Quantifying the cost of nutritious diets and dietary impacts on health: Economic approaches to global food systems and nutrition transition

A Doctoral Dissertation by

Yan Bai, MA MIB

Tufts University

Friedman School of Nutrition Science and Policy

January 28th, 2021

Committee

Dr. William A. Masters (Chair), MA PhD

Dr. Elena N. Naumova, MS PhD

Dr. Gitanjali Singh, MPH PhD

ABSTRACT

Background: The cost of nutritious diets that would improve health outcomes has become a major concern for governments and development agencies involved in food systems around the world. This dissertation uses observed food prices to measure the affordability of healthy diets globally and in East Africa, and uses observed quantities consumed to measure the quality of actual diets in terms of health outcomes in the United States. These studies combine market data with nutritional information on food composition and its consequences for human health, using economic principles to construct and apply metrics that can guide policies and programs with transparent methodologies and robust validation.

Methods: Applying multiple datasets and quantitative methods, we examined two novel indices developed using economic approaches. The Cost of Nutrient Adequacy (CoNA), rooted in the cost-of-living indices, quantifies the least cost of a diet meeting daily nutritional needs, and it is proposed as a measure of access to nutritious diets for low-income populations. Using a global food price dataset of the World Bank, Study 1 assessed CoNA over 20 demographic groups in 172 countries in 2017. Study 2 further measured the seasonality of CoNA in Tanzania, Malawi, and Ethiopia, using time-series subnational price datasets. Study 3 investigated the second index, the International Diet-Health Index (IDHI), inspired by weights in economic indices, measuring the dietary impacts on health given a population's health status and disease burdens. Using data from the National Health and Nutrition Examination Survey (NHANES), we analyzed the trends of IDHI in the US from 2003-2014 and compared IDHI with a modified Alternative Healthy Eating Index (mAHEI), as well as validated indices against all-cause mortality.

Findings: Study 1 finds that CoNA, with a global median of $2.30 (IQR: 1.95-2.75) in 2017, peaked for males and lactating women of 14-18 y/o. Female groups of all ages faced higher costs per 1,000kcal compared to their male peers. Study 2 finds the significant intensity of seasonality in CoNA in Malawi, Tanzania, and Ethiopia (10.0 vs. 6.3 and 4.0%), driven primarily by synchronized price rises for nutrient-dense foods, especially fruits and vegetables. Seasonality in CoNA also presents regional variations within countries. Study 3 reveals that the adjusted-mean of IDHI in the US declined from -0.314 in 2003/04 to -0.325 in 2013/14. Non-Hispanic Black Americans had persistently lower IDHI, and disparities in IDHI widened over time by levels of income, and education. IDHI was more closely correlated with the mAHEI at higher levels of diet quality, and both indices were strongly associated with total mortality.

Conclusions: Nutritious diets are not affordable for the global poor, and the costs differ across demographic groups and over seasons. CoNA could be used as a metric to target populations at risk of undernutrition and to evaluate the nutritional performance of food systems over localities and time. Similarly, IDHI is a valid tool for measuring diet-related health impacts in the context of a population's most prevalent diseases, potentially offering tailored guidance regarding how best to reduce diet-related health disparities. International collaboration on nutrition-sensitive data is urgently needed in global applications of both indices.

ACRONYMS AND ABBREVIATIONS

AHEI: Alternative Healthy Eating Index

AI: Adequate Intakes

AMDR: Acceptable Macronutrient Distribution Ranges

BMI: Body mass index

CDC: Centers for Disease Control and Prevention

CDRR: Chronic Disease Risk Reduction

CI: Confidence Intervals

CMD: Cardio-Metabolic Diseases

CoCA: Cost of Caloric Adequacy

CoNA: Cost of Nutrient Adequacy

CPI: Consumer Price Index

CSA: Central Statistical Agency

DALY: Disability-Adjusted Life Year

DDS: Dietary Diversity Scores

DQIs: Diet Quality Indices

DRI: Dietary Reference Intake

DRV: Dietary Reference Value

EAR: Estimated Average Requirement

EER: Estimated Energy Requirement

FCT: Food Composition Table

GHDx: Global Health Data Exchange

GLM: General Linear Models

GNI: Gross National Income

H-ARs: Harmonized Average Requirements

HEI: Healthy Eating Index

HR: Hazard Ratio

ICP: International Comparison Program

IDHI: International Diet-Health Index

LCU: Local Currency Units

LMICs: Low- and Middle-Income Countries
IRR: Incidence Rate Ratio
mAHEI: modified Alternative Healthy Eating Index
MDD-W: Minimum Dietary Diversity for Women
MDS: Mediterranean Diet Score
NCHS: National Center for Health Statistics
NDI: National Death Index
NHANES: National Health and Nutrition Examination Survey
NRV: Nutrient Reference Value
PAL: Physical Activity Level
OLS: Ordinary Least Squares
PPP: Purchasing Power Parity
RDA: Recommended Dietary Allowances
RR: Relative Risk
SNNPR: Southern Nations, Nationalities, and Peoples' Region
TFP: Thrifty Food Plan
UL: Tolerable Upper Intake Levels
NSDA SR28: National Nutrient Database for Standard Reference 28
WAFC: West African Food Composition Table
WHO: World Health Organization

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Article 1: Global variation in the cost of a nutrient adequate diet by population group

Article 2: Seasonality of diet costs reveals food system performance in east Africa

Article 3: A novel diet-health index reveals worsening trends and widening disparities in the health impacts of diet in the United States

SECTION I: INTRODUCTION

Access to foods/diets and dietary impacts on health are two essential concepts in global nutrition research domains. The relationship between them can be complicated and is rapidly evolving in populations and countries with different development and socioeconomic status. In low-income countries, studies have shown that high or volatile food prices harm people's access to foods and their nutritional status.[1–3] Meanwhile, studies in high and middle income countries indicated that barriers in accessing healthy and affordable food retailers and low-price in foods high in calories, saturated fat and added sugar are key reasons for overnutrition eventually leading to high chronic diseases prevalence.[4–6] However, to the extent of our knowledge, the existing indices measuring these two important concepts vary across countries, projects, and research disciplines applying different methods, which create a major barrier in global studies. There is a critical need to explore and develop internationally comparable indices with transparent methodologies and robust validations.

In this dissertation work, we comprehensively examined two promising indices by applying various quantitative methods of economic and statistical models and using multiple global or national datasets. The first index is the Cost of Nutrient Adequacy (CoNA), defined as the least cost of a diet meeting daily nutritional needs. Such method is rooted in the cost-of-living indices in economics with optimization constraints determined by nutrition sciences rather than consumer utility functions.[7,8] The index is proposed as a measure of access to nutritious diets for low-income populations so as to evaluate the efficiency of food systems to deliver nutritionally complete diets at low cost.[9] Using the food retail prices data from the International Comparison Program (ICP) led by the World Bank, Study 1 assessed CoNA of 20 demographic groups in 172 countries in 2017. Study 2 further analyzed the nationally representative time-series datasets of food retail prices provided by the governments of Tanzania, Malawi and Ethiopia. We assessed the magnitude and timing of seasonality of CoNA for a representative adult female of 19-30 years old, comparing to the least cost of only meeting daily estimated energy requirement (the Cost of Caloric Adequacy, or CoCA).

Study 3 investigated and validated the second index, namely the International Diet-Health Index (IDHI). Inspired by economic indices applying relevant weights in index formulation, IDHI weights dietary intakes of a population by relative risks of those dietary risk factors for diseases of interests and relative burden of diseases in the population, which was designed to measure the dietary impacts on health through diet-related diseases given a population's health status and causes of death and disability.[10] In addition, different from existing dietary quality indices, IDHI's formula clearly specify dietary risk factors and the diseases to be evaluated with potential to cover even more diet exposures and health outcomes. Using data from 8 rounds of the National Health and Nutrition Examination Survey (NHANES) between 2003 and 2014, we analyzed the trends of IDHI in the US for different subpopulations by sex, race, income and education, and compared IDHI with a modified Alternative Healthy Eating Index (mAHEI), as well as validated both indices against all-cause mortality.

Through the studies, we obtained comprehensive understandings and generated rigorous applications of the two novel indices, bridging economic methods and evidence from nutrition and health sciences. As internationally comparable indices, they have the potential to be used as policy tools to measure and evaluate changes in access to nutritious diets of various food systems around the world and health impacts from diets in countries facing complicated and evolving nutritional and health challenges.

SECTION II: STATEMENT OF HYPOTHESIS

1. Global variation in diet costs by population group

As a reference cost-of-living index measuring the cost of nutritious diets, the cost of nutrient adequacy (CoNA) is jointly determined by food prices and availability, as well as energy needs and nutrient requirements of different demographic groups. Controlling on other factors, higher food prices, less available food items, and higher nutrient requirements may cause a higher cost of having a nutritionally satisfied diet. CoNA may be positively or negatively correlated with energy needs because on one hand, higher energy needs require more food intakes, on the other hand provide higher allowance for less nutrient dense and usually cheaper foods to achieve the optimization goal.

2. Seasonality of diet costs in East Africa

Seasonal variations in the least costs of nutritious diets are derived from seasonal price changes of foods and food groups selected in the least cost diets. Synchronized seasonality of food prices will increase the seasonal intensity of diet costs, defined as the average absolute difference between the peak and nadir values in a yearly cycle, while the allowance of substitution across foods and food groups in CoNA may help to alleviate seasonal intensity of diet costs. Also, agricultural production cycles, market infrastructure factors, and temporal variations of food demands may jointly determine the seasonality of CoNA across nations, regions, and localities.

3. Health impacts of diet in the United States

Population in a given age-sex group with higher International Diet-Health Index (IDHI) scores are expected to have higher/lower intakes of healthy/harmful dietary factors which are more critical to diseases causing relatively heavier burdens in terms of disability-adjusted life year (DALY) in such age-sex group. Therefore, a higher IDHI is expected to be associated with a lower mortality rate controlling on covariates such as age, sex, race, smoking, and socioeconomic status of education and income.

SECTION III: REVIEW OF THE LITERATURE

1. Food price indices

Existing food price indices focus either on the prices of traded commodities, such as the FAO Food Price Index[11], or all economic activity (a GDP deflator), market information systems (a producer price index) or retail purchases (a consumer price index). The purpose of most price indices is to capture changes in the cost of what is actually bought and sold, and the relative importance of food items in the index formulation is given by the volumes of consumption, production or trade.[7] Therefore there is no indications for those indices in relation to features of nutrition and health. In contrast, in nutrition and public health domains, researchers have been developing methodologies to track and compare foods, food groups, dietary patterns which are pre-defined from scientific evidence.[12] However, the reference foods or diets may vastly vary according to contexts of studies and sometimes, subjective decisions made by investigators, which is seen as a major barrier for international comparison work.

The Cost of Nutrient Adequacy (CoNA) is a cost-of-living index of market prices, designed to measure the affordability of a nutritious diet based on recommended intake levels for a variety of essential nutrients[9], and it has been applied in some high-impact global studies.[13,14] The least-cost diet method that CoNA is based on has been widely accepted and applied in both economics and nutrition communities. It was originally developed by Stigler[15] and has since then been used for a wide range of purposes. Least-cost diets have attained prominence for use in measuring poverty[8] as well as previous work on tracking the cost of nutritious diets[16–19], and comparing actual choices to least-cost diets within a country.[20,21] The method is also used for making nutritional recommendations for low-income consumers. In the United States, the Thrifty Food Plan (TFP) is the basis for maximum food stamp allotments for low-income Americans.[22] The same method is used internationally, for example, to make recommendations in Denmark[23] and the Netherlands[24], and to help nutrition assistance programs meet the specific needs of children and other vulnerable groups as in the "Cost of Diet" and "Optifood" projects or other approaches.[19,25–27]

With well-defined scientific standards, the CoNA index is designed to measure whether a diet with all sufficient essential nutrients is affordable, as an indicator of access to a nutritious diet in different food systems around the world. Although there is no universal nutrient requirements established across the world yet, international nutrition communities are working towards harmonizing the different nutrient intake recommendations, such as dietary reference intake (DRI), dietary reference value (DRV) and nutrient reference value (NRV), to improve objectivity and transparency.[28] This provides CoNA even stronger foundation to be potentially applied for international comparisons.

2. Food price temporal variations

Regarding the temporal variations of food prices, the study of seasonal variation, or the seasonality, is of policy interest for two reasons. First, seasonality in food prices may translate into changes on individuals and households' food intakes, which have potential negative impacts on people's dietary quality, especially for the urban poor and net food buyers in rural areas in low-income countries.[29–31] Second, seasonal variation of food prices is a more predictable factor than the other drivers of food price spikes and therefore government policies and projects may more accurately target. Although improvement of integrated local markets in Africa may reduce food price volatility[32,33], substantial seasonality of food prices in Africa was still revealed.[30] For example, it was found that the seasonality of staple crops wholesale prices between 16.5% for rice and 33% for maize is two and half to three times larger than in the international reference markets. In the analyses, although the author only included two fruits and vegetables, tomatoes and oranges, both displayed much greater average seasonal variations of 60.8% and 39.8%.[30]

Researchers have also been focused on the relationship between roads and food temporal prices. Transportation costs, determined by road quality and other factors, may take large portion of total wholesale food prices in the locations where do not produce such food item and have relatively bad roads connecting to the major production regions. Therefore, good quality of roads can reduce food price gaps between produce and non-produce regions.[34] Study in Nepal found that roads and bridges density, measured as length of roads and number of bridges per km^2, are significant factors for moderating price levels and price volatility of rice and wheat markets.[33]

3. Dietary quality measurements

Dietary quality measurements are algorithms to evaluate the overall diet in terms of various "healthy" or "nutritious" standards, and categorize individuals in terms of whether their diets are meeting such standards, and the specific methodology and the nutrition standards applied may vary significantly depending on the purpose and populations.[35] For example, Dietary Diversity Scores (DDS), referring to the average count of foods or food groups that people daily consumed, have been proved to be valid indicators with strong positive associations with nutrient adequacy.[36,37]Although food categorizations have always been lacking consensus, it has been largely seen as useful tools reflecting undernutrition or nutrient deficiency issues and is primarily applied in studies in low-income regions.

As prevalence of non-communicable chronic diseases continue growing in both developed and developing countries, more indices have been proposed and applied to address overnutrition issues. Some are based on single dietary factors. For example, Starchy Staple Ratio, defined as the energy contribution from starchy staple foods in a person's daily diet, is found to decline as household incomes increase, indicating that people are deviating from "inferior foods".[38] Similarly, Total or Saturated Fats Ratio, defined as the energy contribution from total or saturated fats in a person's daily diet, is believed to be responsible to a higher global prevalence of obesity and several chronic diseases.[39,40]

However, indices of dietary quality based on a few specific nutrients or risk factors generally predict outcomes less well than broader metrics which assess the quality of a person's diet based on overall nutrition guidelines or dietary patterns.[41] For example, the Healthy Eating Index assesses the adherence to the American Dietary Guideline, the Alternative Healthy Eating Index focuses on the individual diet that is protective to cardiovascular diseases, and the Mediterranean Diet Score indicates the compliance to Mediterranean diet.[42–44] But it was also criticized that such indices may not predict disease or mortality significantly better than individual dietary factors, but can be used to measure the adherence of individuals to dietary patterns.[45] Also, it is difficult to develop indices in countries where evidence-based dietary guidelines have not been robustly developed, which are mostly low- and middle-income countries.

The brief literature review provided in this section is complemented by more specific citations in each section of the dissertation, along with motivation for each step of the research regarding practical application to guide intervention for improved outcomes.

SECTION IV. METHODS AND DATA

All three studies in this dissertation use existing data, combining different kinds of observations using methods that are summarized here and presented in more detail for each article. The software used for data analyses and data visualizations are updated versions of Stata, R, SAS and Excel.

1. The CoNA index

CoNA is computed as the results of the least-cost diet through linear programming method, an optimization tool which have been widely used in the fields of both nutrition and economics.[8,25,46,47] The least-cost diet for CoNA is defined as the solution to:

$$(1)\ min(C) = \Sigma_i p_i q_i \text{ subject to } \Sigma_i n_{ij} q_i \geq NR_{jas} \text{ and } \Sigma_i n_{ie} q_i = E_{as}$$

The objective is to minimize diet cost (C) given the price of each food (p_i), choosing quantities (q_i) to meet or exceed nutrient requirements (NR_{jag}) of nutrient j for a population of age a and sex s given the quantity of each nutrient in each food n_{ij}, within the further constraint of energy balance for nutrient $j=e$ at daily energy intake of E_{as}.

Our initial study only used Estimated Average Requirements (EAR) for adult women between 19-30 years old as the nutrition requirement[9], considering that EAR, the daily nutrient intake level meeting the nutrient needs of half of the healthy individuals in a specific age-sex group, is mostly used to assess nutrient adequacy for a population.[48] Study 1 and 2 applied different DRIs from below categories[49], and detailed description is provided in Section V.

1) Recommended Dietary Allowances (RDA): an estimate of the daily average dietary intake meeting the nutrient needs of 97–98% healthy people of a particular age-sex group;
2) Adequate Intakes (AI): is established if sufficient or adequate scientific evidence is not available to establish an EAR and an RDA. AI is expected to meet or exceed the needs of most individuals in an age-sex group;

3) the Tolerable Upper Intake Levels (UL): is the highest average daily nutrient intake level likely to pose no risk of adverse health effects for nearly all people in an age-sex group;
4) And the Acceptable Macronutrient Distribution Ranges (AMDR): is the range of macronutrient intakes of an energy source that is associated with a reduced risk of chronic disease, yet can provide adequate amounts of essential nutrients.

The Physical Activity Level (PAL) of healthy, well-nourished adults is a major determinant of their total energy requirement. As growth does not contribute to energy needs in adulthood, PAL can be measured or estimated from the average 24-hour TEE and BMR (i.e. PAL = TEE/BMR). Multiplying the PAL by the BMR gives the actual energy requirements.[50]

2. Seasonality and harmonic regression models

Seasonality is defined as seasonal fluctuations in a one-year cycle, and we isolated seasonality from other shocks using harmonic (trigonometric) regression, extracting one specific kind of variation to be compared across countries and regions. Harmonic models have been proved to be more efficient and less biased when the time series is relatively short because it has much less parameters to be estimated comparing to the traditional dummy variable approach model.[30] Detailed model introduction could be found in the Section V.

