

IS DIET QUALITY IMPROVING? DISTRIBUTIONAL CHANGES IN THE UNITED STATES, 1989–2008

TIMOTHY K. M. BEATTY, BIING-HWAN LIN, AND TRAVIS A. SMITH

This article measures changes in the distribution of dietary quality among adults in the United States over the period 1989–2008. 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. 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%) vs. the higher-income population (6.4%).

JEL codes: D39, I14, I32.

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), 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 Thomas 1998). In this article, 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 MyPyramid and in 2011 as MyPlate; federally approved SNAP education programs grew from seven 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 changing the relative prices of foods via taxes or subsidies.

Timothy K. M. Beatty is an associate professor in the Department of Applied Economics at the University of Minnesota. Biing-Hwan Lin is a senior economist in the Economic Research Service at the U.S. Department of Agriculture. Travis A. Smith is a PhD candidate in the Department of Applied Economics at the University of Minnesota. The authors thank two anonymous reviewers and the editor, Brian Roe, for comments and suggestions that greatly strengthened the paper. Research for this study was supported by USDA-ERS Cooperative Agreement No. 58-4000-1-0025. The views expressed here are those of the authors and cannot be attributed to the Economic Research Service or the U.S. Department of Agriculture. All remaining errors are our own.

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

In this article, 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 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 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% 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% of the improvement is attributed to changes in food composition (e.g., decreases in saturated fats, sugars, and sodium). The residual 36.6% is unexplained by either changes in demographics or food composition.

The article proceeds as follows. We begin by describing a widely used measure of dietary quality—the Healthy Eating Index 2005 (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 of 1989–91. We then motivate our empirical approach by providing a brief overview of stochastic dominance. After the

presentation of results, we discuss the economic and policy implications in the final section.

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 HEI). In this article, we focus on dietary quality, and we note that there is evidence that higher quality diets are associated with decreased obesity rates (i.e. improved energy balance) (Epstein et al. 2001, 2008).

We use the HEI—developed in 1995 to measure compliance with the U.S. government’s recommendations for healthful eating—as our 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, the HEI 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. 2004). 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 are revised by the USDA and Health and Human Services 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 Dietary Guidelines for Americans. The HEI was updated in 2005 to reflect the 2005 Dietary Guidelines for Americans (frequently called the HEI-2005; see [Guenther, Reedy, and Krebs-Smith 2008](#)).² Because the HEI-2005 was constructed with the 2005 Dietary Guidelines for Americans 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 twelve 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 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, creating a maximum score of 100. Table 1 provides exact details of the scoring.

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 are considered "needs improvement," and scores of less than 51 are considered "poor." Characterizing a diet based on a single cutoff 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

Data

Our sample uses nationally representative, repeated cross-sectional, individual food intake data from two surveys: the CSFII (1989–91 and 1994–96), and the continuous waves of the 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 United States, with the USDA overseeing the food intake component in both surveys. Finally, for consistency across samples, we focus on adults aged 20 years and older.

The HEI 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 1. As noted, because there is no officially released MPED for the 1989–91 CSFII, the HEI has not been previously computed for surveys before 1994. Of the 3,953 unique foods reported by adults aged 20 and older on day 1 in the 1989–91 CSFII, 3,907 (98.8%) of these foods are also reported in the 1994–96 CSFII. We therefore use the 1994–96 MPED to calculate the HEI for individuals in 1989–91.⁵

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

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 online supplementary appendix, results are robust to using two days of intake.

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

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

Table 1. Healthy Eating Index 2005 Standards for Scoring

Component	Score					
	0	5	8	10	20	
Total fruit	0	→	→	→	→	≥ 0.8 cup eq/1,000 kcal
Whole fruit	0	→	→	→	→	≥ 0.4 cup eq/1,000 kcal
Total vegetables	0	→	→	→	→	≥ 1.1 cup eq/1,000 kcal
Dark green/orange vegetables/legumes	0	→	→	→	→	≥ 0.4 cup eq/1,000 kcal
Total grains	0	→	→	→	→	≥ 3.0 cup eq/1,000 kcal
Whole grains	0	→	→	→	→	≥ 1.5 cup eq/1,000 kcal
Milk	0	→	→	→	→	≥ 1.3 cup eq/1,000 kcal
Meats and beans	0	→	→	→	→	≥ 2.5 oz eq/1,000 kcal
Oils	0	→	→	→	→	≥ 12 g/1,000 kcal
Saturated fat	≥15	→	→	10 →	→	≤ 7% of energy
Sodium	≥2.0	→	→	1.1 →	→	≤ 0.7 g/1,000 kcal
Calories from solid fat, alcohol, and added sugar	≥50	→	→	→	→	≤ 20% of energy

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

Table 2. Healthy Eating Index 2005 Summary Statistics

Population	1989–91	1994–96	2001–04	2005–08
U.S. population	50.16 (13.97) ^a [10.09, 96.42]	51.10 (13.88) ^a [10.69, 97.47]	51.50 (11.91) [13.52, 99.46]	52.46 (12.49) [8.78, 95.38]
No.	9,498	9,867	8,640	9,258
Low-income	48.96 (19.83) ^a [10.09, 90.25]	49.36 (15.45) ^a [10.69, 93.81]	49.65 (13.29) ^a [15.08, 99.46]	51.37 (14.99) [8.78, 94.60]
No.	4,965	3,433	3,551	3,857
Higher-income	50.56 (11.19) ^{a,b} [11.51, 96.42]	51.73 (13.16) ^{a,b} [13.63, 97.47]	52.36 (11.09) ^b [13.52, 93.97]	52.92 (11.29) ^b [10.00, 95.38]
No.	4,533	6,434	5,089	5,401

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

^aWithin-population mean is significantly lower than that in 2005–08 at the 5% level.

