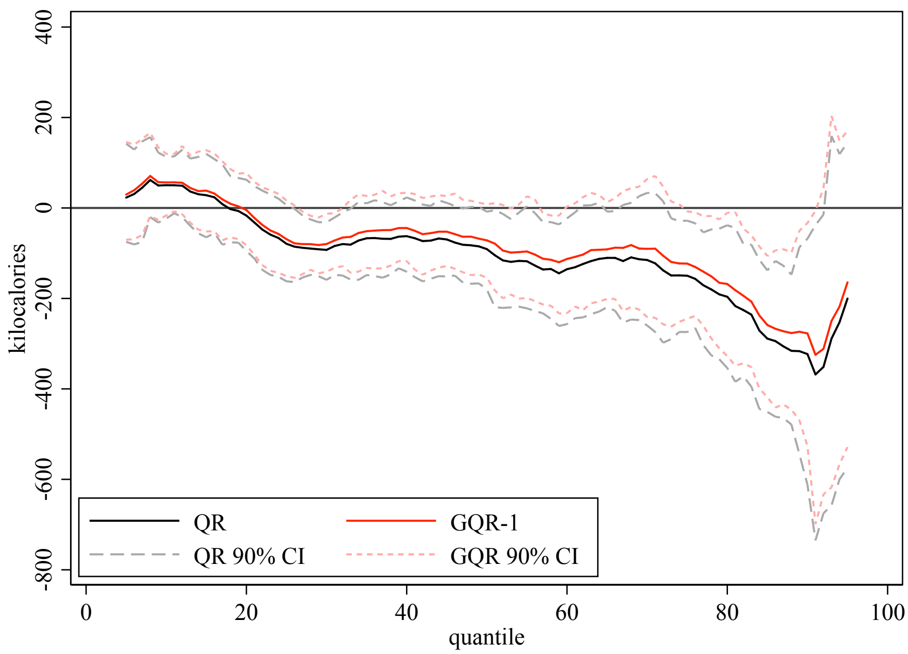


Figure C.4: Cash Out Effect on the Distribution of Kilocalories

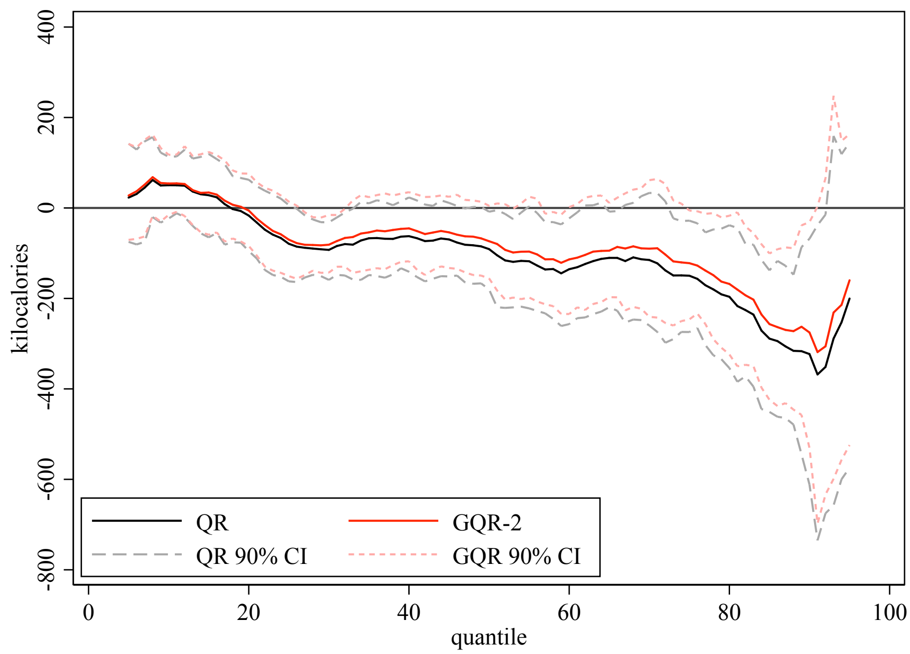


Figure C.5: Cash Out Effect on the Distribution of Kilocalories

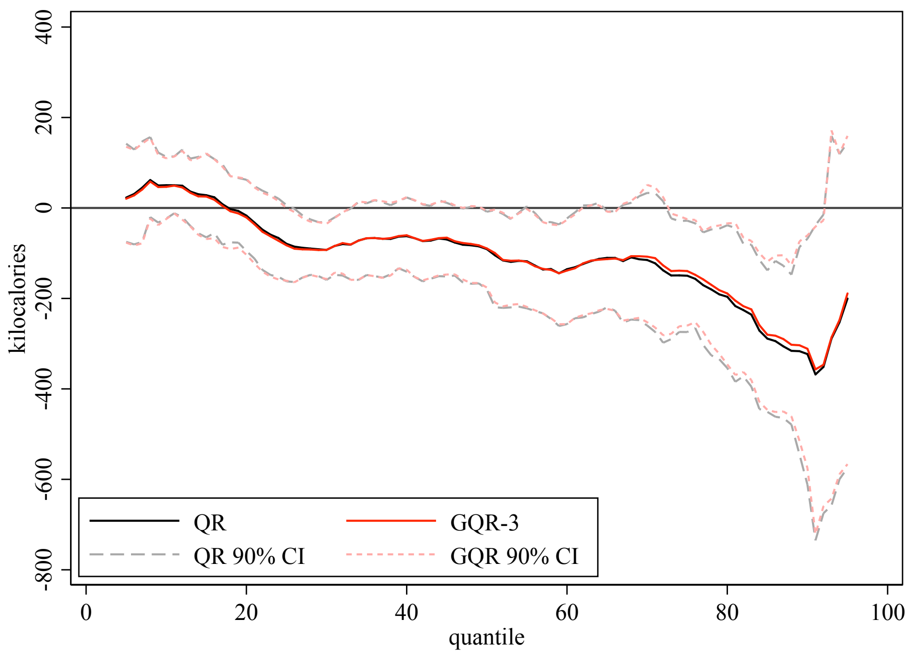