3. The International Diet-Health Index (IDHI)

The IDHI quantifies the overall impact of commonly consumed healthy and harmful dietary factors on diet-related diseases by calculating the dietary risk-weighted sum of diet-related diseases DALYs for each dietary factor using the equation below.

$$IDRI_{as} = \sum_{j=1}^{n} \sum_{k=1}^{m} \left\{ \left[-ln\left(RR_{jka}\right) \right] * Intake_{jas} * \left(DALY_{kas} / \sum_{k=1}^{m} DALY_{kas} \right) \right\}$$

where the dietary quality of dietary factor *j* in terms of a disease *k* for age *a* and sex *s* is the intake of dietary factor j weighted by the negative log of its relative risk (RR) to the disease *k* at age *a*, and the disease's burden measured by the DALY of such disease divided by the total DALYs of diet-related diseases in that age-sex group. The DALY is a measure of overall disease burden, expressed as the number of years lost due to illness, disability, or early death, and was developed in the 1990s as a way of comparing the overall health and life expectancy of different countries.

We also included the intermediate effects of dietary factors on diet-related diseases through BMI and hypertension. In the formula below, β_{ji} refers to the association between intake of dietary factors and hypertension / BMI in different population group a.

$$IDRI_{as} = \sum_{j=1}^{2} \sum_{k=1}^{m} \left\{ [-In(RR_{ika})] * \left(\beta_{jia} * Intake_{jas}\right) * \left(DALY_{kas} / \sum_{k=1}^{m} DALY_{kas}\right) \right\}$$

In the original study, IDHI only includes 12 cardiometabolic diseases and 11 dietary factors.[10] Study 3 expanded IDHI to also include 15 cancers and 1 more dietary factor of dairy products. Detailed introduction of IDHI is described in the Section V.

4. Proportional hazards models

The proportional hazards regression, also called as the Cox model, investigates the effect of multiple variables upon the time a specified event to happen. When the outcome is death for a cohort of people, the incidence rate, defined as the number of new deaths per population at-risk per unit person-time, vary substantially over time, which is then more commonly called the hazard rate.[51] The Cox model aims to estimate the relative hazard ratio, which is the ratio of the hazard rate between the intervention group and the control group. As a semi-parametric model, the Cox model does not require some probability model to represent survival times and is therefore more robust than parametric methods. In addition, it may accommodate both discrete and continuous measures of event times, and incorporate time-dependent covariates, which may change in value over the course of the observation period.[52] Study 3 investigated the association between all-cause mortality and the IDHI, therefore the Cox model would be the main regression model to be applied in the study. To enhance the robustness of the analysis results, we also applied Poisson model, which is then carefully discussed in the Section V.

5. Food price data

In study 1, we used the food retail prices data from the International Comparison Program (ICP) led by the World Bank.[53] The last round available ICP data in 2017 contains food prices of 716 food and beverage items for 172 countries, including globally standardized and region-specific items collected within 5 regions of Africa, Asia and the Pacific, Latin America and the Caribbean, Western Asia, and Commonwealth of Independent States. Food price data for study 2 were nationally representative time-series data provided by the National Bureau of Statistics with the purpose of building up the national consumer price index (CPI). Retail food prices are collected from different types of outlets, including open markets, supermarkets, neighborhood shops, groceries, shopping centers and other retail outlets. Comparing to the wholesale prices, retail prices are less volatile but a better indicator for the consumer's real expenditures. Detailed information about the food price datasets could be found in the Section V.

6. Food composition data

Multiple food composition tables (FCTs) were applied in the analysis, including the Ethiopian FCT[54], Tanzanian FCT[55], FAO's West African Food Composition Table (WAFC)[56], and the National Nutrient Database for Standard Reference (USDA SR28)[57]. In global studies, we prioritized the usage of USDA SR28 compensated by the WAFC. In the country studies, we primarily applied the nutrient content data from local food composition tables.

7. Nutrition and health data

Regarding the study 3, we highly relied on the data from the National Health and Nutrition Examination Surveys (NHANES), which is conducted by the National Center for Health Statistics (NCHS) of the Centers for Disease Control and Prevention (CDC) in the

US.[58] Starting from 1999, the survey examines a nationally representative sample of about 5,000 individuals each year. The NHANES interview includes demographic, socioeconomic, dietary, and health-related questions. The examination component consists of medical, dental, and physiological measurements, as well as laboratory tests administered by highly trained medical personnel.[58]

NCHS linked various surveys with death certificate records from the National Death Index (NDI). The linked NCHS surveys include but not limited to NHANES, NHANES II, NHANES III. Linkage of the NCHS survey participants with the NDI mortality data provides the opportunity to conduct a vast array of outcome studies designed to investigate the association of a wide variety of health factors with mortality.[59]

SECTION V. ARTICLES

ARTICLE 1: GLOBAL VARIATION IN THE COST OF A NUTRIENT ADEQUATE DIET BY POPULATION GROUP

ABSTRACT

Background: Nutrient deficiencies limit human development, and could be caused by the high cost of locally available foods needed to meet nutrient requirements. In this study, we identify the populations whose nutrient needs are most difficult to meet with existing global food systems.

Methods: Using the most recent collection of global food prices (ICP 2017), we compute least-cost diets within upper and lower bounds for energy and 20 nutrients for healthy population across 20 demographic groups in 172 countries.

Findings: Diet costs per day have a global median of $2.30 (IQR: 1.95-2.75) in 2017, and are highest for males 14-18 y/o at $2.85 [95%CI: 2.83, 2.88]. For females, cost per day is highest during pregnancy and lactation, exceeding cost for adult males 19-50 y/o. Energy-adjusted costs per 1,000 kcal have a global median of $0.94 (IQR: 0.79-1.12), and are higher for females throughout the life course, peaking for adolescent girls 9-13 y/o and women above 70 y/o at $1.17 [95% CI: 1.15, 1.19] and $1.18 [95% CI: 1.17, 1.19]. Meeting needs at least cost uses more animal-source foods in higher income countries, and more plant-source foods in lower income countries. Diet costs are most sensitive to requirements for calcium, iron, zinc and vitamins C and E, and upper bounds on carbohydrates and sodium.

Interpretation: Diets with adequate nutrients are unaffordable for many demographic groups, especially women. Results can help guide agriculture and food policy or transfer programs to support populations at risk of inadequate intake.

Key words: Food Prices, Diet Costs, Food System, Nutrient Adequacy, Malnutrition

INTRODUCTION

Malnutrition is a multifaceted phenomenon that limits human potential, and is often linked to insufficient or excessive intake of essential nutrients. Malnutrition in low-income populations typically involves multiple deficiencies, often from diets high in starchy staples that lack adequate diversity to ensure adequate intake of vitamins and minerals and balance across macronutrients.[60–63] Mortality attributed to micronutrient deficiency has declined in recent decades but remains high,[64] contributing to the development of diseases, deficits in human development and losses in economic productivity[65] with great disparities across populations.[66] High food prices can limit the ability of low-income people to acquire diets with adequate nutrients,[1,2,67,68] and many policies and programs aim to improve diets by bringing nutritious foods within reach.[69] Although it is well known that nutrient deficiencies are most prevalent among children, adolescents, and pregnant and lactating women, the cost and composition of complete diets that meet all nutrient requirements for each population has not yet been quantified globally.

This study uses the cost of nutrient adequacy (CoNA), based on the most affordable set of locally available foods that meet all requirements for essential nutrients in the target population, to identify how variation in nutrient needs affects diet costs around the world. Our use of CoNA builds on a long tradition of least-cost diet calculations, in this case to measure the ability of a food system to deliver nutrients in required amounts for each demographic group.[70] Initial formulations of a least-cost diet used only lower bounds for a few nutrients leading to unpalatable combinations of a few items,[15] but adding and updating nutrient requirements has allowed the method to be widely used for dietary recommendations[15,46] and poverty measurement.[8,70,71] Previous work on diet costs worldwide has focused on the needs of a single representative population group such as adult women[18], or on the cost of adhering to a reference diet recommended for all people[19], with only limited exploration to date of variation in the cost of meeting requirements by age, sex and reproductive status.[11]

In this study we use the latest global food price data from the International Comparison Program (ICP) led by the World Bank to measure the cost per day and cost per calorie of meeting nutrient needs in each country for healthy males and females from childhood and adolescence into adulthood, pregnancy and lactation, and old age. We analyze the composition of these least-cost diets by food groups, revealing how the most affordable foods needed to meet nutrient needs vary among people by age, sex and reproductive status. Finally, we compute the contribution of each nutrient requirement to the minimum cost of meeting all needs, showing which nutrients are most expensive to obtain given foods commonly available in local markets.

METHODS

Food price and composition

Diet cost are based on average retail prices from 172 countries in 2017. These data were released by the ICP in May 2020, and reported by national statistical agencies for standard items drawn from a global list of 193 foods and non-alcoholic beverages that may be sold worldwide, complemented by 523 additional items that may be sold within specific geographic regions. From this list of 716 foods, each country reports a single national-average price for the items that were actually available in 2017 at a nationally representative sample of diverse retail outlets in both rural and urban areas, typically visited every month or quarterly throughout the year.[53] The ICP reports prices for 175 countries and territories, but for this study we dropped three locations with missing income data (Anguilla, Bonaire and Montserrat, whose combined population in 2017 was around 40,000 people). For reporting purposes, we grouped the remaining 172 countries by income level as classified by the World Bank into high, upper-middle, lower-middle and low based on their 2017 gross national income (GNI) per capita.[74] Price data provided in ICP are in local currency units (LCU) per unit of each food as purchased, which we converted into LCU/kg. To determine the edible portion and nutrient content of each item, we matched each item to its entry in the USDA National Nutrient Database for Standard Reference (SR28),[57] complemented by three specialized Food Composition Tables (FCTs) for products not listed in SR28, notably several region-specific fish and meat products as well as whole coconuts and cassava leaves.[75–77] We excluded items that were non-caloric, specialized infant foods, or condiments whose quantity consumed would be negligible, and items whose size or composition was unknown, yielding a total of 545 food items (Figure S1). Each item's price was reported by an average of 40 countries, and each country reported an average of 125 prices, from a maximum of 261 items in Kazakhstan to a minimum of 37 items in Bermuda. For reporting purposes only we categorized items into 6 major food groups, as either: starchy staples; pulses, nuts & seeds; animal-source foods; fruits & vegetables; oils & fats; or sweets & beverages. All countries report prices for at least 1 item in each of these food groups with the exception of El Salvador which reported no prices in the pulses, nuts & seeds group (Table S1).

Calculation of the Cost of Nutrient Adequacy (CoNA)

CoNA is defined here as the least cost diet that meets Dietary Reference Intake (DRI) requirements for 20 nutrients, within energy balance plus upper and lower bounds for 3 macronutrients of protein, lipids and carbohydrates, and 17 micronutrients of calcium, zinc, iron, magnesium, phosphorus, copper, selenium, sodium, vitamin C, thiamin, riboflavin, niacin, vitamin B6, folate, vitamin B12, vitamin A and vitamin E (Table S2). The DRIs are from the U.S. Institute of Medicine,[49] and include upper and lower bounds from the Acceptable Macronutrient Distribution Range (AMDR) for total carbohydrates, protein and fats; the Upper Level (UL) to avoid toxicity for 13 micronutrients; and an upper bound on sodium for Chronic Disease Risk Reduction (CDRR).[78] For micronutrients, the lower bound in our main specification is the Estimated Average Requirement (EAR) to meet the median level of need in each population, which we compare to the Recommended Dietary Allowance (RDA) to meet or exceed the needs of 97.5 percent of each population. For iron and zinc, we applied higher Harmonized Average Requirements (H-ARs)[79] to reflect the lower bioavailability of those nutrients from least cost diets which are composed primarily of starchy staples, leguminous grains and other vegetal foods. Finally, for total estimated energy requirements (EERs) we used an active level of physical activity and the median weights and heights of the World Health Organization (WHO) reference population over the 20 demographic groups, which include 7 age categories (4-8 y/o, 9-13 y/o, 14-18 y/o, 19-30 y/o, 31-50 y/o, 51-70 y/o and 70+ y/o) for males and females plus requirements for pregnant and lactating women in 3 age groups (14-18 y/o, 19-30 y/o and 31-50 y/o).[80]

Mathematically, CoNA is defined as the solution to:

(1) $\min_{cg}\{CoNA_{cg} = \Sigma_i p_{ic} \times q_{icg}\}$, subject to:

(2) $\Sigma_i a_{ie} \times q_{icg} = EER_g$

(3) $\Sigma_i a_{fn} \times q_{icg} \geq LB_{ng}$

(4) $\Sigma_i a_{fn} \times q_{icg} \leq UB_{ng}$

(5) $q_{icg} \geq 0$

This metric identifies the most affordable diet for each demographic group (g) in each country (c), minimizing their total cost ($CoNA_{cg}$) summed over all food and beverage items (i) by selecting that population's quantity of each item (q_{icg}) to meet their estimated energy requirements (EER) as well as lower bounds (LB) or upper bounds (UB) for each nutrient (n), given each item's price in that country (p_{ic}) and each item's energy content (a_{ie}) and nutrient composition (a_{in}). All constraints refer to usual intake, and specify each demographic group's energy balance (EER_g) as well as various kinds of lower or upper bounds for each nutrient (LB_{ng}, UB_{ng}). For equation (3) there are lower bounds on the 3 macronutrients plus 16 micronutrients, namely calcium, copper, iron, magnesium, phosphorous, selenium and zinc plus vitamins A, B6, B12, C and E as well as folate, niacin, riboflavin and thiamin, and for equation (4) there are upper bounds on all nutrients except vitamin B12, riboflavin and thiamin, plus an upper bound on vitamin A in retinol form, and a CDRR upper bound for sodium.

We calculated CoNA for each of the 20 demographic groups in all 172 countries, using linear programming to solve equations (1)-(4) for a total of 3,440 different diets in each case. Computations were first done in nominal LCUs and then converted into 2017 USD using 2017 purchasing power parity (PPP) conversion factors from the World Bank.[53] Results are reported as cost per day, and to compare diet composition on an energy-adjusted basis we also report cost per 1,000 kcal, referring to the cost of diet quality. Our primary specification uses with EARs, reflecting the median level of requirement for a healthy population in each age-sex group. To address heterogeneity in requirements within each demographic group, we also report results for all 34,400 diets using RDAs, reflecting the 97.5th percentile to meet requirements for almost all healthy people in each group (Table S2). To test for statistically significant differences among population groups, we regressed

CoNA and CoNA per 1,000kcal on indicator variables for sex and age groups, and their interaction terms, adjusting for differences in national price levels using an indicator for each country.

Sensitivity of diet costs to nutrient requirements

The definition of CoNA implies that, if each food had only one nutrient, equation (1) would simply add up their cost, and every nutrient listed in equations (2) and (3) would have its own most affordable food in the least-cost diet. In practice, actual foods provide multiple nutrients, and mathematical programming is needed to compute the most cost-effective combination of available items to meet all requirements. At the least-cost solution, the number of foods equals the number of constraints that are binding. The incremental difference in CoNA caused by small changes in each nutrient requirement reveals the degree to which locally available foods complement each other to meet human needs to be met at low cost. The level of incremental cost associated with each nutrient is known as the shadow price of each constraint, measured in terms of the currency units of equation (1) per unit of each nutrient in the relevant constraint from equations (2), (3) or (4). Non-binding constraints have a shadow price of zero. To show variation in the sensitivity of diet costs to each nutrient requirement across demographic groups, we focus on percent changes and report each constraint's shadow price elasticity as defined in Equation (6):

$$(6)\ e_{sp}^{j} = \left| \frac{\%\Delta CoNA^{*}}{\%\Delta DRI_{j}} \right|$$

Here the elasticity (e_{sp}^{j}) of the shadow price (sp) for each constraint (j) is the percent change in the least-cost solution ($\%\Delta CoNA^{*}$) for a one percent change in that requirement ($\%\Delta DRI_{j}$). For clarity these are presented in absolute value terms, as the rise in costs associated with any increase in energy required by equation (2), any increase in nutrient requirements specified in equation (3), or any decrease in the upper-bound limits imposed by equation (4).

Variation in cost by demographic group within countries

To quantify differences in diet costs across demographic groups, we show the full distribution of diet costs for each group using box plots, and then compute the difference in least-cost diets caused by variation in requirements across groups using general linear models (GLM). The regressions focus on the marginal effect of differences in age, sex, pregnancy and lactation, on average across all countries, adjusted for each country's average level of diet costs. Including an indicator variable for each country absorbs all cross-country differences associated with national food systems described in other studies (i.e. fixed effects).[72] This approach also absorbs differences in price reporting, such as the artifactually higher level of CoNA in countries that use shorter food lists thereby omitting items that would have been included in the least-cost diet (Figure S2).

Computational and statistical methods

We conducted linear programing using the lpSolve package in R.[81] The results were visualized using box and whisker plots, presenting median, 25% and 75% quantile, as well as the range of median $\pm 1.58\times IQR/N^{1/2}$.[82] We estimated CoNA and CoNA per 1,000kcal over demographic groups by employing GLM adjusting categorical variables of sex, age groups, and their interaction terms, addressing heterogeneity within each population by controlling for countries' fixed effects. We then presented the trends of food quantities in the least cost diets by the 6 major food groups, and shadow price elasticities of major binding nutrients, across all demographic groups, as well as the differences across country's national income groups based on estimation results by GLMs. We used Stata/SE 15.1 and RStudio Version 1.2.5033 for computation, data processing, statistical analyses, and visualizations.

RESULTS

The cost of nutrient adequate diets around the world

CoNA of all demographic groups across 172 countries in 2017 had a global median of \$2.30 (IQR: 1.95-2.75). Adolescent boys between 14-18 y/o had the highest cost with a median of \$2.71 (IQR: 2.29-3.12), followed by lactating girls between 14-18 y/o with a median of \$2.64 (IQR: 2.29-3.13). Across countries of different income levels, middle-income countries faced the highest CoNA with a median of \$2.56 (IQR: 2.16-2.97) for the upper-middle-income, and of \$2.38 (2.01-2.78) for the lower-middle-income countries (Table S3). When we change the nutrient requirements to meet the needs of 97.5% healthy population (RDAs), the median CoNA increases by 17.8% to \$2.71 (IQR: 2.30, 3.24) while the trends across demographic groups and country income levels generally maintain (Table S4).

Presented in Figure 1, the median CoNA is in most cases above the updated international poverty line of \$2.10 in 2017USD,[83] and in all cases is above \$1.32 per day, which is the portion (63%) that can credibly be reserved for food expenditures.[68] The cost over demographic groups display similar trends across country income levels. The ratio between CoNA and food expenditure per capita increases with country income level; in low-income countries, CoNA is well above 100% of food expenditures per capita per day for almost all population groups, indicating that the nutrient adequate diets are relatively unaffordable in those countries. Globally, people need to spend at least \$0.94 (IQR: 0.79-1.12) and \$1.10 (IQR: 0.94-1.31) per 1,000kcal to meet nutrient adequacy requirements for 50% and 97.5% healthy populations (Table S5 and Table S6). The cost of diet quality over a life course is N-shaped, with the highest costs for adolescents and the elderly who require more nutrient dense diets. Figure 1 also reveals that female groups in all ages face higher costs per calorie compared to their male peers, again reflecting the need for more nutrient dense diets.