^bWithin-year higher-income mean is significantly different from low-income mean at the 5% level.

bound on the cutoff for many federal nutrition assistance programs. During our sample period, the cutoff for SNAP was 130%, and the cutoff for WIC was 185%. 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 reports the mean HEI scores for the population as a whole and for individuals above and below 185% of the poverty line for each of the periods in our sample.⁷

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

Table 2 shows a consistent pattern of increasing dietary quality across groups over time. Comparing the most recent period of 2005–08 to the earlier periods, we see a significant increase (at the 5% level) for the population at large over the periods 1989–91 and 1994–96. Low-income individuals appear to have a stagnant HEI score over the period 1989–2004, and then a significant increase in the period 2005–08. We also compare low- and higher-income individuals within year and find that higher-income individuals have significantly higher mean HEI scores for all years in the data, although in the final year of the data, the mean HEI gap between groups is smallest.

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.

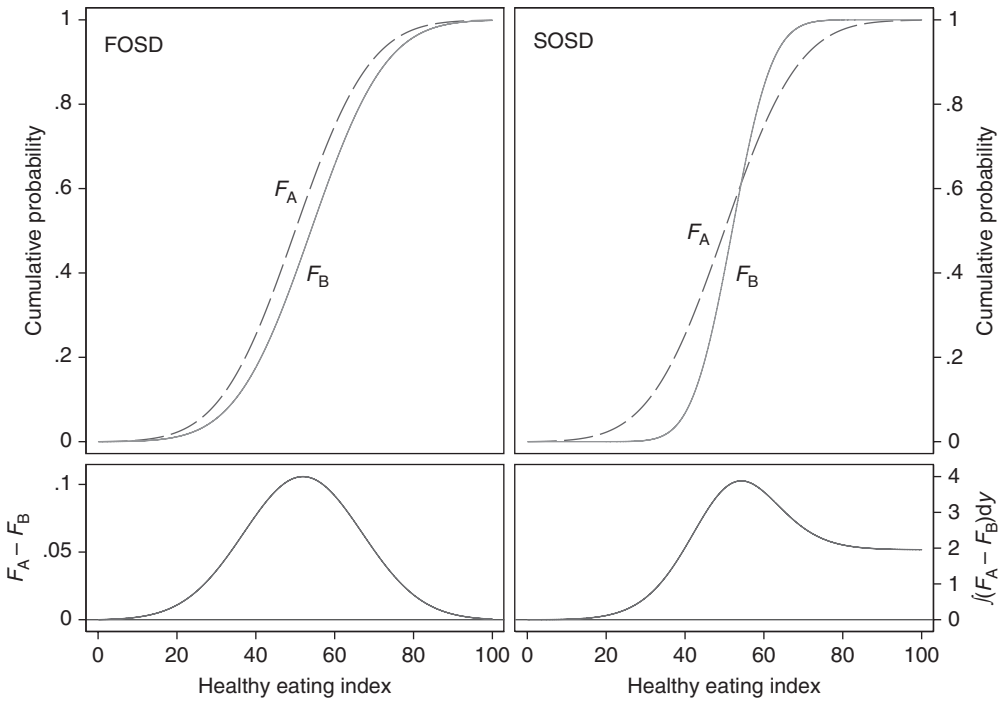


Figure 1. First- and second-order stochastic dominance

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 among 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

Definitions

Consider two distributions of HEI scores with cumulative distribution functions (CDFs) $F_A(z)$ and $F_B(z)$, for a population of interest in two distinct time periods or, alternatively, for two mutually exclu-

sive subpopulations within a single time period. We say that distribution B first-order stochastically dominates 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 1.

Distributional studies of well-being often look to higher orders of stochastic dominance, notably second-order stochastic dominance (SOSD). Whereas first-order stochastic dominance (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 dominates

⁸ Stochastic dominance approaches have also been used to study changes in body mass index (Madden 2012) and environmental quality (Maasoumi and Millimet 2005) and extended to qualitative health measures (Allison and Foster 2004).

F_A if

$$\int_0^z [D_B^1(y) - D_A^1(y)] dy \leq 0 \quad \forall z$$

with a strict inequality for some value of z . This relationship is illustrated in the right panel of figure 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,

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

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 first-order stochastically dominates A , then for any social welfare function W defined on the distribution of diet quality $F(z)$ such that $W(F) = \int U(z) dF(z)$ 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.

Estimation

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

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

Integrating the empirical analogue of (1) by parts leads to a natural estimator of $D_j^s(z)$

$$(2) \quad \hat{D}_j^s(z) = \frac{1}{\hat{N}_j(s-1)!} \times \sum_{i=1}^{n_j} \theta_i (z - y_i)^{s-1} I(y_i \leq z)$$

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

$$(3) \quad \hat{D}_j^1(z) = \hat{F}_j(z) = \frac{1}{\hat{N}_j} \sum_{i=1}^{n_j} \theta_i I(y_i \leq z)$$

and the statistic for the second-order case follows directly.