Figure C.6: Cash Out Effect on the Distribution of HEI scores

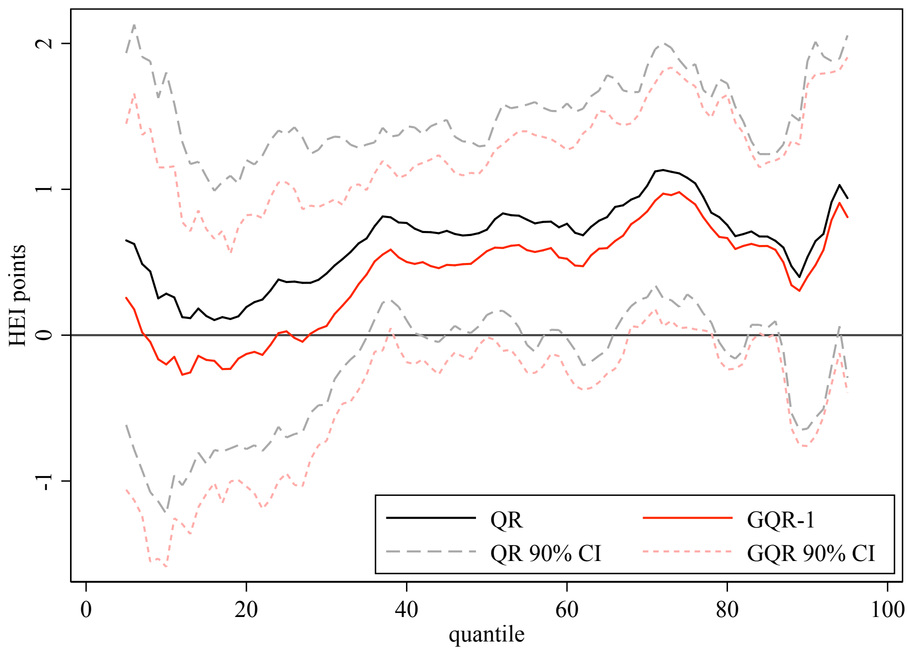


Figure C.7: Cash Out Effect on the Distribution of HEI scores

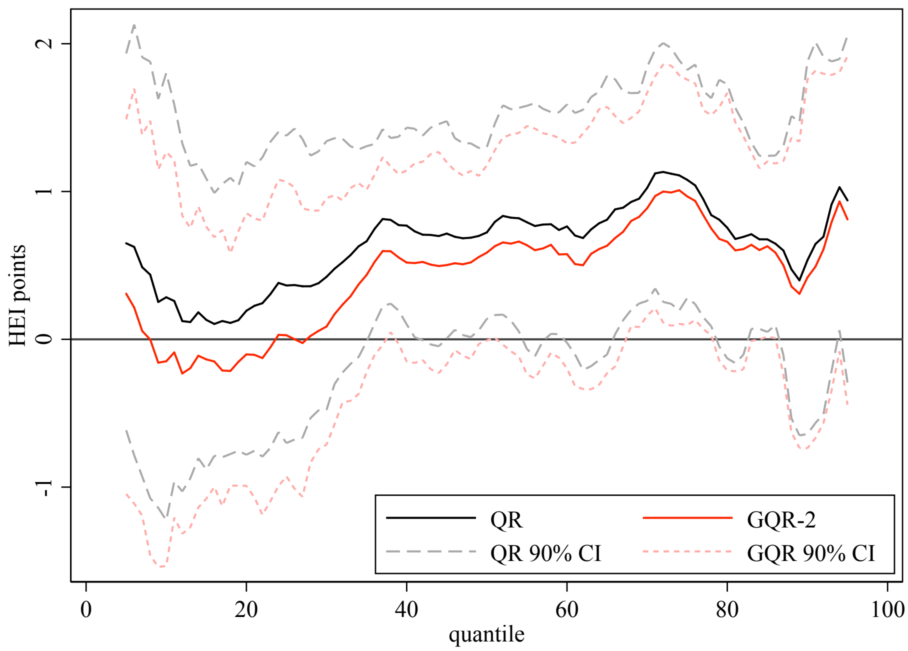


Figure C.8: Cash Out Effect on the Distribution of HEI scores