Regression results presented in Figure 2 show variation in CoNA caused by differences in requirements between each demographic group, controlling for the average level of diet costs in each country. Among all groups, diet cost per day is estimated to peak at \$2.85 [95% CI: 2.83, 2.88] in 14-18 y/o boys. Lactating girls of 14-18 y/o faces a higher CoNA of \$2.81 [95% CI: 2.79, 2.84]. CoNA of lactating women, adolescents, and adults over 70 y/o are also significantly higher than a 19-30 y/o adult. Younger children between 4-8 y/o have significantly lower CoNA due to lower energy needs (Table S7). Regarding the energy-adjusted cost of diet quality, young adolescent girls of 9-13 y/o and elderly female over 70 y/o have the highest estimates of \$1.17 [95% CI: 1.15, 1.19] and \$1.18 [95% CI: 1.17, 1.19] per 1,000 kcal. All female groups face significantly higher energy-adjusted costs than males in the same age range, and younger people under 18 and adults over 50 have significantly higher CoNA per 1,000kcal than adults aged 19 to 30 (Table S7).

Food selection in the least cost diets

Least cost diets can meet all nutrient requirements using large quantities of starchy staples, complemented by a mix of nutrient-dense foods. On average over all demographic groups, least-cost diets include 429g of the most affordable local starchy staples, providing 1,314kcal per day, or about 44% of total weight and 53% of total dietary energy (Table S8 and S9). To meet all nutrient needs at least cost, the quantities required from diverse types of food vary significantly by demographic group. As shown in Figure 3, least-cost diets for people aged 14-50 include a much larger volume of starchy staples than the diets of younger or older people. Animal-source foods are included in larger quantities for adolescents and the elderly, while significantly more fruits and vegetables are selected to meet nutrient needs among lactating women. These data are global averages. As found in earlier studies,[72] eggs and dairy are more expensive in low- and middle-income countries (LMICs) than in upper-middle and high income countries, so the least-cost diets in LMICS include smaller quantities of those animal-source foods than the global average shown in Figure 4. On average the least cost diets in low-income countries include 137 g/day less animal-source foods than in high-income countries, and 139 g/day more of pulses, nuts and seeds and 118 g/day of fruits and vegetables to meet all nutrient requirements (Table S10).

Sensitivity of diet costs to each nutrient constraint

The nutrient requirements to which diet costs are most sensitive typically include energy balance, lower bounds for calcium, iron, and other micronutrients, and the upper bound on carbohydrates for macronutrient balance (Table S11). Across demographic groups, as shown in Figure 4 we find that the lower bound on calcium is particularly expensive for children, adolescents, and older adults, while meeting iron requirements is particularly expensive for adolescents (especially boys) and non-pregnant and non-lactating females. For adults (especially males), diet costs are more sensitive to total energy requirements, and to the AMDR upper bound on carbohydrates as a source of energy. Other nutrient requirements are more similar in their impact on CoNA across demographic groups. Least-cost diets are slightly more sensitive to zinc requirements among older males, and more sensitive to vitamin C and E requirements among lactating females. Vitamin E plays a somewhat larger role in requirements for older females, and diet costs for children and adolescents are somewhat more sensitive to the sodium CDRR. As noted for Figure 3, these data are global averages. In lower-income countries, the higher cost of dairy products may explain the observation that least-cost diets more sensitive to calcium requirements than in upper-middle and high-income countries; other differences in the cost and composition of available foods make diet costs more sensitive to the carbohydrate AMDR in higher income countries, and more sensitive to the sodium CDRR in lower-income countries (Table S12) .

DISCUSSION

This is the first study to quantify variation in the cost of meeting nutrient requirements by demographic group in the global population, for males and females from childhood to old age including periods of pregnancy and lactation. Diet costs use the most affordable combination of available foods at national average prices in 172 countries observed during 2017, to meet an updated set of global dietary requirements for a healthy population of each age and sex. We find three critical results.

First, we extend previous observations that meeting nutrient needs is often unaffordable for low-income people. We find that even least-cost diets would exceed per-capita income for people below global poverty lines, in every one of our demographic groups and in all regions of the world. The median cost per day across countries is highest for adolescents aged 14-18 at \$2.71 for males, or \$2.44 for females, plus an additional \$0.04 per day during pregnancy and \$0.20/day during lactation. Diet costs are lowest for small children aged 4-

8, at \$1.66 for boys and \$1.64 for girls. These costs exceed what households at the international poverty line can spend on food, which averages of 63% of all expenditure or \$1.32 per day,[68,72] thereby confirming that higher income or lower prices would be needed for individuals to obtain adequate nutrients at any stage of life.

Second, we find significant variation in least-cost diets across demographic groups. Focusing on differences in nutrient density and diet quality, costs per 1,000 kcal are higher for females than males at any age, and highest for adolescent girls aged 9-13 and elderly women above age 70 at \$1.17 and \$1.18 respectively. This is well above the global median diet cost of \$0.94 per 1,000 kcal, due to the need for higher density of the nutrients that are most expensive to obtain from the foods available in each country. The nutrient requirements to which diet costs are most sensitive are calcium and iron, followed by zinc and vitamins C and E, as well as the cost of keeping diets below the upper bound on sodium. Many micronutrient supplementation programs include iron and zinc, and sometimes calcium; our results add evidence to verify that these are indeed nutrients that are particularly out of reach for many people, especially for some groups (e.g. iron for women of reproductive age). Conversely, we do not see vitamin A as a very expensive nutrient (Table S11), even in low-income countries, suggesting that preventing vitamin A deficiencies could be handled well with affordable vitamin A-rich foods.

Third, we find that meeting nutrient needs at least cost requires a different mix of foods for each demographic group, and that this pattern varies across countries. Least-cost diets for adolescents, pregnant and lactating women, and the elderly include more nutrient-dense foods (pulses, nuts & seeds, animal-source foods, and fruits & vegetables) and less starchy staples than for adult males. Greater attention to these foods in agriculture and food systems could help reduce inequities between men and women in the cost of nutritious diets. The mix of foods that provide all required nutrients at lowest cost includes large quantities of dairy and eggs in higher-income countries, but these items are more expensive in low-income countries where least cost diets include 139 g/day less animal source foods than they do in high-income countries, and correspondingly more pulses, nuts and seeds as well as fruits and vegetables. This substitution between food groups to meet DRI requirements of all essential nutrients raises concerns about other aspects of healthy diets as specified in national dietary guidelines, which call for larger quantities of non-starchy food groups that tend to be high cost, as documented elsewhere.[68,73] For example in high income countries, animal-source foods are cheaper sources of nutrients, which may play a role in driving food choices away from consumption of abundant plant source foods as recommended in dietary guidelines.

A major limitation of this study is having only one national average price for each food in each country, and having a list of foods and markets for price collection that may not represent all the items and purchase locations used by low-income people. The ICP provides by far the most comprehensive collection of internationally standardized food price data, but without prices for every food in every location it remains likely that low-income consumers could obtain required nutrients at lower cost than our results. Lower cost options might include less attractive versions of the items included here, or entirely different items such as bushmeat or wild fruits at particular places and times, but use of different local food price datasets with variation within countries provides cost per day estimates of similar magnitude to our results.[68,84]

Our study reveals that the worldwide cost of nutrient adequate diets varies by demographic group but exceeds available income for all kinds of people in extreme poverty. The total costs are highest for adolescents and lactating women, while the cost per calorie is highest for females of all ages, particularly over age 50. We demonstrated that certain food groups or nutrients may be particularly critical to reduce the costs of nutritious diets, including nutrient-dense plant foods (fruits, vegetables, pulses, nuts) in low-income countries, and calcium and iron globally. Our results show that food systems around the world do not provide affordable nutritious diets to all, and can help identify target populations at risk of inadequate intake.

MAIN FIGURES

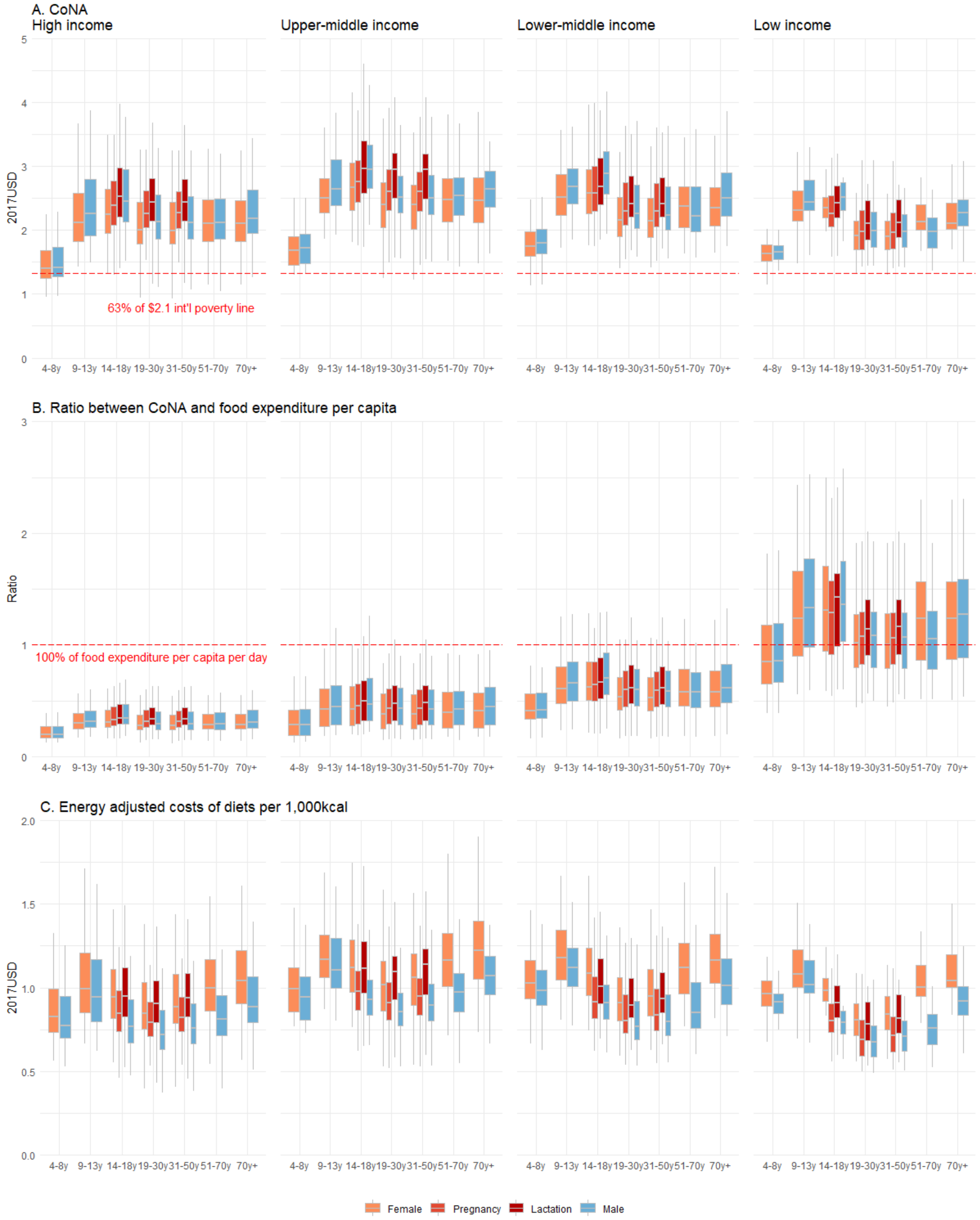


Figure 1: Diet costs for each of 20 demographic groups in 172 countries, by national income category (A) presents the range of cost per day for sufficient quantities of the most affordable foods available in each country to meet median daily nutrient requirements for a healthy population in each demographic group. (B) presents the range of ratios between that least-cost diet and the observed level of food expenditure per capita per day. (C) presents the range of diet costs per 1000 kcal for sufficient quantities of the most affordable foods available in each country to meet median daily nutrient requirements for a healthy population in each demographic group. Box plots show the 25th and 75th percentile of each range around its median, with vertical lines to the median ± 1.58×IQR/(N)1/2. The horizontal dashed line in (A) facilitates comparison to a global poverty line with food expenditure levels of $1.32 per day, corresponding to 63% of $2.10 per day in expenditure on all good and services in 2017 USD. The horizontal line at 1.0 in (B) refers to the country's actual average food expenditure per capita.

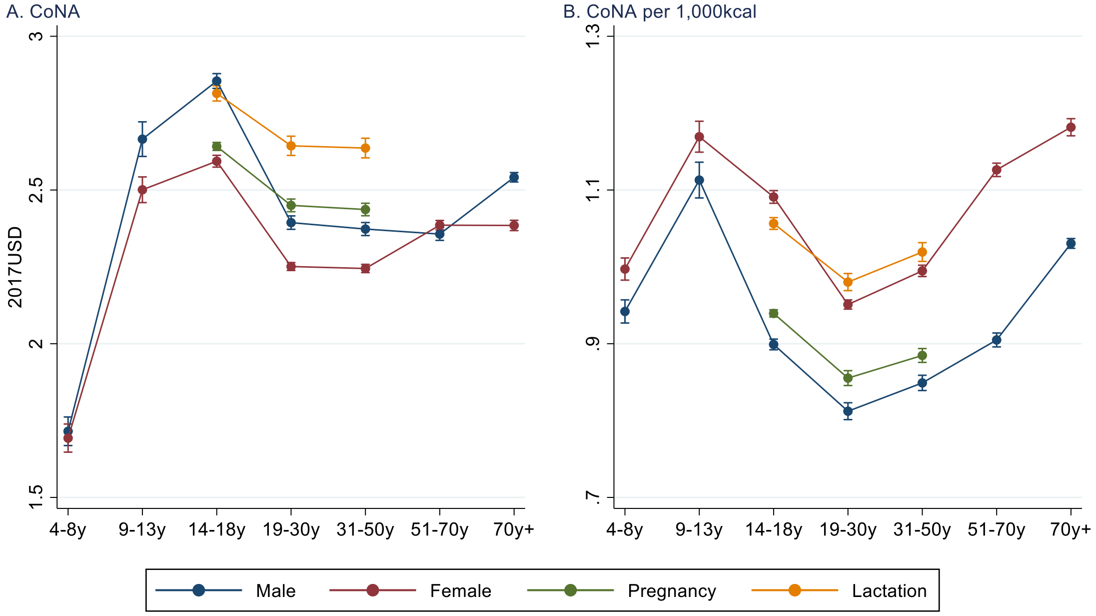


Figure 2: Variation in diet cost among demographic groups
(A) presents cost per day. (B) presents costs per 1,000kcal. All data shown refer to the least-cost diet meeting all nutrient requirements for the median person in a healthy population from each group. Error bars show 95% confidence intervals around the mean for each group, from a regression with 172 country fixed effects that controls for each country's differences in average costs for all demographic groups.

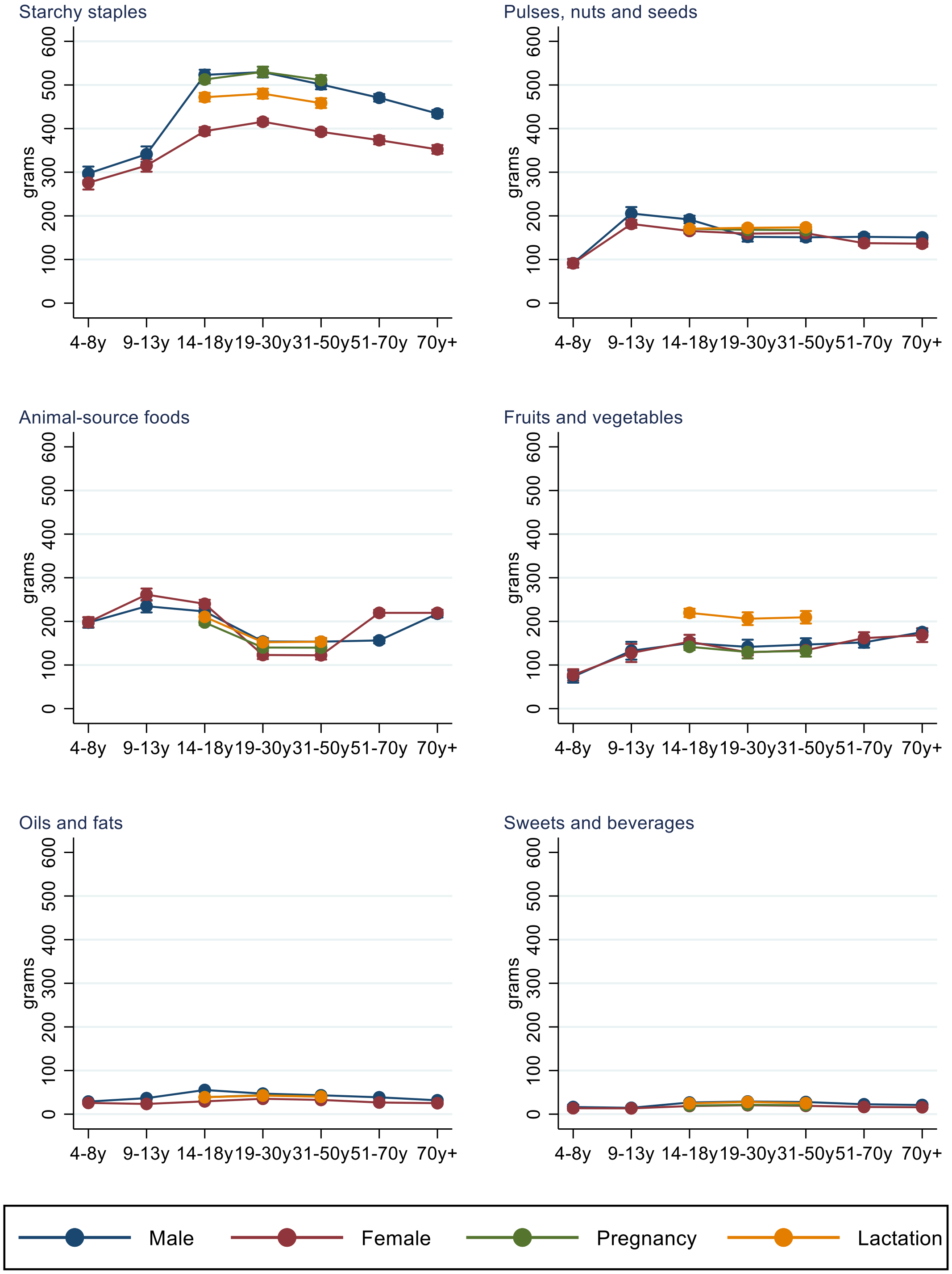


Figure 3: Composition of least-cost diets by food category and demographic group
Results shown are food quantities (in g/day), in least-cost diets meeting all nutrient requirements for the median person in a healthy population from each group. Error bars show 95% confidence intervals around the mean for each group from a regression with 172 country fixed effects that controls for each country's differences in average costs for all demographic groups.