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 to be the more recent time period, the null hypothesis of an increase in dietary quality at order $s \in \{1, 2\}$ is,

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

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) 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 Bishop, Formby, and Thistle (1989) procedure have been proposed, such as multiple comparison tests (Anderson 1996;

Davidson and Duclos 2000) or Kolmogorov–Smirnov-type tests (Barrett and Donald 2003; Bennett 2010; Linton, Massoumi, and Whang 2005; McFadden 1989). Multiple comparison approaches are based on arbitrarily chosen ordinates, which can lead to test inconsistency (Barrett and Donald 2003; Davidson and Duclos 2000). Therefore, in this study we use a Kolmogorov–Smirnov-type statistic that compares all objects within the support of the two distributions.

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

$$(4) \quad \hat{d}_s^+ = \sup_{z \in Z} [\hat{D}_B^s(z) - \hat{D}_A^s(z)]$$

$$(5) \quad \hat{d}_s^- = \sup_{z \in Z} [\hat{D}_A^s(z) - \hat{D}_B^s(z)].$$

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 $H_0^{s+} : \hat{d}_s^+ \leq 0$, and similarly for decreased diet quality (F_A dominating F_B) using \hat{d}_s^- . When the distributions are mutually independent, Kolmogorov–Smirnov-type tests based on \hat{d}_s^\pm are consistent (McFadden 1989).⁹ Test statistics are calculated using

$$(6) \quad \hat{T}_s^\pm = \sqrt{\frac{n_A n_B}{n_A + n_B}} \hat{d}_s^\pm.$$

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 to construct the null distribution. The least favorable configuration is the point in the null distribution that is least favorable to the alternative hypothesis (i.e., $F_A = F_B$). As a result, the test is conservative; rejection of the null under the least favorable configuration implies rejection at any point in the null distribution. We

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 the Counterfactual Analysis section, we relax this assumption to ignorability (also called conditional independence or unconfoundedness) to construct aggregate counterfactual decompositions.

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) 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 (4) and (5). The recentered bootstrap t statistics are

$$(7) \quad \hat{T}_s^{*\pm} = \sqrt{\frac{n_A n_B}{n_A + n_B}} \hat{d}_s^{*\pm}.$$

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

$$(8) \quad \hat{p}_s^\pm \simeq \frac{1}{B} \sum_{i=1}^B I(\hat{T}_s^{*\pm} > \hat{T}_s^\pm).$$

The p values allow for a test of stochastic dominance 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 Bishop, Formby, and Thistle (1989) 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 .

Robustness Check for FOSD

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 below B, or the curves cross. The Bishop,

¹⁰ Our samples are constructed using multistage stratification where each stratum is clustered by two primary sampling units. Test statistics based on a simple random bootstrap sample 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 primary sampling unit 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) to create the bootstrap distribution of $\hat{T}_s^{*\pm}$.

Formby, and Thistle (1989) 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 overclassification 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 Bishop, Formby, and Thistle (1989) procedure.

A second drawback to the Bishop, Formby, and Thistle (1989) 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 Bishop, Formby, and Thistle (1989) 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 Bishop, Formby, and Thistle (1989) 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 prespecified 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 preset 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%) 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 (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

$$(9) \quad \lim_{n_A, n_B \rightarrow \infty} \mathbf{P}[K_{min} \leq b | K_{max} > a] \rightarrow \frac{2[G_1(b) - G(a, b)]}{1 - G_2(a)}$$

where¹¹

$$(10) \quad G_1(b) = 1 - e^{-2b^2}$$

$$(11) \quad G_2(a) = \frac{\sqrt{2\pi}}{a} \sum_{k=1}^{\infty} e^{-(2k-1)^2\pi^2/(8a^2)}$$

$$(12) \quad G(a, b) = \sum_{k=-\infty}^{\infty} e^{-2(k(a+b))^2} - \sum_{k=-\infty}^{\infty} e^{-2(a+k(a+b))^2}.$$

The two-stage p values (denoted $p_{1,\alpha}^{2S}$) are calculated from (9) 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.

Results

Our main results are summarized in tables 3 and 4 and depicted in figures 2–4. Tables

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

Table 3. Tests of Stochastic Dominance among U.S. Adults

Distribution		Bootstrap Tests				Two-Stage		Inferred Ranking
A	B	\hat{p}_1^-	\hat{p}_2^-	\hat{p}_1^+	\hat{p}_2^+	$p_{1,0.1}^{2S}$	$p_{1,0.01}^{2S}$	
1989–91	1994–96	0.007	0.010	0.900	0.660	0.507	0.362	$A <_1 B^{***}$
	2001–04	0.028	0.002	1.000	0.937	0.999	0.999	$A <_1 B^{**}$
	2005–08	0.002	0.000	1.000	0.877	0.981	0.966	$A <_1 B^{***}$
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	$A <_1 B^{***}$
2001–04	2005–08	0.133	0.051	0.991	0.790	0.972	0.950	$A <_2 B^*$

Notes: The \hat{p}_s^\pm values refer to one-sided tests of the null hypothesis H_s^\pm using equation (8). The asymptotic $p_{1,\alpha}^{2S}$ values are calculated from (9), where $\alpha = 0.1, 0.01$. The notation $A <_s B$ 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% (indicated by a triple asterisk), 5% (indicated by a double asterisk), and 10% (indicated by a single asterisk).