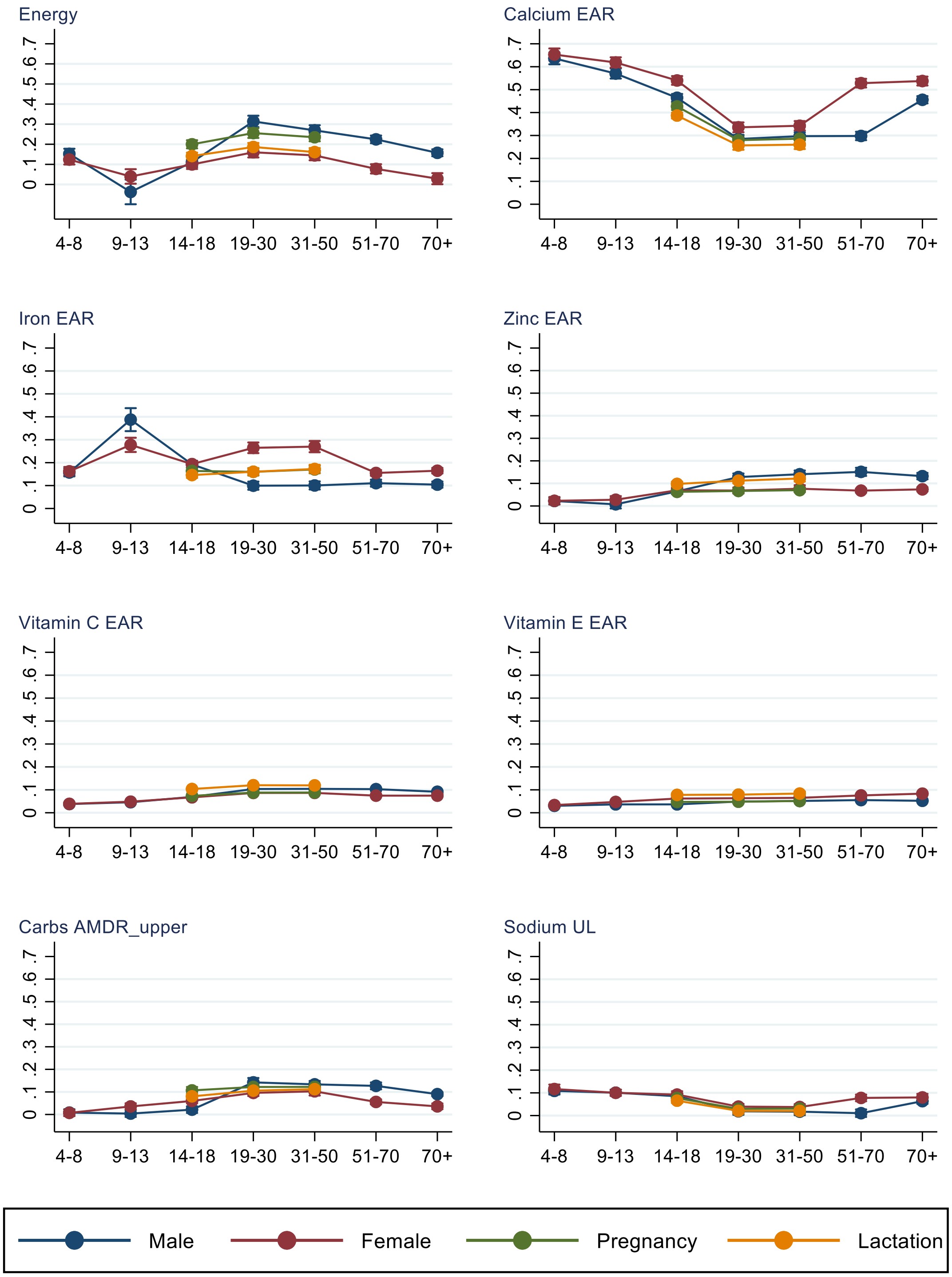


Figure 4: Sensitivity of diet costs to selected nutrient requirements by demographic group
Results shown are shadow price elasticities, defined as the percent change in diet cost for each one percent change in the nutrient requirement shown, for the 8 most important constraints. Diets meet nutrient adequacy requirements for the median person in each demographic group. Error bars show 95% confidence intervals around the mean for each group from a regression with 172 country fixed effects that controls for each country's differences in average costs for all demographic groups.

SUPPLEMENTARY FIGURES

Figure S1: Flow chart of food item selection from the 2017 ICP dataset
From 716 food and non-alcoholic beverage items in the ICP-2017 database, we excluded items that were non-caloric, specialized infant foods, or condiments whose quantity consumed would be negligible, and items whose size or composition was unknown, yielding a total of 545 food items to generate CoNA in 172 countries.

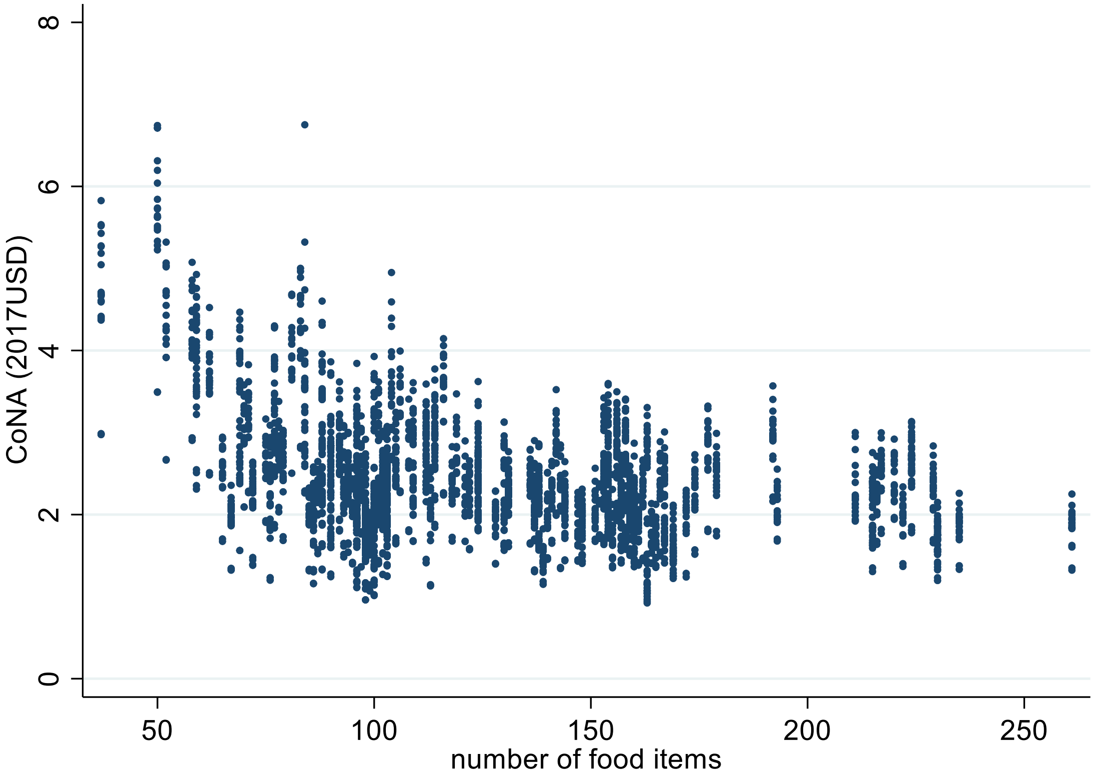


Figure S2: Measured CoNA is higher in countries that report prices for very few foods
CoNA is negatively correlated with the number of food items included in the ICP database for a country, with a Pearson's correlation coefficient of -0.3516 ($P<0.001$) and a Spearman's rho of -0.3101 ($P<0.001$).

SUPPLEMENTARY TABLES

Table S1: Number of food items included in the price dataset for cona generation, by country and food group

N	Territories	Starchy staples	Pulses, nuts and seeds	Animal-source foods	Fruits and vegetables	Oils & fats	Sweets and beverages	Total food items
1	Aruba	22	6	41	23	7	15	114
2	Angola	16	2	33	14	5	7	77
3	Albania	21	4	31	16	6	15	93
4	United Arab Emirates	28	10	61	44	7	15	165
5	Argentina	17	1	26	18	5	8	75
6	Armenia	44	7	75	44	11	48	229
7	Antigua and Barbuda	13	3	23	18	5	8	70
8	Australia	19	4	34	19	7	15	98
9	Austria	22	3	31	19	6	15	96
10	Azerbaijan	37	9	83	40	11	42	222
11	Burundi	28	6	53	24	11	16	138
12	Belgium	22	3	26	19	7	13	90
13	Benin	33	4	65	23	11	17	153
14	Burkina Faso	32	5	63	24	12	18	154
15	Bangladesh	48	8	73	25	14	25	193
16	Bulgaria	23	4	37	19	7	15	105
17	Bahrain	26	8	50	39	6	13	142
18	Bahamas, The	15	1	18	10	3	5	52
19	Bosnia and Herzegovina	21	4	29	19	5	14	92
20	Belarus	40	8	75	40	9	44	216
21	Belize	15	3	31	24	5	12	90
22	Bermuda	7	1	13	6	4	6	37
23	Bolivia	12	2	17	15	5	8	59
24	Brazil	18	3	34	20	6	11	92
25	Barbados	14	4	28	13	5	8	72
26	Brunei Darussalam	32	8	55	30	12	29	166
27	Bhutan	25	4	23	26	5	19	102
28	Botswana	23	6	48	21	9	17	124
29	Central African Republic	28	6	50	19	9	18	130
30	Canada	18	4	31	15	6	14	88
31	Switzerland	23	3	31	19	6	14	96
32	Chile	19	4	30	18	7	12	90
33	China	52	5	88	32	13	30	220
34	Cote d'Ivoire	33	5	65	23	12	19	157
35	Cameroon	32	6	66	24	12	18	158

N	Territories	Starchy staples	Pulses, nuts and seeds	Animal-source foods	Fruits and vegetables	Oils & fats	Sweets and beverages	Total food items
36	Congo, Dem. Rep.	31	6	65	24	12	18	156
37	Congo, Rep.	33	6	67	24	12	18	160
38	Colombia	16	3	24	16	7	13	79
39	Comoros	18	4	46	10	9	17	104
40	Cape Verde	29	6	52	24	8	17	136
41	Costa Rica	12	2	23	11	6	11	65
42	Curacao	23	5	34	21	7	13	103
43	Cayman Islands	23	6	42	25	7	15	118
44	Cyprus	23	4	36	18	7	15	103
45	Czech Republic	21	3	36	18	6	13	97
46	Germany	22	3	31	18	7	13	94
47	Djibouti	29	5	47	20	8	15	124
48	Dominica	17	3	30	20	6	12	88
49	Denmark	22	4	32	18	7	13	96
50	Dominican Republic	21	6	39	25	7	14	112
51	Algeria	24	5	43	17	6	17	112
52	Ecuador	15	3	22	20	5	11	76
53	Egypt	40	12	76	48	12	23	211
54	Spain	23	4	32	19	7	15	100
55	Estonia	21	3	33	18	7	14	96
56	Ethiopia	33	5	57	23	12	17	147
57	Finland	21	2	28	17	5	13	86
58	Fiji	20	4	34	17	5	17	97
59	France	23	4	36	17	6	14	100
60	Gabon	21	6	50	22	10	15	124
61	United Kingdom	23	3	33	18	7	15	99
62	Ghana	31	6	65	24	10	18	154
63	Guinea	27	6	58	24	11	18	144
64	Gambia, The	33	6	66	24	12	19	160
65	Guinea-Bissau	33	6	67	24	12	18	160
66	Equatorial Guinea	33	6	67	24	12	19	161
67	Greece	23	4	35	19	7	15	103
68	Grenada	16	4	15	22	7	13	77
69	Guyana	22	6	40	16	7	15	106
70	Hong Kong SAR	29	5	74	34	10	22	174
71	Honduras	11	3	21	18	3	6	62
72	Croatia	23	3	34	19	7	15	101
73	Haiti	15	3	22	17	4	10	71
74	Hungary	22	4	33	17	6	14	96
75	Indonesia	34	3	70	32	10	28	177

N	Territories	Starchy staples	Pulses, nuts and seeds	Animal-source foods	Fruits and vegetables	Oils & fats	Sweets and beverages	Total food items
76	India	54	10	92	34	14	31	235
77	Ireland	21	4	33	17	7	14	96
78	Iran	25	8	50	36	6	15	140
79	Iraq	31	10	59	44	8	15	167
80	Iceland	19	3	29	18	6	13	88
81	Israel	20	4	35	18	7	14	98
82	Italy	22	4	35	19	7	15	102
83	Jamaica	10	1	23	15	3	6	58
84	Jordan	31	10	61	44	8	15	169
85	Japan	11	3	22	15	6	12	69
86	Kazakhstan	54	10	84	52	12	49	261
87	Kenya	28	6	50	24	11	18	137
88	Kyrgyz Republic	44	9	79	39	11	42	224
89	Cambodia	35	3	59	28	10	21	156
90	St. Kitts and Nevis	21	6	36	24	7	15	109
91	Korea, Rep.	14	3	29	17	6	14	83
92	Kuwait	32	10	57	42	8	15	164
93	Lao PDR	32	3	64	28	4	25	156
94	Liberia	30	6	50	21	7	17	131
95	St. Lucia	20	6	36	25	6	15	108
96	Sri Lanka	28	7	45	29	11	18	138
97	Lesotho	25	4	48	20	8	16	121
98	Lithuania	23	4	33	18	7	13	98
99	Luxembourg	23	3	35	19	7	14	101
100	Latvia	22	4	34	19	7	15	101
101	Morocco	47	13	87	47	11	25	230
102	Moldova	46	9	85	37	11	42	230
103	Madagascar	26	5	45	21	7	15	119
104	Maldives	12	1	14	20	9	13	69
105	Mexico	20	4	33	17	6	14	94
106	North Macedonia	21	3	30	16	5	13	88
107	Mali	33	6	65	24	12	19	159
108	Malta	23	4	35	19	7	15	103
109	Myanmar	53	8	81	33	13	29	217
110	Montenegro	20	3	34	19	5	14	95
111	Mongolia	37	4	38	28	11	26	144
112	Mozambique	33	6	65	24	11	19	158
113	Mauritania	29	6	47	21	10	17	130
114	Mauritius	28	6	58	24	9	18	143
115	Malawi	33	5	59	24	8	19	148
116	Malaysia	27	5	61	31	11	27	162

N	Territories	Starchy staples	Pulses, nuts and seeds	Animal-source foods	Fruits and vegetables	Oils & fats	Sweets and beverages	Total food items
117	Namibia	30	5	48	21	8	16	128
118	Niger	30	5	65	23	11	18	152
119	Nigeria	32	6	64	24	12	19	157
120	Nicaragua	14	2	24	24	5	9	78
121	Netherlands	23	4	31	19	7	14	98
122	Norway	20	2	29	17	6	13	87
123	Nepal	35	7	35	24	7	23	131
124	New Zealand	16	3	24	13	5	11	72
125	Oman	31	10	60	43	8	15	167
126	Pakistan	52	10	79	31	12	31	215
127	Panama	18	2	27	19	7	11	84
128	Peru	23	6	36	23	7	14	109
129	Philippines	33	6	81	31	12	29	192
130	Poland	21	4	34	19	7	15	100
131	Portugal	22	4	33	19	7	14	99
132	Paraguay	18	2	27	20	6	8	81
133	West Bank and Gaza	27	8	55	39	6	13	148
134	Qatar	32	10	59	41	6	15	163
135	Romania	21	4	36	18	6	15	100
136	Russian Federation	18	4	30	16	5	13	86
137	Rwanda	30	5	55	22	10	17	139
138	Saudi Arabia	32	10	60	44	8	15	169
139	Sudan	30	10	56	36	9	13	154
140	Senegal	33	6	67	23	12	18	159
141	Singapore	33	3	67	31	12	26	172
142	Sierra Leone	32	6	60	24	12	19	153
143	El Salvador	10	0	22	8	2	8	50
144	Serbia	22	4	35	19	7	15	102
145	Sao Tome and Principe	28	6	42	24	7	15	122
146	Suriname	24	6	39	25	7	15	116
147	Slovak Republic	20	4	37	19	6	14	100
148	Slovenia	23	4	35	19	7	15	103
149	Sweden	18	4	29	15	6	13	85
150	Eswatini	30	6	52	24	9	17	138
151	Sint Maarten	12	2	21	11	5	8	59
152	Seychelles	33	4	52	22	9	17	137
153	Turks and Caicos Islands	21	5	37	24	6	12	105
154	Chad	33	5	65	23	11	18	155
155	Togo	24	3	44	8	8	11	98

N	Territories	Starchy staples	Pulses, nuts and seeds	Animal-source foods	Fruits and vegetables	Oils & fats	Sweets and beverages	Total food items
156	Thailand	30	3	62	30	10	28	163
157	Tajikistan	42	9	71	31	14	48	215
158	Trinidad and Tobago	24	6	38	25	7	14	114
159	Tunisia	20	6	46	18	7	16	113
160	Turkey	21	4	30	18	7	13	93
161	Taiwan	25	2	64	30	11	26	158
162	Tanzania	32	6	66	24	12	19	159
163	Uganda	33	6	59	23	11	19	151
164	Uruguay	15	5	21	12	6	8	67
165	United States	14	3	30	15	5	9	76
166	Uzbekistan	39	7	68	45	12	53	224
167	St. Vincent and the Grenadines	19	4	32	24	7	10	96
168	British Virgin Islands	19	5	33	25	5	13	100
169	Vietnam	37	5	72	30	9	26	179
170	South Africa	15	2	33	17	5	12	84
171	Zambia	29	5	56	24	9	18	141
172	Zimbabwe	32	6	61	24	11	19	153
	Average	**26**	**5**	**46**	**23**	**8**	**17**	**125**