Table 4. Tests of Stochastic Dominance among U.S. Adults by Income Group

Distribution		Bootstrap Tests				Two-stage		Inferred Ranking
A	B	\hat{p}_1^-	\hat{p}_2^-	\hat{p}_1^+	\hat{p}_2^+	$p_{1,0.1}^{2S}$	$p_{1,0.01}^{2S}$	
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	$A <_1 B^{***}$
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	$A <_1 B^{***}$
2001–04	2005–08	0.031	0.033	0.997	0.818	0.998	0.996	$A <_1 B^{**}$
Higher income								
1989–91	1994–96	0.007	0.006	0.880	0.684	0.429	0.289	$A <_1 B^{***}$
	2001–04	0.004	0.000	0.999	0.897	0.999	0.999	$A <_1 B^{***}$
	2005–08	0.002	0.000	1.000	0.906	0.985	0.973	$A <_1 B^{***}$
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	$A <_1 B^{***}$
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 (8). The asymptotic $p_{1,\alpha}^{2S}$ values are calculated from (9), where $\alpha = 0.1, 0.01$. The notation $A <_s B$ 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% (indicated by a triple asterisk), 5% (indicated by a double asterisk), and 10% (indicated by a single asterisk).

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 economically important improvement in the HEI scores over the period under study. For any level of dietary quality, more Americans have higher HEI scores in 2005–08 than they did in the period 1989–91. However, there are differences between income

groups with regards to when and where the improvements occurred (table 5).

Between Periods

Figure 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 toward 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 subpanel. The area under

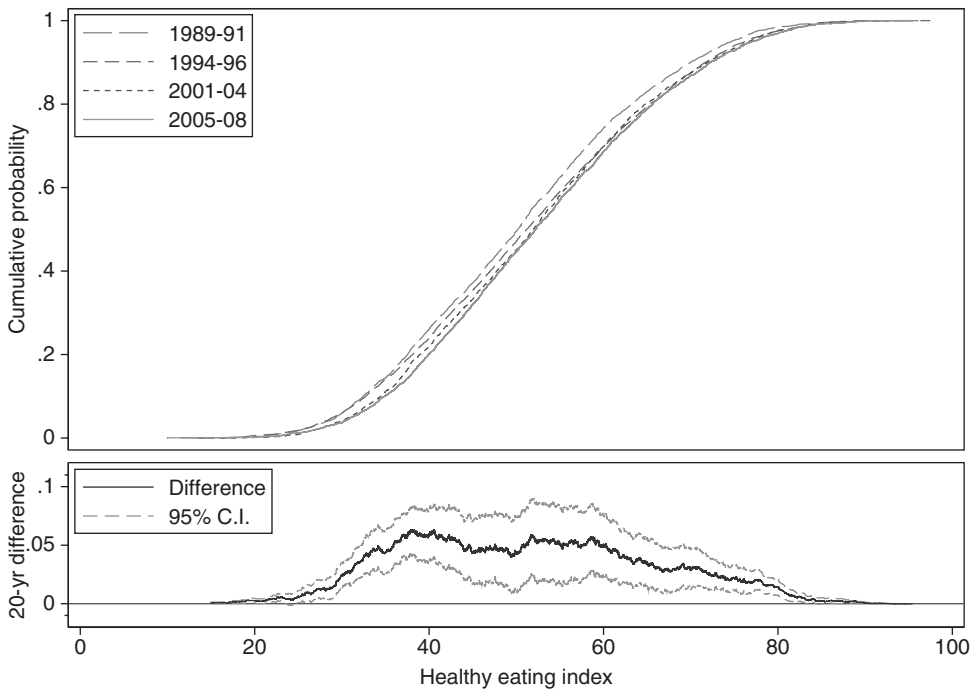


Figure 2. Distribution of adult HEI-2005 scores in the U.S. population

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

the difference curve in the subpanel 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.

From the first three rows of table 3, we see that in comparing the period 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 the periods 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 the period 2005–08 first-order stochastically dominates the period 1994–96, but the results are fairly weak with regards to an ordering of the periods 2005–08 and 2001–04.

Some care is required in interpreting the last two columns of table 3 because they report results from Bennet's two-stage test. As noted, these p values are for the null hypothesis that the CDFs do not cross, as determined by both K_{max} and K_{min} 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 in that there is no statistical evidence that the 1989–91 distribution crosses any of the later years.

Between Income Groups

We now turn our attention to direct comparisons of individuals above and below 185% of the poverty guideline. As noted, we choose 185% of the poverty line as our cutoff because it is an upper limit on the threshold for many federal nutrition assistance programs.¹² Panel (a) of figure 3 presents the empirical CDFs and the difference between the periods 1989–91 and 2005–08 for low-income individuals; panel (b) likewise for higher-income individuals. Table 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 the period 2005–08 first-order stochastically dominates

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 online supplementary appendix for interested readers.

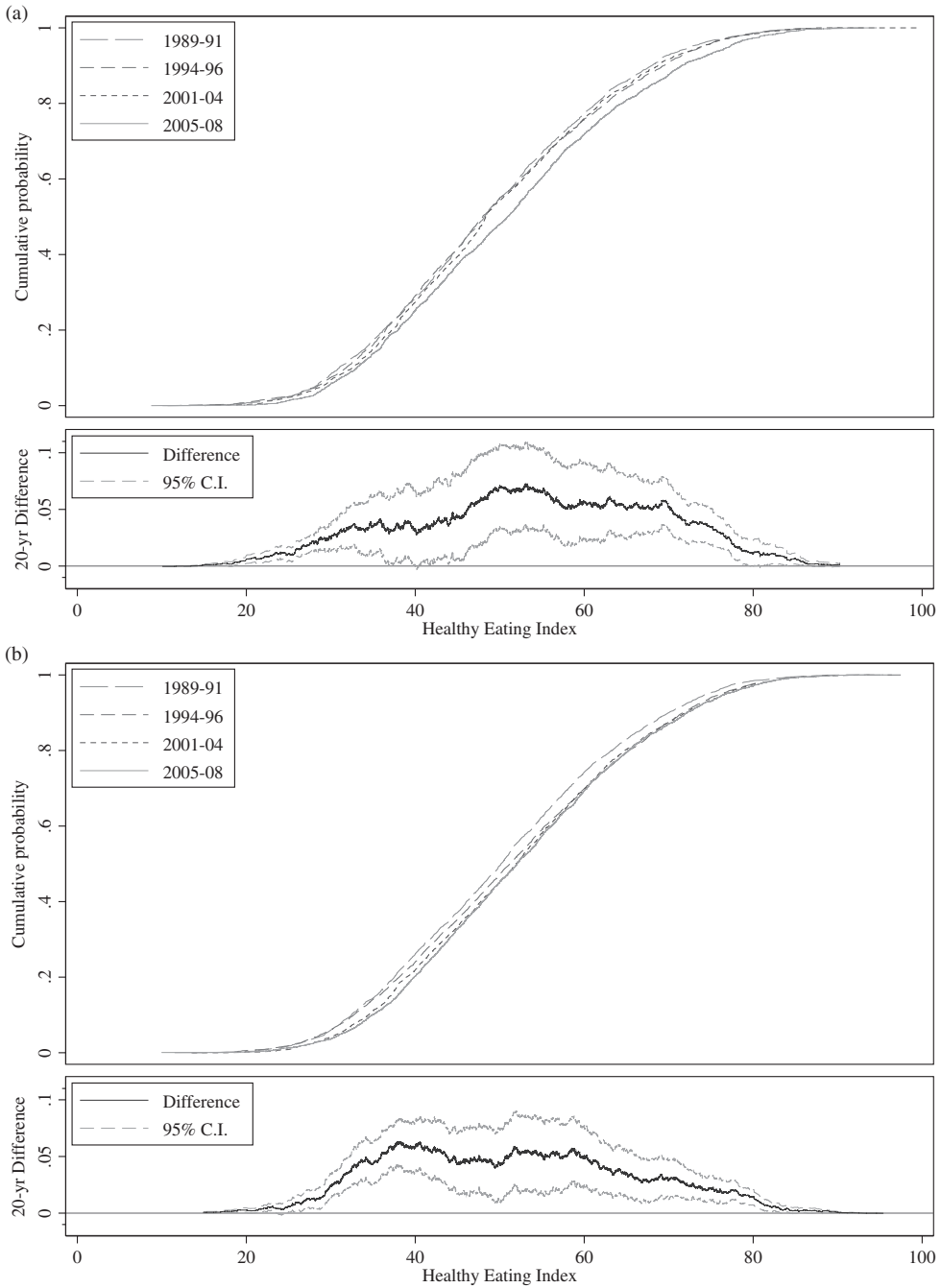


Figure 3. Distribution of adult HEI-2005 scores by income group

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

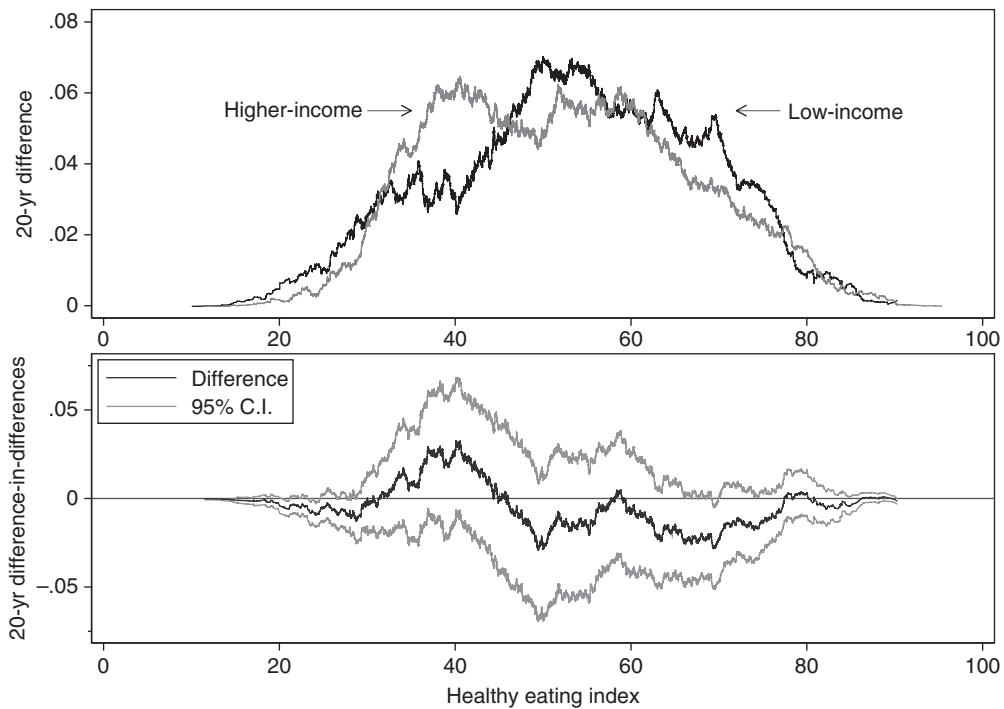


Figure 4. Differences in dietary improvements among low- and higher-income populations

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

the distribution in the earliest period, with no evidence of a crossing.¹³

Results support the observation in table 2 that a significant portion of dietary improvement among low-income individuals occurred over the period 2001–08. For example, in comparing the period 1989–91 to the period 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 the period 1989–91 to the period 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 the period 1989–91 is first-order dominated by each subsequent period, with no evidence of a crossing. Comparing the period 1994–96 to the period 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 the period 2005–08 first-order stochastically dominates the period 1994–96, but we cannot rank the two most recent time periods.