Table S2: Dietary Reference Intakes applied in the calculation of CoNA

DRIs	Description	Nutrients
EAR	Estimated average requirement to meet the needs of at least 50% of a healthy population in each demographic group[1]	Protein, carbohydrates, calcium, iron, magnesium, phosphorus, zinc, copper, selenium, vitamin C, thiamin, riboflavin, niacin, vitamin B6, folate, vitamin B12, vitamin A, vitamin E
RDA	Recommended dietary allowance to meet the needs of at least 97.5% of a healthy population in each demographic group[1]	Protein, carbohydrates, calcium, iron, magnesium, phosphorus, zinc, copper, selenium, vitamin C, thiamin, riboflavin, niacin, vitamin B6, folate, vitamin B12, vitamin A, vitamin E
H-AR[4]	Harmonized average requirements to meet the needs of at least 50% of a healthy population in each demographic group, in settings with low bioavailability[2]	Iron and zinc
UL[5]	Upper levels to avoid toxicity and pose no risk of adverse health effects in a healthy population[1]	Calcium, iron, phosphorus, zinc, copper, selenium, vitamin C, vitamin B6, vitamin A (retinol)
CDRR	Chronic disease risk reduction level of intake to limit chronic disease risk in a healthy population.[3]	Sodium
AMDR	Acceptable macronutrient distribution ranges within which there is a reduced risk of chronic disease.[1]	Protein, carbohydrates and lipids

Notes:

1. Institute of Medicine (IOM). 2006.

2. Allen et al., 2019

3. National Academies of Sciences, Engineering, and Medicine. 2019.

4. We used the requirement for iron with a low absorption assumption, and for zinc of a "semi-unrefined diet."

5. UL of magnesium, niacin, folate and vitamin E refer to pharmacological agent or supplements, and therefore are excluded from our linear programing constraints

Table S3: CoNA per day over 20 demographic groups and 4 country income levels using EARs/H-ARs

No.	Categories	Countries, n	Mean	Standard Deviation	P25	P50	P75
	Demographic groups						
	Male						
1	4-8y	172	1.72	0.43	1.40	1.66	1.92
2	9-13y	172	2.67	0.81	2.11	2.51	2.94
3	14-18y	172	2.85	0.77	2.29	2.71	3.12
4	19-30y	172	2.39	0.65	1.97	2.26	2.70
5	31-50y	172	2.37	0.65	1.94	2.23	2.67
6	51-70y	172	2.36	0.65	1.92	2.22	2.65
7	70y+	172	2.54	0.67	2.07	2.40	2.88
	Female						
8	4-8y	172	1.69	0.43	1.39	1.64	1.89
9	9-13y	172	2.50	0.72	1.98	2.34	2.78
10	14-18y	172	2.59	0.72	2.08	2.44	2.90
11	19-30y	172	2.25	0.66	1.83	2.12	2.52
12	31-50y	172	2.24	0.67	1.82	2.10	2.50
13	51-70y	172	2.39	0.65	1.92	2.25	2.67
14	70y+	172	2.38	0.66	1.94	2.22	2.66
	Female, pregnancy						
15	14-18y	172	2.64	0.71	2.16	2.48	2.95
16	19-30y	172	2.45	0.67	2.00	2.33	2.76
17	31-50y	172	2.44	0.67	2.00	2.31	2.76
	Female, lactation						
18	14-18y	172	2.81	0.79	2.29	2.64	3.13
19	19-30y	172	2.64	0.77	2.14	2.48	3.00
20	31-50y	172	2.64	0.78	2.13	2.49	2.98
	Country income levels						
1	High income	1,260	2.34	0.80	1.84	2.16	2.63
2	Upper-middle income	880	2.62	0.69	2.16	2.56	2.97
3	Lower-middle income	760	2.50	0.76	2.01	2.38	2.78
4	Low income	540	2.21	0.54	1.83	2.14	2.44
	Total	3,440	2.43	0.74	1.95	2.30	2.75

Note: Data shown are reported in USD at PPP exchange rates for 2017, using EARs and H-ARs (for iron and zinc) to reflect an adequate level of intake for at least 50% of a healthy population.

Table S4: CoNA per day over 20 demographic groups and 4 country income levels using RDAs

No.	Categories	Countries, n	Mean	Standard Deviation	P25	P50	P75
	Demographic groups						
	Male						
1	4-8y	172	2.05	0.50	1.71	1.98	2.29
2	9-13y	172	3.05	0.91	2.44	2.85	3.45
3	14-18y	172	3.26	0.87	2.65	3.05	3.56
4	19-30y	172	2.77	0.74	2.27	2.63	3.10
5	31-50y	172	2.76	0.75	2.26	2.60	3.07
6	51-70y	172	2.75	0.75	2.26	2.60	3.08
7	70y+	172	2.98	0.79	2.46	2.81	3.29
	Female						
8	4-8y	172	2.03	0.51	1.68	1.95	2.30
9	9-13y	172	2.90	0.83	2.35	2.71	3.29
10	14-18y	172	3.02	0.86	2.46	2.82	3.32
11	19-30y	172	2.64	0.77	2.14	2.44	2.90
12	31-50y	172	2.64	0.79	2.14	2.44	2.90
13	51-70y	172	2.84	0.79	2.37	2.62	3.14
14	70y+	172	2.85	0.80	2.38	2.61	3.16
	Female, pregnancy						
15	14-18y	172	3.30	0.92	2.72	3.15	3.61
16	19-30y	172	2.95	0.83	2.41	2.79	3.28
17	31-50y	172	2.95	0.85	2.40	2.78	3.24
	Female, lactation						
18	14-18y	172	3.45	0.96	2.82	3.22	3.84
19	19-30y	172	3.10	0.89	2.53	2.93	3.45
20	31-50y	172	3.10	0.90	2.54	2.95	3.44
	Country income levels						
1	High income	1,260	2.75	0.94	2.16	2.54	3.06
2	Upper-middle income	880	3.08	0.82	2.55	2.95	3.50
3	Lower-middle income	760	2.95	0.92	2.37	2.83	3.29
4	Low income	540	2.70	0.66	2.25	2.60	3.02
	Total	3,440	2.87	0.88	2.30	2.71	3.24

Note: Data shown are reported in USD at PPP exchange rates for 2017, using RDAs to reflect an adequate level of intake for at least 97.5% of a healthy population.

Table S5: CoNA per 1,000kcal over 20 demographic groups and 4 country income levels using EARs and H-ARs

No.	Categories	Countries, n	Mean	Standard Deviation	P25	P50	P75
	Demographic groups						
	Male						
1	4-8y	172	0.94	0.24	0.77	0.91	1.06
2	9-13y	172	1.11	0.34	0.88	1.05	1.23
3	14-18y	172	0.90	0.24	0.72	0.85	0.98
4	19-30y	172	0.81	0.22	0.67	0.77	0.92
5	31-50y	172	0.85	0.23	0.70	0.80	0.95
6	51-70y	172	0.90	0.25	0.74	0.85	1.02
7	70y+	172	1.03	0.27	0.84	0.97	1.17
	Female						
8	4-8y	172	1.00	0.25	0.82	0.96	1.11
9	9-13y	172	1.17	0.34	0.93	1.09	1.30
10	14-18y	172	1.09	0.30	0.87	1.03	1.22
11	19-30y	172	0.95	0.28	0.77	0.90	1.06
12	31-50y	172	0.99	0.30	0.81	0.93	1.11
13	51-70y	172	1.13	0.31	0.91	1.06	1.26
14	70y+	172	1.18	0.33	0.96	1.10	1.32
	Female, pregnancy						
15	14-18y	172	0.94	0.25	0.77	0.88	1.05
16	19-30y	172	0.86	0.23	0.70	0.81	0.97
17	31-50y	172	0.88	0.24	0.72	0.84	1.00
	Female, lactation						
18	14-18y	172	1.06	0.30	0.86	0.99	1.18
19	19-30y	172	0.98	0.29	0.79	0.92	1.11
20	31-50y	172	1.02	0.30	0.82	0.96	1.15
	Country income levels						
1	High income	1,260	0.95	0.31	0.75	0.87	1.07
2	Upper-middle income	880	1.07	0.27	0.88	1.02	1.20
3	Lower-middle income	760	1.02	0.31	0.82	0.97	1.13
4	Low income	540	0.91	0.24	0.74	0.88	1.01
	Total	3,440	0.99	0.30	0.79	0.94	1.12

Note: Data shown are reported in USD at PPP exchange rates for 2017, using EARs and H-ARs (for iron and zinc) to reflect an adequate level of intake for at least 50% of a healthy population.

Table S6: CoNA per 1,000kcal over 20 demographic groups and 4 country income levels using RDAs

No.	Categories	Countries, n	Mean	Standard Deviation	P25	P50	P75
	Demographic groups						
	Male						
1	4-8y	172	1.12	0.28	0.94	1.09	1.26
2	9-13y	172	1.27	0.38	1.02	1.19	1.44
3	14-18y	172	1.03	0.27	0.83	0.96	1.12
4	19-30y	172	0.94	0.25	0.77	0.89	1.05
5	31-50y	172	0.99	0.27	0.81	0.93	1.10
6	51-70y	172	1.06	0.29	0.87	1.00	1.18
7	70y+	172	1.21	0.32	1.00	1.14	1.34
	Female						
8	4-8y	172	1.20	0.30	0.99	1.15	1.36
9	9-13y	172	1.36	0.39	1.10	1.27	1.54
10	14-18y	172	1.27	0.36	1.03	1.19	1.40
11	19-30y	172	1.11	0.33	0.90	1.03	1.23
12	31-50y	172	1.17	0.35	0.95	1.08	1.28
13	51-70y	172	1.34	0.37	1.12	1.24	1.48
14	70y+	172	1.41	0.39	1.18	1.29	1.57
	Female, pregnancy						
15	14-18y	172	1.17	0.33	0.97	1.12	1.28
16	19-30y	172	1.03	0.29	0.84	0.97	1.14
17	31-50y	172	1.07	0.31	0.87	1.01	1.18
	Female, lactation						
18	14-18y	172	1.30	0.36	1.06	1.21	1.44
19	19-30y	172	1.15	0.33	0.94	1.09	1.28
20	31-50y	172	1.20	0.35	0.98	1.14	1.33
	Country income levels						
1	High income	1,260	1.12	0.37	0.88	1.02	1.23
2	Upper-middle income	880	1.25	0.32	1.03	1.19	1.41
3	Lower-middle income	760	1.20	0.37	0.97	1.15	1.34
4	Low income	540	1.11	0.29	0.90	1.07	1.24
	Total	3,440	1.17	0.35	0.94	1.10	1.31

Note: Data shown are reported in USD at PPP exchange rates for 2017, using RDAs to reflect an adequate level of intake for at least 97.5% of a healthy population.

Table S7: Regression results on cost of diet and cost of diet per 1,000kcal across demographic groups

		Cost of diet				Cost of diet per 1,000kcal			
N	Categorical variables	Beta coef	P-value	95% CI		Beta coef	P-value	95% CI	
	Gender								
1	Male	(base)				(base)			
2	Female	(0.14)	0.00	(0.17)	(0.12)	0.14	0.00	0.13	0.15
3	Pregnancy	0.06	0.00	0.03	0.09	0.04	0.00	0.03	0.06
4	Lactation	0.25	0.00	0.21	0.29	0.17	0.00	0.15	0.18
	Age_group								
5	4-8y	(0.68)	0.00	(0.73)	(0.63)	0.13	0.00	0.11	0.15
6	9-13y	0.27	0.00	0.21	0.33	0.30	0.00	0.28	0.33
7	14-18y	0.46	0.00	0.43	0.49	0.09	0.00	0.07	0.10
8	19-30y	(base)				(base)			
9	31-50y	(0.02)	0.18	(0.05)	0.01	0.04	0.00	0.02	0.05
10	51-70y	(0.04)	0.02	(0.07)	(0.01)	0.09	0.00	0.08	0.11
11	70y+	0.15	0.00	0.12	0.17	0.22	0.00	0.21	0.23
	Gender # Age_group								
12	Female#4-8y	0.12	0.00	0.05	0.19	(0.08)	0.00	(0.11)	(0.06)
13	Female#9-13y	(0.02)	0.57	(0.10)	0.05	(0.08)	0.00	(0.12)	(0.05)
14	Female#14-18y	(0.12)	0.00	(0.16)	(0.08)	0.05	0.00	0.04	0.07
15	Female#31-50y	0.01	0.42	(0.02)	0.05	0.01	0.44	(0.01)	0.02
16	Female#51-70y	0.17	0.00	0.14	0.21	0.08	0.00	0.06	0.10
17	Female#70y+	(0.01)	0.43	(0.05)	0.02	0.01	0.18	(0.01)	0.03
18	Pregnancy#14-18y	(0.27)	0.00	(0.31)	(0.23)	(0.00)	0.76	(0.02)	0.01
19	Pregnancy#31-50y	0.01	0.73	(0.03)	0.05	(0.01)	0.47	(0.03)	0.01
20	Lactation#14-18y	(0.29)	0.00	(0.34)	(0.24)	(0.01)	0.27	(0.03)	0.01
21	Lactation#31-50y	0.01	0.62	(0.04)	0.07	0.00	0.84	(0.02)	0.02

Note: Data shown are GLM regression results, adjusted for countries' fixed effects, using CoNA based on EARs and H-AR levels that meet requirements for at least 50% of a healthy population.

Table S8: Composition of least-cost diets by demographic group and income level (g/day)

No.	Categories	Countries, n	Starchy staples	Legumes, nuts and seeds	Animal-source foods	Fruits and vegetables	All fats and oils	Sweets and beverages
	Demographic groups							
	Male							
1	4-8y	172	297	91	197	73	29	16
2	9-13y	172	341	205	235	133	37	15
3	14-18y	172	523	192	223	150	55	27
4	19-30y	172	530	152	154	142	47	29
5	31-50y	172	501	151	154	147	43	28
6	51-70y	172	471	152	156	152	39	23
7	70y+	172	434	151	218	176	32	21
	Female							
8	4-8y	172	276	91	198	77	26	14
9	9-13y	172	315	181	261	128	24	13
10	14-18y	172	394	165	240	153	29	19
11	19-30y	172	416	159	123	129	35	20
12	31-50y	172	393	160	122	134	33	19
13	51-70y	172	374	138	220	162	27	17
14	70y+	172	352	136	220	169	25	16
	Female, pregnancy							
15	14-18y	172	513	170	197	142	39	20
16	19-30y	172	530	168	140	130	44	22
17	31-50y	172	511	167	140	132	41	21
	Female, lactation							
18	14-18y	172	472	171	211	220	39	25
19	19-30y	172	480	173	152	206	43	28
20	31-50y	172	458	174	153	209	40	25
	Country income levels							
1	High income	1,260	443	133	270	99	38	39
2	Upper-middle income	880	465	131	160	158	45	17
3	Lower-middle income	760	433	147	115	170	35	4
4	Low income	540	332	272	132	217	20	8
	Total	3,440	429	157	186	148	36	21

Note: Data shown are simple averages, not weighted by population size, over all demographic groups. Diets are based on EARs and H-ARs that meet requirements for at least 50% of a healthy population.

Table S9: Composition of least-cost diets by demographic group and income level (kcal/day)

No.	Categories	Countries, n	Starchy staples	Legumes, nuts and seeds	Animal-source foods	Fruits and vegetables	All fats and oils	Sweets and beverages
	Demographic groups							
	Male							
1	4-8y	172	957	324	208	33	251	49
2	9-13y	172	1,041	701	248	57	317	31
3	14-18y	172	1,663	683	219	73	480	57
4	19-30y	172	1,720	561	145	69	403	49
5	31-50y	172	1,610	554	146	71	370	45
6	51-70y	172	1,466	554	150	72	330	33
7	70y+	172	1,322	541	221	76	272	34
	Female							
8	4-8y	172	874	319	210	33	223	39
9	9-13y	172	957	632	270	53	202	25
10	14-18y	172	1,191	592	245	65	252	32
11	19-30y	172	1,274	571	127	59	302	34
12	31-50y	172	1,184	575	128	60	279	29
13	51-70y	172	1,082	491	221	67	230	28
14	70y+	172	1,002	485	221	69	217	25
	Female, pregnancy							
15	14-18y	172	1,568	613	196	65	335	35
16	19-30y	172	1,647	609	134	64	374	36
17	31-50y	172	1,566	605	135	63	351	34
	Female, lactation							
18	14-18y	172	1,385	613	207	93	336	31
19	19-30y	172	1,429	627	145	90	367	38
20	31-50y	172	1,346	628	147	91	344	29
	Country income levels							
1	High income	1,260	1,359	468	232	51	335	33
2	Upper-middle income	880	1,336	461	153	80	383	66
3	Lower-middle income	760	1,371	571	177	68	280	10
4	Low income	540	1,062	969	162	84	175	27
	Total	3,440	1,314	564	186	66	312	36

Note: Data shown are simple averages, not weighted by population size, over all demographic groups. Diets are based on EARs and H-ARs that meet requirements for at least 50% of a healthy population.

Table S10: Differences in composition of least-cost diets by national income levels

	Intake (gram)				Energy (kcal)			
Categorical variables	**Beta coef**	**P-value**	**95% CI**		**Beta coef**	**P-value**	**95% CI**	
Starchy staples								
Upper-middle income	21.2	0.03	2.4	39.9	(11.0)	0.63	(55.9)	33.8
Lower-middle income	(10.1)	0.26	(27.5)	7.4	21.2	0.40	(28.5)	71.0
Low income	(111.8)	0.00	(133.5)	(90.2)	(297.2)	0.00	(358.7)	(235.8)
Pulses, nuts and seeds								
Upper-middle income	(1.4)	0.76	(10.2)	7.5	(16.1)	0.30	(46.9)	14.6
Lower-middle income	14.3	0.03	1.8	26.7	96.1	0.00	48.7	143.4
Low income	139.3	0.00	118.4	160.2	500.8	0.00	441.7	560.0
Animal-source foods								
Upper-middle income	(109.9)	0.00	(123.8)	(96.1)	(78.4)	0.00	(88.9)	(67.9)
Lower-middle income	(154.9)	0.00	(168.2)	(141.6)	(65.5)	0.00	(80.3)	(50.6)
Low income	(137.3)	0.00	(154.6)	(120.0)	(69.8)	0.00	(85.5)	(54.1)
Fruits and vegetables								
Upper-middle income	58.9	0.00	45.5	72.3	28.3	0.00	18.7	37.9
Lower-middle income	71.3	0.00	54.2	88.4	12.0	0.01	3.5	20.6
Low income	118.0	0.00	87.5	148.5	32.5	0.00	19.3	45.7
Oils and fats								
Upper-middle income	6.9	0.00	5.2	8.6	46.4	0.00	34.7	58.1
Lower-middle income	(2.7)	0.02	(5.0)	(0.5)	(44.4)	0.00	(59.4)	(29.3)
Low income	(18.2)	0.00	(19.9)	(16.4)	(159.9)	0.00	(175.3)	(144.5)
Sweets and beverages								
Upper-middle income	(22.4)	0.00	(27.6)	(17.3)	30.9	0.00	17.9	44.0
Lower-middle income	(35.7)	0.00	(39.9)	(31.5)	(19.5)	0.00	(26.3)	(12.6)
Low income	(30.9)	0.00	(36.3)	(25.5)	(6.4)	0.29	(18.3)	5.5

Note: Data shown are GLM regression results, adjusted for each of 20 demographic groups, using CoNA based on EARs and H-AR levels that meet requirements for at least 50% of a healthy population.