We can compare the total twenty year improvements in each income group by examining subpanels (a) and (b) in figure 3. We see that low-income individuals experienced relatively smaller increases over the bottom tail of the relevant range of HEI compared with their higher-income counterparts. We can more formally investigate this finding by taking the difference (between above and below 185% of the poverty line) in the differences (between the earlier and later periods). Figure 4 superimposes the subfigures in panels (a) and (b) of figure 3 in

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% significance levels.

Table 5. Location and Time-Path of Dietary Improvement

HEI range	Between Period			Total
	$\hat{F}_{89} - \hat{F}_{94}$	$\hat{F}_{94} - \hat{F}_{01}$	$\hat{F}_{01} - \hat{F}_{08}$	$\hat{F}_{89} - \hat{F}_{08}$
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 twenty-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 $\hat{F}_{08} = \hat{F}_{2005-08}$.

the top panel and then plots the difference between the two in the bottom panel—that is, in the subpanel of figure 4 we plot:

$$(13) \quad DD = \left[\hat{D}_{high,89}^1(z) - \hat{D}_{high,08}^1(z) \right] - \left[\hat{D}_{low,89}^1(z) - \hat{D}_{low,08}^1(z) \right].$$

As shown in figure 4, considering lower levels of dietary quality below a HEI of 45, we find higher-income individuals experienced a greater improvement over the period 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 than higher-income individuals.

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, and 60.¹⁴ Table 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 United States over the 20-year period that occurred in the bottom quartile (<40) between 1989–91 and 1994–96 was 2.8%. The last column of table 5 measures the overall improvements over the period 1989–2008 within each quartile of the distribution of dietary quality.

For the U.S. adult population, improvements below the median (HEI <50) occurred steadily over the period 1989–2008. In the upper range of dietary quality (HEI above 50), however, virtually all of the gains occurred over the periods 1989–96 and 2001–08. Overall, there were slightly higher gains

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.

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 the period 2001–08. This is in contrast to the higher-income population, which saw the majority of their improvements occurring over the period 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 the period 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 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 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.

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 and 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.

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](#) found 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.

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

To construct the distribution of dietary quality in the period 1989–91 as if food were formulated in the period 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 1989–91 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 score based on updated nutrient and MPED values for each respondent in the 1989–91 sample. Figure 5 displays the results from the reformulation counterfactual, as well as results from the next section.

The distribution of HEI that accounts for reformulation lies everywhere to the right of the original 1989–91 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 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 the period 1989–91 and the period 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 the period

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

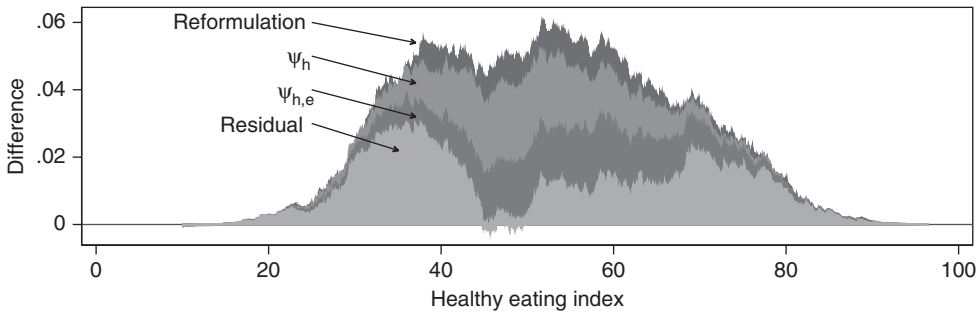


Figure 5. Differences between the 1989–91 counterfactual distributions and observed 2005–08 distribution

Notes: “Reformulation” is defined in the subsection *Food Reformulation*. ψ_h is a reweighting function (see subsection *Demographic Changes*), which includes dummies for female sex and race/ethnicity fully interacted with age groups, and $\psi_{h,e}$ additionally includes education dummies, all as defined in table 5.

1989–91 might have chosen different foods had their foods been formulated as they were in the period 2005–08. Nevertheless, it shows how food reformulation, all else equal, can play an important role in changing dietary quality.

Demographic Changes

The United States of 2005–08 is an older, more diverse, and better educated country than the United States of 1989–91. To the extent that these factors are correlated with healthy eating, they may explain some of the improvements in dietary quality. Table 6 illustrates demographic changes using data from our sample and from the U.S. Census. There is a clear decrease in the population aged 30–44 years and a concomitant rise in the population aged 45–64 years. 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. 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 DiNardo, Fortin, and Lemieux 1996 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 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

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

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

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

Table 6. U.S. Population Characteristics, Adults Aged 20 Years and Older

Demographic	1989–91 CSFII	2005–08 NHANES	1990 Census ^a	2005–07 Census ^b
Aged 20–29 years ^c	21.7	19.4	22.7	19.1
Aged 30–44 years	35.9	28.3	33.5	29.1
Aged 45–64 years	26.2	35.4	26.1	35.0
Aged ≥65 years	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
Total No.	9,377	9,257		

^aU.S. Census Bureau, General Population Characteristics (CP-1, 3-4).