Table S11: Sensitivity of diet costs to binding nutrient constraints

No.	Nutrients	High income	Upper-middle income	Lower-middle income	Low income	All countries
	Energy and Macronutrients					
1	Energy	0.16	0.18	0.19	0.03	0.15
2	Protein, AMDR, lower	0.00	0.01	0.01	0.00	0.01
3	Lipid, AMDR, lower	0.02	0.02	0.03	0.03	0.03
4	Lipid, AMDR, upper	0.01	0.01	0.01	0.03	0.01
5	Carbs, AMDR, lower	0.00	0.00	0.01	0.04	0.01
6	Carbs, AMDR, upper	0.10	0.10	0.07	0.01	0.08
	Minerals					
7	Calcium, EAR	0.32	0.36	0.49	0.66	0.42
8	Iron, H-AR	0.23	0.18	0.13	0.16	0.18
9	Magnesium, EAR	0.01	0.01	0.01	0.00	0.01
10	Zinc, H-AR	0.05	0.13	0.09	0.04	0.08
11	Copper, UL	0.00	0.00	0.00	0.01	0.00
12	Selenium, UL	0.00	0.00	0.00	0.01	0.00
13	Sodium, UL	0.03	0.07	0.07	0.08	0.06
	Vitamins					
14	Vitamin C, EAR	0.10	0.08	0.07	0.07	0.08
15	Riboflavin, EAR	0.02	0.02	0.01	0.01	0.01
16	Niacin, EAR	0.07	0.03	0.01	0.01	0.04
17	Vitamin B6, EAR	0.03	0.02	0.02	0.00	0.02
18	Folate, EAR	0.01	0.00	0.01	0.00	0.01
19	Vitamin B12, EAR	0.04	0.04	0.01	0.01	0.03
20	Vitamin A, EAR	0.03	0.02	0.02	0.03	0.02
21	Vitamin E, EAR	0.04	0.07	0.07	0.06	0.06

Note: Data shown are simple averages, not weighted by population size, over all demographic groups and countries in each national income category. Diet costs are based on EARs and H-ARs that meet requirements for at least 50% of a healthy population. Sensitivity is measured as the shadow price elasticity of each constraint, defined as the percent change in cost per day for each one percent change in the lower or upper bound, shown here in absolute value of the ratio (for example, a 1 percent increase in energy requirements would raise diet costs by 0.15 percent.) Nutrient constraints whose average shadow price elasticity is below 0.01 for all income levels are not shown.

Table S12: Differences in sensitivity of diet costs to nutrient constraints, by income level

	Shadow price elasticities							
Categorical variables	**Beta coef**	**P-value**	**95% CI**		**Beta coef**	**P-value**	**95% CI**	
	Estimated Energy Requirement				**Zinc, EAR**			
Upper-middle income	0.03	0.04	0.00	0.05	0.08	0.00	0.06	0.09
Lower-middle income	0.04	0.01	0.01	0.06	0.03	0.00	0.02	0.05
Low income	(0.13)	0.00	(0.16)	(0.11)	(0.01)	0.10	(0.02)	0.00
	Carbohydrate, AMDR, Lower bound				**Sodium, CDRR (upper limit)**			
Upper-middle income	0.00	0.55	(0.01)	0.02	0.04	0.00	0.03	0.05
Lower-middle income	(0.02)	0.00	(0.04)	(0.01)	0.04	0.00	0.03	0.05
Low income	(0.09)	0.00	(0.10)	(0.08)	0.05	0.00	0.03	0.06
	Calcium, EAR				**Vitamin C, EAR**			
Upper-middle income	0.04	0.00	0.02	0.06	(0.02)	0.00	(0.02)	(0.02)
Lower-middle income	0.17	0.00	0.15	0.19	(0.03)	0.00	(0.04)	(0.03)
Low income	0.34	0.00	0.32	0.37	(0.03)	0.00	(0.04)	(0.02)
	Iron, H-AR				**Vitamin E, EAR**			
Upper-middle income	(0.05)	0.00	(0.07)	(0.03)	0.03	0.00	0.02	0.04
Lower-middle income	(0.10)	0.00	(0.12)	(0.08)	0.03	0.00	0.02	0.04
Low income	(0.07)	0.00	(0.10)	(0.04)	0.02	0.00	0.01	0.03

Note: Data shown are GLM regression results, adjusted for each of 20 demographic groups, using CoNA based on EARs and H-AR levels that meet requirements for at least 50% of a healthy population. The dependent variable is shadow price elasticities as defined in Table S11, showing results for the eight nutrients to which diet costs are most sensitive.

ARTICLE 2: SEASONALITY OF DIET COSTS REVEALS FOOD SYSTEM PERFORMANCE IN EAST AFRICA

ABSTRACT

Seasonal fluctuations in food prices reflect interactions between climate and society, measuring the degree to which predictable patterns of crop growth and harvest are offset by storage and trade. Previous research on seasonality has found significant impacts but focused on specific commodities. This study accounts for substitution between items to meet nutritional needs, computing seasonal variation in local food environments using monthly retail prices for 191 items across Ethiopia, Malawi and Tanzania from 2002 through 2016. We computed over 25,000 least-cost diets meeting nutrient requirements at each market every month, then measured the magnitude and timing of seasonality in diet costs. We find significant intensity in Malawi, Tanzania and Ethiopia (10.0, 6.3 and 4.0%), driven primarily by synchronized price rises for nutrient-dense foods. Results provide a metric to map nutritional security, pointing to opportunities for more targeted investments to improve the year-round delivery of nutrients.

INTRODUCTION

High food prices limit consumption and harm well-being for low income people.[3,85,86] This study addresses the predictable component of price fluctuations, focusing on recurring seasonal peaks of consistent timing and intensity.[87] All kinds of food price volatility may affect nutrition and health[31,88–90] but seasonality is of particular interest because it measures the degree to which people have improved agriculture and food systems sufficiently to overcome predictable climate fluctuations. Improvements in storage and transport have helped stabilize prices over time[32,33], but there remains significant seasonality in wholesale prices at many market locations in Africa.[30] This study is the first to measure seasonal variation in retail prices across all food groups and diet costs in a way that allows substitution among items to meet nutrient needs.

Our study uses government file data on monthly retail prices and harmonic regression analysis to measure the timing and intensity of seasonality in three East African countries, Tanzania, Malawi and Ethiopia, chosen due to their vulnerability to malnutrition and also variation in geography north and south of the equator, as well as variation in altitude and distance from ocean ports or land transport routes. The inclusion of these three countries is also due to availability of relatively high quality of food price data. From Tanzania we have prices for 61 foods at 21 market locations from 2011 through 2015, in Malawi we have 48 foods at 29 markets from 2007 through 2016, and in Ethiopia prices are for 82 foods at 120 markets from 2002 through 2016. The total number of market-month observations is 3,480 in Malawi, 1,236 in Tanzania, and 20,806 in Ethiopia. Our harmonic model uses sine and cosine functions to estimate smooth, symmetric fluctuations of each item's price or diet costs over time, in this case with one cycle each year reflecting the region's unimodal rainfall (Fig. 1A). The seasonal intensity of price variation is the difference between its annual peak and nadir normalized to a unit-free percentage of the nadir. This approach allows us to measure the magnitude and timing of peaks for different combinations of foods at different locations, using 95% confidence intervals (CI) around the estimated intensity to test for statistical significance.

The prices we use were originally collected to measure inflation for each country's consumer price index, and are repurposed here to track the least-cost sources of 22 essential nutrients in the proportions needed for an active and healthy adult woman. We and others focus on diet costs for women of reproductive age because they are often at risk of malnutrition, with severe consequences for themselves and for child health.[91] To allow for substitution among foods in delivering nutrients, we computed the least-cost combination of foods at each place and time needed to meet all requirements, and compared that to bare subsistence cost of daily energy from starchy staples only.[8,9] Each food list includes a wide variety of nutrient sources, including starchy staples, pulses/nuts/seeds, animal foods, fruits/vegetables, oils/fats, and sweets. Not all foods are available at each market every month, but only 102 of the 25,522 market-months in our study had an insufficient variety of foods to meet all nutrient needs, and all of those were in Ethiopia. After computing least-cost diets we used harmonic regression to extract the seasonal component of variation in cost of nutrients and daily energy at each location, and report differences in timing and intensity as a novel metric of food system performance and vulnerability to climatic fluctuations. Our method would also be useful to identify price anomalies due to disruptions such as armed conflict or disease outbreaks.

Measuring seasonality in the cost of nutrients over all major food groups, allowing for substitution among items as their relative prices change, allows us to compare the ability of local farmers and traders to deliver year-round access to all essential nutrients in the proportions needed by people. This permits us to quantify the nutritional performance of local agroecosystems, distinguishing nutrition security from food security, and identify how each type of food contributes to seasonal variation so as to guide interventions that could improve year-round access to a nutritious diet.

RESULTS

Seasonality of individual food prices

The timing of harvest leads to seasonality in prices at each market location, if not offset by storage and trade with other places. Fig. 1A reveals the national average pattern in rainfall and temperature over each calendar year. Tanzania and Malawi have a cooler dry season approximately from May to October, and Ethiopia, located north of the equator, has a dry season from November to March. We found that these recurring cycles lead to statistically significant seasonality in most food items in all three countries (36 of 61 items in Tanzania, 31 of 48 items in Malawi, and 72 of 82 items in Ethiopia; Table S1-Table S3).

To visualize these data in Fig. 1B we show the estimated seasonal intensity and peak timing for 22 standard items from 6 major food groups. Fruits and vegetables generally have stronger seasonality than other food groups, especially in Malawi. For example, tomatoes have high seasonal intensity of 25.8% (18.7%, 33.3%) in Tanzania, 60.3% (46.1%, 75.9%) in Malawi and 38.7% (31.7%, 46.2%) in Ethiopia. High seasonal intensity was also found in prices of locally representative dark leafy vegetables, notably 12.8% (7.5%, 18.4%) for mchicha (amaranth leaves) in Tanzania, 32.7% (22.2%, 44.2%) for rape leaves and 20.7% (10.7%, 31.6%) for pumpkin leaves in Malawi, and 46.9% (38.0%, 56.4%) for kale in Ethiopia. Potatoes and sweet potatoes also have high seasonality in their prices, while cereal grains and pulses, nuts and seeds have less seasonal fluctuation, and animal sourced foods have little or no seasonality in these data. Seasonal peaks in Tanzania and Malawi were synchronized for starchy staples, pulses/nuts/seeds in the late rainy seasons before harvesting, while fruits and vegetables have diverse price peaks which could help to stabilize diet costs if they offer similar nutrients allowing substitution among them over the course of each year.

Seasonality of diet costs

The ability of local food systems to deliver all nutrients needed for health is revealed by the cost of nutrient adequacy from all foods, which we abbreviate CoNA. We compare that to the cost of caloric adequacy from starchy staples, abbreviated CoCA, which is what would be needed for bare subsistence at each location every month. National average levels of CoNA over the period of observation were TZS 912.1 ($1.50 in 2011 USD at PPP prices) in Tanzania, MWK 129.6 ($1.21) in Malawi, and ETB 6.74 ($1.34) in Ethiopia. These costs were 2.41, 3.11 and 3.49 times the country's average level of CoCA required for subsistence (Table S4-S6).

Seasonal fluctuation in the overall cost of all nutrients is large and statistically significant. As shown in Fig. 2A and Table S4-S6, intensity was much stronger in Malawi with seasonal intensity of 10.0% (5.7%, 14.6%), compared to 6.3% (3.7%, 9.0%) in Tanzania and 4.0% (2.5%, 5.5%) in Ethiopia. Seasonal intensity in CoCA was significant in all three countries. The intensity was strongest at 13.9% (12.2%, 15.6%) in Ethiopia, and 8.0% (1.5%, 14.9%) in Malawi and 5.9% (0.8%, 11.3%) in Tanzania. The premium for nutrients above dietary energy, measured by the gap between CoNA and CoCA, also has significant seasonality with intensity of 6.3% (2.4%, 10.4%) in Tanzania, 9.0% (2.7%, 15.6%) in Malawi and 5.3% (3.8%, 6.8%) in Ethiopia.

Peak timings of the three indicators in Malawi and Tanzania were estimated to be about 3 months before the harvest season starting in May. In Ethiopia, although CoCA was estimated to peak in late August, which is about 2 months before the start of harvest season in November, CoNA and CoNA premium peaked earlier, respectively in late July and mid-April. The timing and magnitude of these peaks reflect the limited degree to which different foods can substitute for each other to deliver all required nutrients around the year. As shown in Fig. 3A, the cost of each food group in a least-cost diet varies over time, with high levels of overall seasonality in Malawi driven by its seasonality in fruit and vegetable prices. A different view of these substitutions is presented in Fig. 4, as each food group's contribution of total calories which has significant seasonality in Malawi and Ethiopia but not in Tanzania. In Malawi, energy intake from starchy staples in CoNA becomes minimum before the harvest season, and therefore more energy from fruits and vegetables, animal foods, and sweets. Fig. 3B also reveals time trends in CoNA, for which the national averages increased from $1.31 to $1.56 over the 2011-2015 period in Tanzania, from $0.96 to $1.46 over the 2007- 2016 period in Malawi, and from $1.04 to $1.68 over the 2002-2016 period in Ethiopia. In both Malawi and Tanzania, seasonality in the cost of fruits and vegetables, contributed the most in the seasonality of CoNA (Fig. 3A) although fruits and vegetables do not take a large portion in total cost or energy of CoNA (Fig. 2B, Fig. 4B and Table S4-S6).

Seasonal intensity in CoNA also presents great regional variations within countries. Regional results are shown in Fig. 5 and Fig. 6, where 12 of 21 regions in Tanzania, 14 of 25 districts in Malawi and 27 of 57 zones in Ethiopia showed significant results. In Tanzania, inland region of Singida and west border region of Kigoma showed strong seasonality in CoNA with intensity of 24.7% (8.2%, 43.7%) and 18.2% (9.9%, 27.2%). In Malawi, 5 districts suffered severe seasonality with intensity more than 20%, among which the Dowa district, close to the capital of Lilongwe City, showed seasonal intensity of 35.2% (15.5%, 58.2%) and its peak timing was estimated approximately one month earlier than the national estimation. In Ethiopia, three zones had unusual higher seasonality in CoNA than the rest of the country, which are Kemashi with intensity of 25.2% (16.9%, 34.1%) and Agnuak with intensity of 27.7% (3.4%, 57.8%) on the west borders to Sudan and South Sudan, as well as Yem, a special woreda in the Southern Nations, Nationalities, and Peoples' Region (SNNPR), with seasonal intensity of 36.7% (22.7%, 52.3%). Finally, we note the role of variation in individual dietary requirements which affects the level of cost but has little effect on seasonality. For example, a higher level of physical activity would require 12% more daily energy which raised CoNA by about 4% but led to negligible differences in the timing or intensity of seasonality.

DISCUSSION

This paper introduced a novel combination of techniques to characterize spatio-temporal variation in food prices across three countries in East Africa, measuring the ability of local farmers and traders to achieve year-round delivery of all essential nutrients at low cost despite climatic fluctuations. We used government file data on a total of 191 items at 170 locations in various years from 2002 through 2016, solved for the least-cost combination of foods needed to meet requirements for 22 essential nutrients at each of 25,522 market-months, and then applied harmonic regression to estimate seasonal intensity and peak timing of diet costs at each location. Three important findings were discovered:

First, most individual foods have significant seasonality in retail prices, extending previous observations about major commodities to all food categories. Fruits and vegetables have the largest seasonal price variations, which averages over 20% for 7 of 21 items in Tanzania, 14 of 17 items in Malawi and 8 of 24 items in Ethiopia. Items such as carrots, mangoes, papaya, oranges, avocado, tomatoes, green peppers and onions are important not only for the essential nutrients they provide, but also for other aspects of diet quality and local livelihoods. Foods that are more easily stored and transported, such as cereal grains and pulses, nuts or seeds, have lower levels of seasonality than the highly perishable fruits and vegetables. We also find that seasonality in the prices of widely traded grains is lower on retail markets than previous studies had found in wholesale prices on commodity markets[30], implying that wholesale-to-retail margins help stabilize consumer prices. Nonetheless peak times for various food groups tend to be synchronized before harvests in all three countries, limiting year-round access to all essential nutrients.

Next, even after allowing for substitution among foods, overall diet costs using the least-cost sources of nutrients and energy fluctuate seasonally in ways that are statistically and nutritionally significant. Substitution away from fruits and vegetables worsens diet quality during the lean season [92], and we find that scarcity of nutrient-dense foods typically precedes scarcity of calories from starchy staples as the peak timing for cost of nutrient adequacy is earlier than the peak for cost of caloric adequacy. We also find large regional variation in the seasonality of diet costs, revealing how local food systems differ in their ability to deliver low-cost nutrients around the year. Reducing and stabilizing the cost of acquiring a nutritious diet is important not only for those who buy all their food, but also for farmers who use markets to complement what they grow. Purchased foods from local markets contribute substantially to the diets of agricultural households in Africa [93], and are especially important in lean seasons and for diet diversity beyond what can be produced and stored on the household's own farm.[94,95]

Our third major finding is that prices for animal sourced foods had the least seasonality. This is one reason why the cost of nutrient adequacy had less seasonality in Tanzania and Ethiopia than in Malawi, since their least-cost nutrient sources included more animal products. Overall, these findings point to opportunities for further improvement in low-cost, relatively stable supplies of animal sourced foods, in addition to improvements in market access that would help people overcome seasonality in local production of plant-based foods.

Our analysis reveals the potential for high-frequency, high-density price observations to reveal the ability of local agroecosystems and food markets to deliver nutritionally complete diets at low cost, using data on food composition to compute the least-cost combination of foods that meet all essential nutrient requirements at each time and place. Protocols and software tools to automate the computation of least-cost diets allow us to extract nutrient costs from food price data over a total of 25,522 market-months, thereby measuring food system performance in ways that directly inform efforts to improve year-round access to nutritious diets in both rural and urban areas. Future studies may apply this method to identify the causes of differences in seasonality including local agricultural calendars, trade opportunities and storage costs interacting with consumer demand, affecting both peaks in diet costs that harm consumers as well as seasonal lows that affect farm income and farming-dependent populations.