^bU.S. Census Bureau, 2005–07 Annual Community Survey three-year sample.

^cAll numbers are expressed as the percentage of the adult population aged 20 and older

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

$$(16) \quad \psi(h) = \frac{\Pr(t_h = 08|h)}{\Pr(t_h = 89|h)} \cdot \frac{\Pr(t_h = 89)}{\Pr(t_h = 08)}.$$

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. Because 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).

Although long-run demographic changes such as sex, 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

Because of the aggregate decompositional nature of the DiNardo, Fortin, and Lemieux methodology (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 (sex, age, and race/ethnicity) and denote this reweighting function by ψ_h . We then construct a counterfactual distribution accounting for changes in demographics and changes in education levels, denoted by $\psi_{h,e}$.¹⁷ We investigate the effects of the ordering herein.

Figure 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

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.

¹⁷ The conditional probability model includes a dummy for sex, 16 cells of race/ethnicity fully interacted with age dummies, and three education dummies, all as described in table 6. Results of the model are available in the online supplementary appendix.

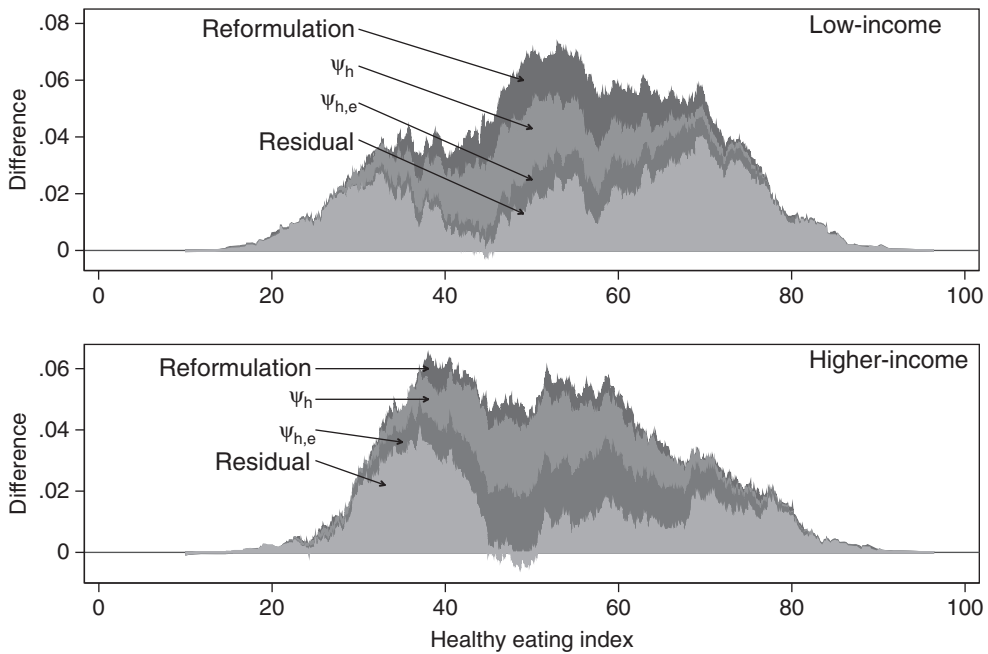


Figure 6. Differences between the 1989–91 counterfactual distributions and observed 2005–08 distribution by income group

Notes: “Reformulation” is defined in the subsection *Food Reformulation*. ψ_h is a reweighting function (see subsection *Demographic Changes*), which includes dummies for female sex and race/ethnicity fully interacted with age groups, and $\psi_{h,e}$ additionally includes education dummies, all as defined in table 5.

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 sex, 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 toward 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. Although this assumption is economically unappealing, the exercise provides insight into the effects of changing

demographics on diet quality via clear and tractable analytical techniques.

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 6.

Changes in food composition account for a substantially larger percentage of the dietary improvement for lower-income individuals (19.6%) compared with 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 sex, age, and race/ethnicity account for a similar

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

Table 7. DiNardo, Fortin, and Lemieux (1996) Counterfactual Improvements

Order	U.S. Population	Low Income	Higher Income
Original order			
1. Reformulation	10.1	19.6	6.4
2. Sex, age, race/ethnicity	26.6	25.3	26.8
3. Education	26.7	13.5	27.0
Total: reformulation & demographics	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. Sex, Age, Race/Ethnicity	37.7	33.7	37.7
Total: reformulation & demographics	63.4	58.4	60.2

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

share of the improvement for low-income (25.3%) versus higher-income individuals (26.8%). For low-income individuals, changes in educational attainment account for half that of higher-income individuals (13.5% versus 27.0%). The remaining residual share of the twenty-year improvement is larger within the low-income population (41.6%) than the higher-income population (39.8%). This suggests that further research into the determinants of diet quality of low-income individuals may be warranted.

Robustness

The order in which we construct counterfactual distributions using the DiNardo, Fortin, and Lemieux (1996) 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 in 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 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.

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 toward 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.

Although 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

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

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 scores for both women (over the period 1984–2008) and men (over the period 1986–2008) who were free of chronic disease at baseline. For example, those in the second quintile were 7% less likely to report a chronic disease than those in the lowest quintile, all else equal. One way to assess 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 the period 1989–91, the twentieth percentile of the HEI distribution was 37.3. In 2005–08, a HEI value of 37.3 represented the 15.4th percentile of the HEI distribution. In other words, 4.6% of individuals moved out of this higher-risk category between 1989 and 2008 because of 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.