One key limitation of these analyses is that governments may not collect prices for all foods that could be low-cost sources of essential nutrients, at the times and locations where they are needed by people at risk of malnutrition. Other limitations include variation in the nutrient composition of each food especially after cooking, variation in peoples' nutrient requirements, and variation in retail prices within the month at each market, all of which are subject to further research. Finally, our measure of seasonality in this paper is limited to harmonic fluctuations which is just one component of all variation. Future work could address different kinds of price differences, and identify ways to improve agricultural production, storage and transport to stabilize diet costs and improve year-round affordability of nutritious diets.

DATA AND METHODS

Data sources

The food prices used in this study are historical file data provided by national statistical services in each country. Prices were originally collected for the purpose of measuring inflation using a consumer price index (CPI), based on a list of all goods and services needed to represent national average per capita consumption in that country over an entire year. Since individuals can substitute foods seasonally, and observed diets may not actually meet their nutritional needs, to address the impact of climate fluctuations on cost of nutrients we link food prices with the nutrient composition of each item, and model the least-cost combination of foods needed to meet human requirements of each nutrient.

For Tanzania, the National Bureau of Statistics collected monthly retail food prices of 71 food and non-alcoholic beverage items from all 21 regions of mainland Tanzania between January 2011 and December 2015. Prices data are collected from different types of outlets, including open markets, supermarkets, neighborhood shops, groceries, shopping centers and other retail outlets. The monthly price surveys are conducted in urban regional headquarters in all 21 regions in approximately four outlets per item. For non-processed food items, price collectors go to the shops/markets on three consecutive days for price collection, and retain the median of those three observations.

In Malawi, the National Statistical Office assembled monthly price data for 55 food items in 29 market locations across 25 administrative districts between January 2007 and December 2016. Unlike the Tanzania dataset, all 29 markets in this dataset are in rural towns, 17 of which are the district capitals known as “boma” markets, and the remaining 12 are in other towns. The data are collected during the first two weeks of each month usually from three retail shops preselected by NSO or vendors subject to data collectors’ judgement. They retain the geometric mean of the three observations. To reduce the disproportional effect of extreme values on model results we have winsorized outliers beyond the top 1% of all ratios between reported price and the median for each item, replacing those outliers with the cutpoint value for that item.

The Ethiopia price were obtained from the Consumer Price Survey, collected by the Central Statistical Agency (CSA). Monthly retail food prices considered in this study cover 97 food items in 120 markets from 57 zones of 11 administrative regions between January 2002 and December 2016 of 15 years. Like the dataset in Tanzania, the surveys are conducted in towns and cities. To ensure the survey to be national representative, the CSA also assigns the number of markets in each region to be proportional to the region's share of total urban population in Ethiopia. CSA enumerators collect three price quotations from traders, retailers and consumers in the first 15 days of each month. They retain the median of those three prices, and prior to our receipt of the data also trimmed outliers below the 1st and above the 99th percentile of each item.

After assembling each country's archival price data, we converted their units of measure to local currency per kilogram of edible matter (LCU/kg), and matched the item's description to entries in local food composition tables (FCTs)[56,96,97] where available. To fill gaps where no local composition data is available we used the USDA National Nutrient Database for Standard Reference (SR28).[57] For data visualization and analysis, we also converted food prices to LCU per 100kcal, and classified foods based on an adjusted form of the Minimum Dietary Diversity for Women (MDD-W) guidelines[37] into 6 major mutually exclusive food groups as: 1) Grains, white roots and tubers, and plantains ("Starchy staples"); 2) Pulses, nuts and seeds; 3) Dairy and eggs, meat, poultry and fish ("Animal foods"); 4) Fruits and vegetables; 5) Oils and fats; and 6) Sweets. Finally, we dropped food items that have nutrients but would not be included in significant quantities for adult meal plans such as infant foods and condiments. There are finally 61, 48 and 82 food items respectively included in the analysis for Tanzania, Malawi and Ethiopia, representing all 6 major food groups. Descriptive statistics and numerical results are reported in the annex of extended data.

Computation of least-cost diets

To identify the most affordable sources of all essential nutrients, we automate the computation of least-cost diets at every time and place using linear programming approaches that were originally formulated to solve this and related problems during the second world war.[15] With each food's market price and nutrient composition as fixed parameters, we obtain the quantity of each food that delivers all nutrients within fixed lower and upper bounds at lowest total cost. This least-cost diet for all nutrients is defined as the solution to:

$\min\{C = \Sigma_i p_i \times q_i\}$, subject to 6 kinds of constraint:

(1) $\Sigma_i a_{ij} \times qi \geq EAR_j$

(2) $\Sigma_i a_{ij} \times qi \leq UL_j$

(3) $\Sigma_i a_{ij} \times qi \leq AMDR_{j,upper} \times E / e_j$

(4) $\Sigma_i a_{ij} \times qi \geq AMDR_{j,lower} \times E / e_j$

(5) $\Sigma_i a_{ie} \times q_i = E$

(6) $q_1 \geq 0, q_2 \geq 0, q_3 \geq 0, \ldots, q_i \geq 0$

The objective is lowest diet cost given the price of each food (p_i), choosing quantities (q_i) to meet or exceed the population's estimated average requirement (EAR) for nutrient j given the quantity of nutrient j in each food n_{ij}, within the further constraint of overall estimated energy needs (E), while remaining below upper levels (UL) for most micronutrients and the chronic disease risk reduction (CDRR) upper bound for sodium, and within a range for macronutrients determined by acceptable macronutrient distribution ranges ($AMDR_{lower}$ and $AMDR_{upper}$) as percentages of daily energy needs (E). The reference number e_j is the energy density of macronutrients, which is 4 kcal per gram of protein and carbohydrate and 9 kcal per gram of lipid. In the analysis, we included 22 nutrients, including 3 macronutrients of protein, fat, carbohydrate, 8 minerals of calcium, iron, magnesium, phosphorous, zinc, copper and selenium, sodium and 10 vitamins of vitamin C, thiamin, riboflavin, niacin, vitamin B6, folate, vitamin B12, vitamin A, retinol and vitamin E. Using this same framework we also computed the cost of caloric adequacy for daily subsistence, using only starchy staples to meet the constraint of energy needs along.

All the dietary reference intakes (DRIs) applied in our analysis include the most updated EAR, UL, AMDR and estimated energy requirement developed by the US Institute of Medicine[98], and we used healthy, not pregnant and lactating women of 57kg and 163cm between 19 and 30 years old with low active physical activity level as reference population group. EAR is the amount of nutrient intake value meeting the requirement of half healthy population. For nutrients other than sodium, the upper limit indicates the UL, which is the highest level of daily nutrient intake that is likely to pose no risk of adverse health effects for general population; for sodium, we used the Chronic Disease Risk Reduction Intake (CDRR) developed in 2019 as the UL considering the beneficial effect of reducing sodium intake on cardiovascular disease risk, hypertension risk, systolic blood pressure, and diastolic blood pressure.[99] The AMDR provide a range of intakes for macronutrients that is associated with reduced risk of chronic disease.

To automate computations, we call the lpSolve package in R[81] to return solutions for each location every month. Those computations are done in nominal local currency terms to reflect choices at each place and time. Then, for comparison over time and across countries we converted each diet cost into constant US dollars using 2011 purchasing power parity (PPP) exchange rates provided by the World Bank.[100] Since local inflation occurs from month to month but PPP conversion factors are reported for each calendar year, we smooth over 12 months using the least-squares technique as implemented in Stata using the -denton- command.[101]

Measurement of seasonality

We extracted the magnitude and timing of seasonal fluctuations using harmonic regression, also known as a trigonometric model. This approach uses sine and cosine functions over time, offering a parsimonious representation using just two parameters to estimate smooth, symmetric rise and fall of a variable. The harmonic approach has been shown to be more efficient than traditional monthly indicator models that estimate one coefficient for each month, and the harmonic form offers a closer fit for many seasonal patterns than other functional forms.[30] The model specification is shown below:

$$\ln(C_{kt}) = \beta_0 + \beta_s \times \sin(2\pi\omega t) + \beta_c \times \cos(2\pi\omega t) + \beta_T \times T(t) + \beta_y \times Y_t \tag{1}$$

where C_{kt} is the monthly time series of food price or diet cost, in market k at month t. Coefficients of *sin* and *cos* terms, β_s and β_c, measure the magnitude (A) and peak timing (P) of seasonality where ω is a constant equal to 1/12, indicating 12 months per annual cycle. T(t) is a cubic polynomial term of t, controlling the trend of time series. Y_t controls the fixed effect of crop years. In Tanzania and Malawi, the first month of a crop year is May, while it is October in Ethiopia.[102]

Seasonal intensity is defined as the difference between annual peak and nadir prices normalized to a unit-free percentage of the nadir price, expressed as $\exp\{(2A) - 1\}$, where A is the amplitude of the seasonality. Therefore, the seasonality is comparable across different food items, price indicators, countries, and over time. The estimates of amplitude (A) and peak timing (P) and their variances are calculated using the δ-method and equations below[87]:

$$A = \delta\sqrt{\beta_s^2 + \beta_c^2}, \text{ where } \delta=1, \text{ if } \beta_c > 0, \text{ and } \delta=-1, \text{ if } \beta_c < 0 \text{ and} \tag{2}$$

$$\mathrm{Var(A)} = (\sigma_s^2\beta_s^2 + \sigma_c^2\beta_c^2 + 2\sigma_{sc}\beta_s\beta_c)/(\beta_s^2 + \beta_c^2) \tag{3}$$

$$P = \frac{12\left(1-\frac{\varphi}{\pi}\right)}{2}, where\ \varphi = -\arctan(\beta_s/\beta_c) \tag{4}$$

$$\mathrm{Var}(\varphi) = (\sigma_s^2\beta_c^2 + \sigma_c^2\beta_s^2 - 2\sigma_{sc}\beta_s\beta_c)/(\beta_s^2 + \beta_c^2)^2 \tag{5}$$

where σ_s, σ_c and σ_{sc} are the standard deviations of β_s and β_c parameters, and their joint covariance. We also calculated the 95% confidence intervals (CI) for A and P using a standard constant from a t-distribution of 1.96. The 95% CI of the amplitude is from $A - 1.96\sqrt{Var(A)}$ to $A + 1.96\sqrt{Var(A)}$. The harmonic regression models allow for assessing the significance of seasonal components, e.g. the significance of β_s and/or β_c parameters (for sin and cos terms, respectively). Thus, the peak timing estimates can be formally compared. If the 95% CI does not contain the value of zero, seasonality will be determined significant.[103]

For the seasonality analysis of energy intake compositions and cost components of CoNA contributed by different food groups, we applied a different harmonic model specification:

$$\mathrm{I_{kt}} = \beta_0 + \beta_s\times\sin(2\pi\omega t) + \beta_c\times\cos(2\pi\omega t) + \beta_T\times T(t) + \beta_y\times Y_t \tag{6}$$

where $\mathrm{I_{kt}}$ is the energy intake compositions of CoNA in kcal and the cost components of CoNA from each food group in k-market and time t. In this analysis, seasonal intensity is defined as the average absolute difference between the peak and nadir values in a yearly cycle, or simply the double of amplitude, $2A$, estimated from the model (6).

We compared results from both harmonic regression and traditional monthly indicator models, for diet costs and individual food items across three countries, and the comparison results are shown in Fig. S1-S4. The model specification for harmonic model followed the equation (6) above, and the specification for dummy variable approach is equation (7) and (8) as below:

$$I_{kt} = \beta_0 + \Sigma_m \beta_m \times M_m + \beta_T \times T(t) \tag{7}$$

$$\ln(C_{kt}) = \beta_0 + \Sigma_m \beta_m \times M_m + \beta_T \times T(t) \tag{8}$$

where I_{kt} is the diet costs or food prices in the k-th market at month t. Mm is the dummy variable for calendar months, and we selected November as the base month in the analyses.

We employed multivariate mixed effect model in estimations where observations are from multiple markets, with random intercepts and coefficients on the seasonal terms (sin and cos terms) by markets. If an estimation was based on observations from single market, ordinary least squares (OLS) model was applied instead. All regression models were run in Stata/SE 15.1.

FIGURES AND TABLES

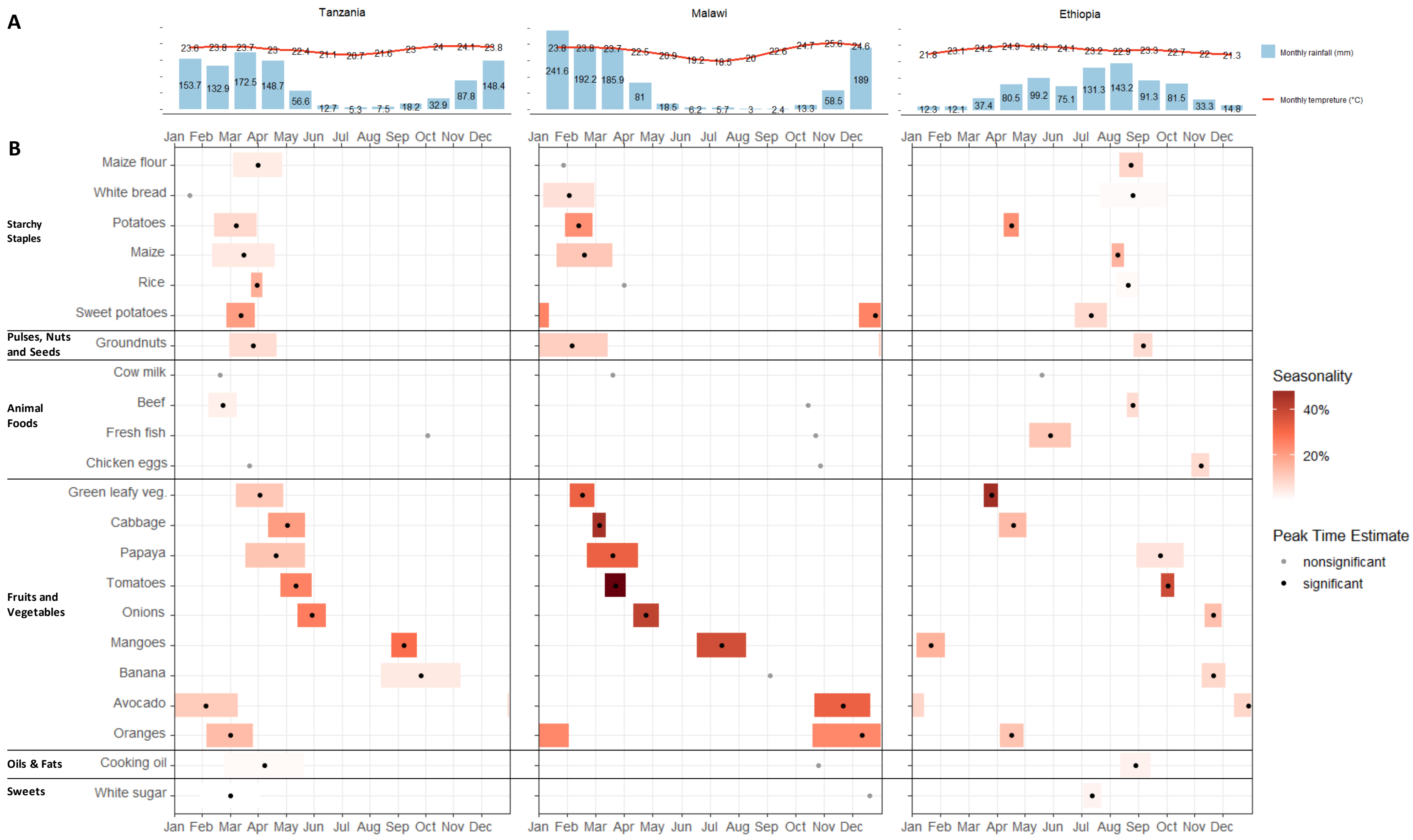


Fig. 1. Intensity and timing of seasonality in market prices for commonly consumed foods in Tanzania, Malawi and Ethiopia. **(A)** shows the national average monthly rainfall (mm) and temperature (°C) between 1991 and 2016 (32). **(B)** shows 95% confidence intervals around the peak month for each food, shown as a black dot, with the magnitude of intensity shown by the color gradation of each bar. Grey dots show the peak month for foods without statistically significant harmonic seasonality. Price variation is estimated from data in local currency units (LCU) per item, on average over all market locations in each country shown.