Supplementary material

Supplementary material is available at http://oxfordjournals.org/our_journals/ajae/ online.

References

- Alkire, S., and J. E. Foster. 2011. Counting and Multidimensional Poverty Measurement. *Journal of the Public Economics* 95 (7): 476–87.
- Allison R. A., and J. E. Foster. 2004. Measuring Health Inequality using Qualitative Data. *Journal of Health Economics* 23 (3): 505–24.
- Anderson, G. 1996. Nonparametric Tests for Stochastic Dominance in Income Distributions. *Econometrica* 64 (5): 1183–93.
- Barrett, G. F., and S. G. Donald. 2003. Consistent Tests for Stochastic Dominance. *Econometrica* 71 (1): 71–104.
- Bennett, C. 2010. On Bidirectional Tests for Stochastic Dominance. Paper presented at the Annual Meetings of the Midwest Econometrics Group, Saint Louis, MO.
- Bennett, C. 2013. Inference for Dominance Relations. *International Economic Review*, 54(4): 1309–1328.
- Bhattacharya, D. 2007. Inference on Inequality from Household Survey Data. *Journal of Econometrics* 137(2): 674–707.
- Bhattacharya, J., J. Currie, and S. Haider. 2004. Poverty, Food Insecurity and Nutritional Outcomes in Children and Adults. *Journal of Health Economics* 23 (4): 839–62.
- . 2006. Breakfast of Champions? The School Breakfast Program and the Nutrition of Children and Families. *Journal of Human Resources* 41 (3): 445–66.
- 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): 1149–54.
- Billingsley, P. 1968. *Convergence of Probability Measures*. New York: John Wiley & Sons.
- Bishop, J.A., J.P. Formby, and P.D. Thistle. Statistical Inference, Income Distributions, and Social Welfare. In *Research on Economic Inequality*, Volume 1, ed. D.J. Slottje, 49–82. Greenwich, CT: JAI Press.
- 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): 153–76.
- 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): 504–13.
- 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): 1009–18.
- Dardanoni, V., and A. Forcjan. 1999. Inference for Lorenz Curve Orderings. *Econometrics Journal* 2 (1): 49–75.
- Davidson, R., and J. Y. Duclos. 2000. Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality. *Econometrica* 68 (6): 1435–64.
- Deaton A. 1997. *The Analysis of Household Surveys: A Microeconomic Approach to Development Policy*. Baltimore: John Hopkins University Press.
- 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): 1001–44.
- 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.
- Drewnowski, A., and A. Barratt-Fornell. 2004. Do Healthier Diets Cost More? *Nutrition Today* 39 (4): 161–168.
- Duclos, J. Y., D. E. Sahn, and S. D. Younger. 2006. Robust Multidimensional Poverty Comparisons. *Economic Journal* 116 (514): 943–68.
- 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): 171–78.
- 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): 318–26.
- 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.
- 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): 1922–29.
- 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): 1529–34.
- 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): 1725–9.
- Gastwirth, J. L., and T. K. Nayak. 1999. Comments on “Tests of Significance for Lorenz Partial Orders” by J. A. Bishop and J. P. Formby. In *Handbook of Income Inequality Measurement*, ed. J. Silber, 336–9. Amsterdam: Springer.
- 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.
- 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.
- 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): 1854–64.
- 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.
- Guo, X., B. A. Warden, S. Paeratakul, and G. A. Bray. 2004. Healthy Eating Index and Obesity. *European Journal of Clinical Nutrition* 58 (12): 1580–6.
- Hall, K. D., S. B. Heymsfield, J. W. Kemptz, 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.
- 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.
- Kant, A. 1996. Indexes of Overall Diet Quality: A Review. *Journal of the American Dietetic Association* 96 (8): 785–91.
- 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): 346–63.
- Kolenikov, S. 2010. Resampling Variance Estimation for Complex Survey Data. *Stata Journal* 10 (2): 165–99.
- Kourlaba, G., and D. B. Panagiotakos. 2009. Dietary Quality Indices and Human Health: A Review. *Maturitas* 62 (1): 1–8.
- 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.
- 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.
- Maasoumi, E., and D. L. Millimet. 2005. Robust Inference Concerning Recent Trends in Environmental Quality. *Applied Econometrics* 20 (1): 55–77.
- Madden, D. 2012. A Profile of Obesity in Ireland: 2002–2007. *Journal of the Royal Statistical Society: Series A* 175 (4): 893–914.
- McFadden, D. 1989. Testing for Stochastic Dominance. In *Studies in the Economics of Uncertainty: In Honor of Josef Hadar*, ed. T. B. Fomby and T. K. Seo, 113–134. New York: Springer.
- 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): 2112–8.
- 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 (10): 715–20.
- 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): 209–17.
- Ravaillon, M. 1996. Issues in Measuring and Modelling Poverty. *Economic Journal* 106 (438): 1328–43.
- 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.
- 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): 342–7.
- Strauss, J., and D. Thomas. 1998. Health, Nutrition and Economic Development. *Journal of Economic Literature* 36 (2): 766–817.
- U.S. Department of Health and Human Services. 2013. Poverty Guidelines, Research, and Measurement. Available at: <http://aspe.hhs.gov/poverty/>.
- 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): 562–3.
- 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): 959–71.