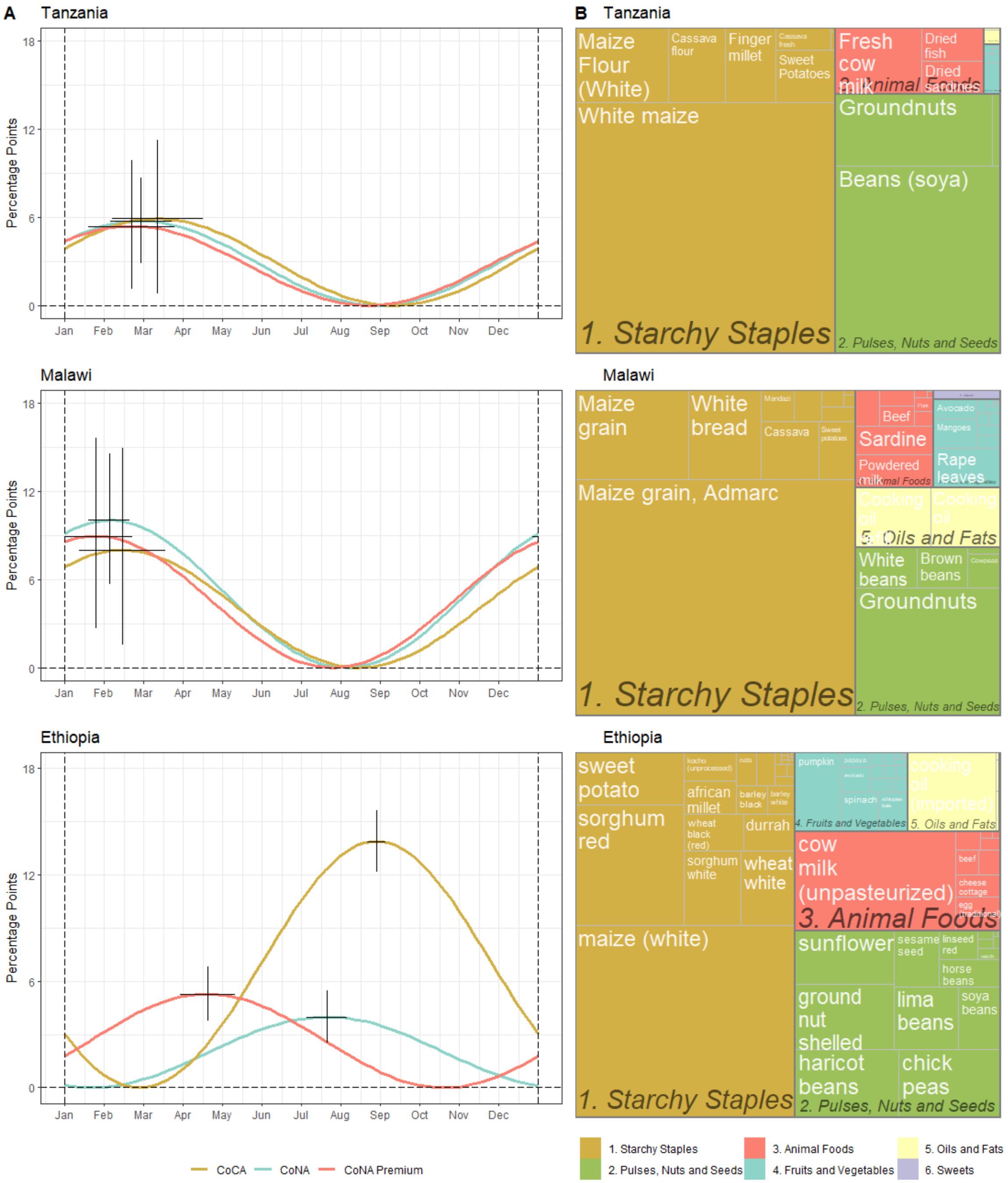


Fig. 2. Seasonality and composition of diet costs for nutrient adequacy (CoNA) vs the cost of caloric adequacy (CoCA)

(A) displays estimated harmonic seasonality over a one-year cycle for the three indicators with error bars showing 95% confidence intervals around the magnitude of seasonal intensity along the vertical axis and peak month along the horizontal axis. **(B)** shows the average energy composition by food group and item of the least-cost diet selected for CoNA over all observations in each country. CoCA is a least-cost diet that meets energy needs using only starchy staples. The CoNA premium is the cost of meeting nutrient requirements beyond daily energy, defined as CoNA-CoCA, in local currency units per day.

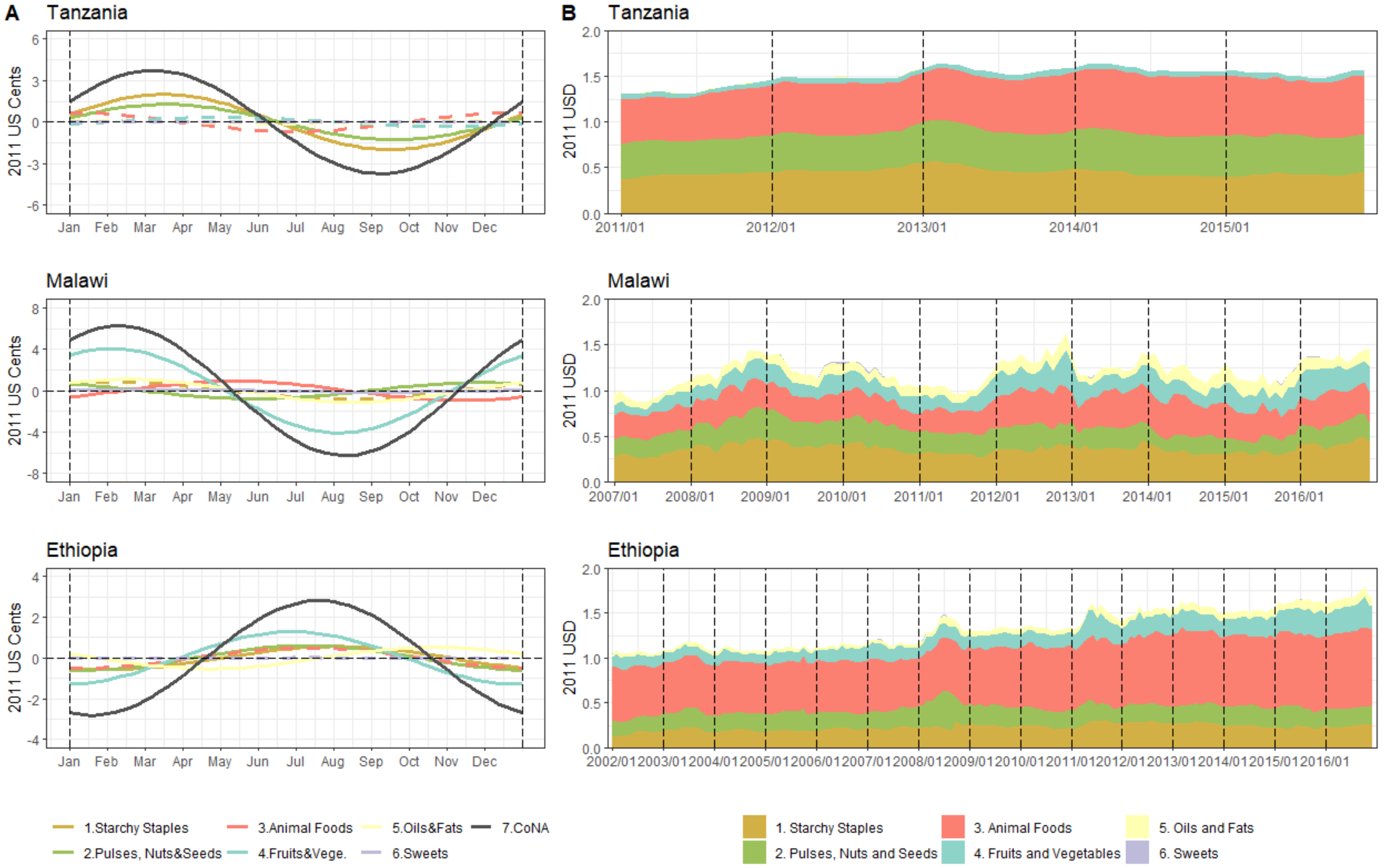


Fig. 3. Seasonality in diet costs by food group over time
(A) displays estimated harmonic seasonality over a one-year cycle for the overall cost of nutrient adequacy (CoNA) and for the selected components of that diet from each of six food groups. Dashed lines are not statistically significantly different from zero. **(B)** shows the contribution of each food group to the cost of nutrient adequacy (CoNA) each month, averaged over all marketplaces in the country shown. Diet costs are converted to USD at PPP exchange rates.

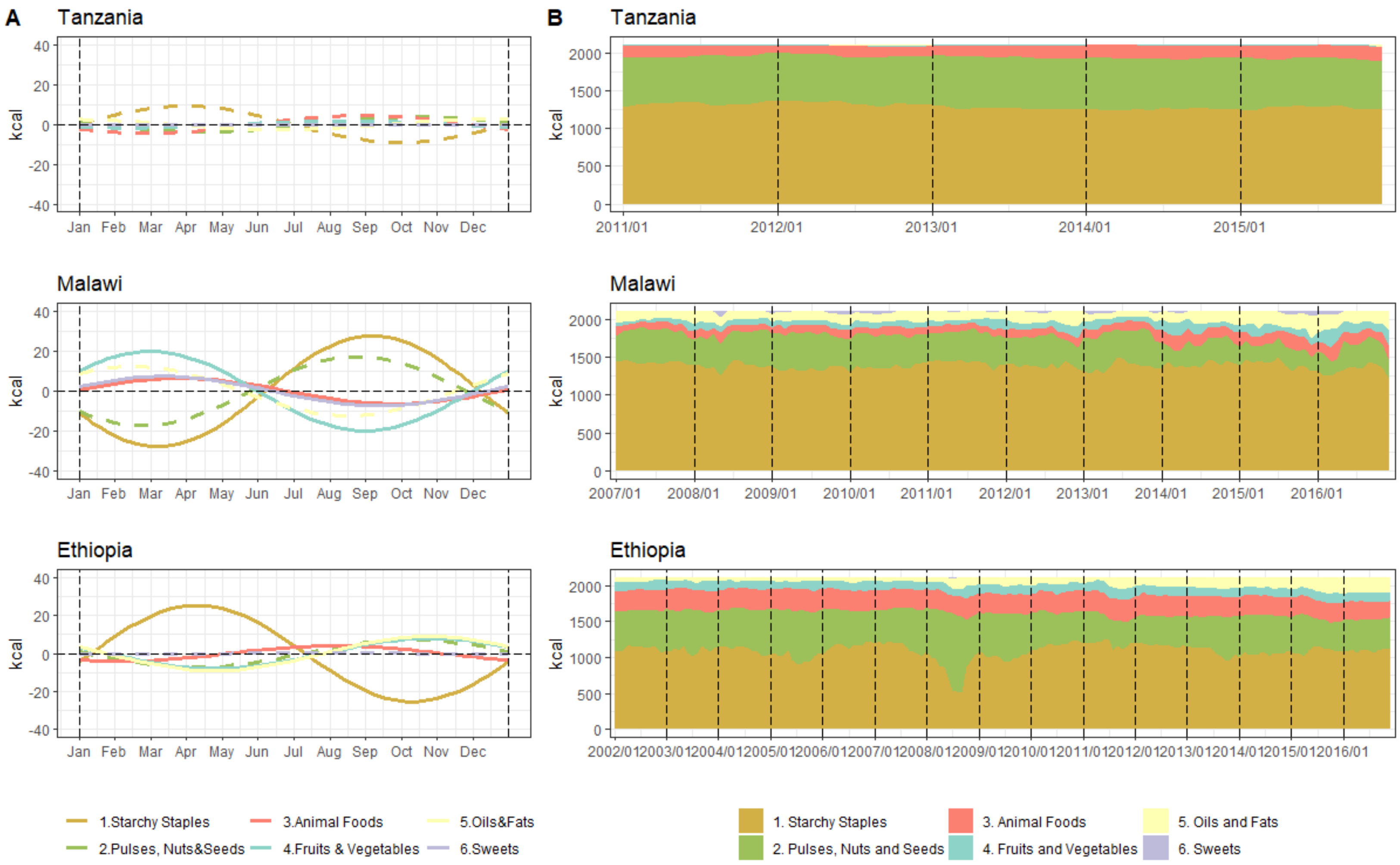


Fig. 4. Seasonality in the composition of least-cost diets by food group over time (kcal/day)
(**A**) displays the predicted seasonal curves over a year cycle of energy intakes in kcal from six food groups, and the dashed line means insignificant result. (**B**) shows the average energy compositions of CoNA over markets contributed by six food groups. The total daily energy intake is 2107.6kcal, required by a woman between 19-30 years old under low active physical activity level with a height of 163cm and a weight of 57kg.

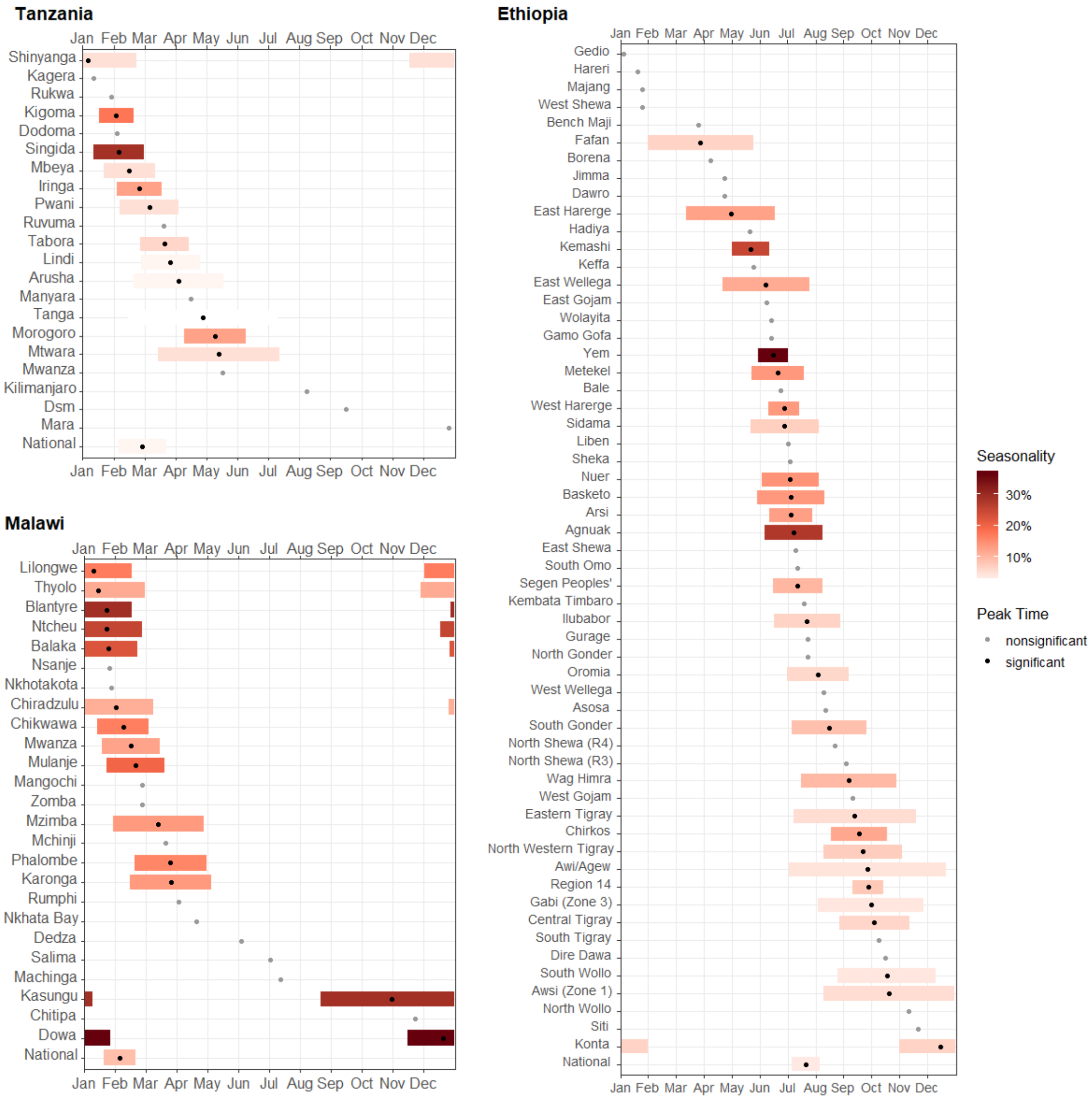


Fig. 5. Intensity and timing of seasonality in diet costs across 21 regions of Tanzania, 25 districts of Malawi and 57 zones of Ethiopia

Data shown are 95% confidence intervals around the peak month in each location, shown as a black dot, with the magnitude of intensity shown by the color gradation of each bar. Grey dots show the peak month in locations without statistically significant seasonality in diet costs, as measured by the cost of nutrient adequacy (CoNA).

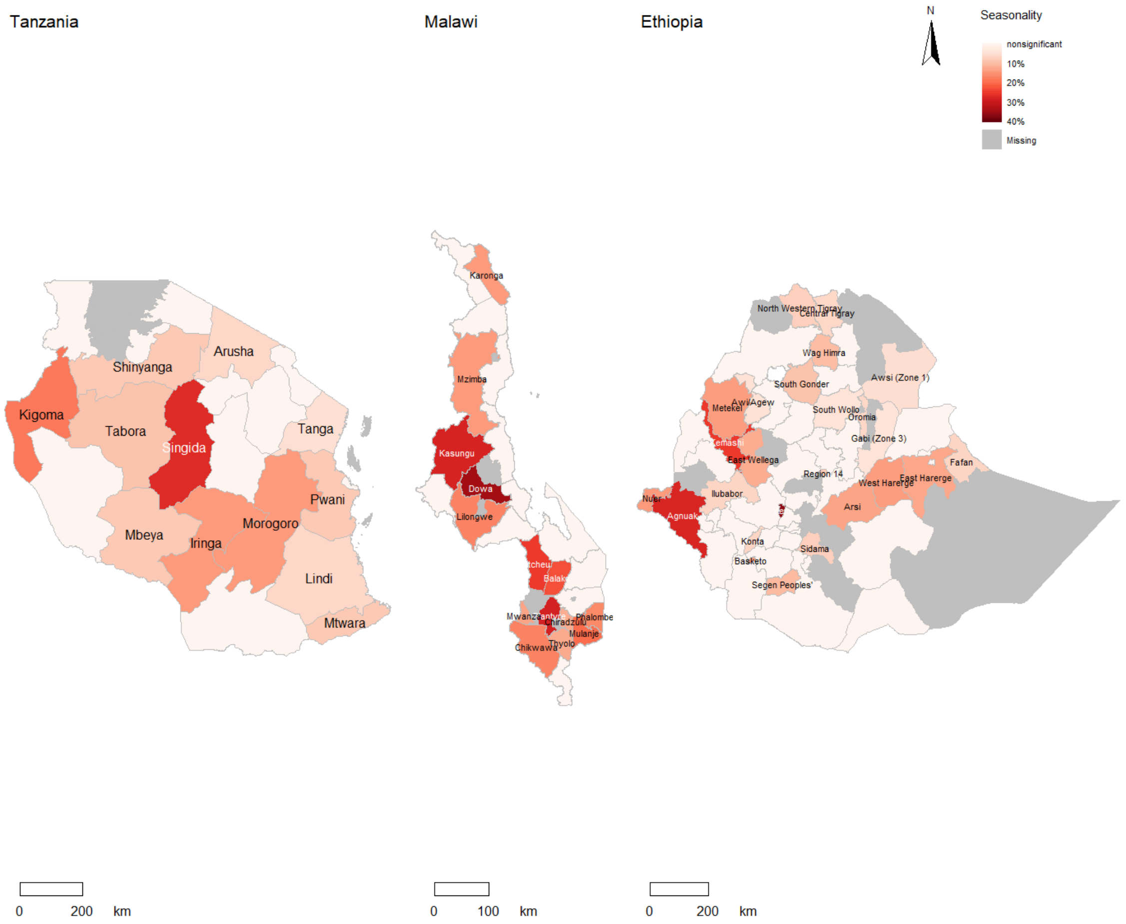


Fig. 6. Geographical distribution of seasonality in diet costs within Tanzania, Malawi and Ethiopia
Color gradations show the magnitude of estimated seasonal intensity in the cost of nutrient adequacy (CoNA).

SUPPLEMENTARY MATERIALS

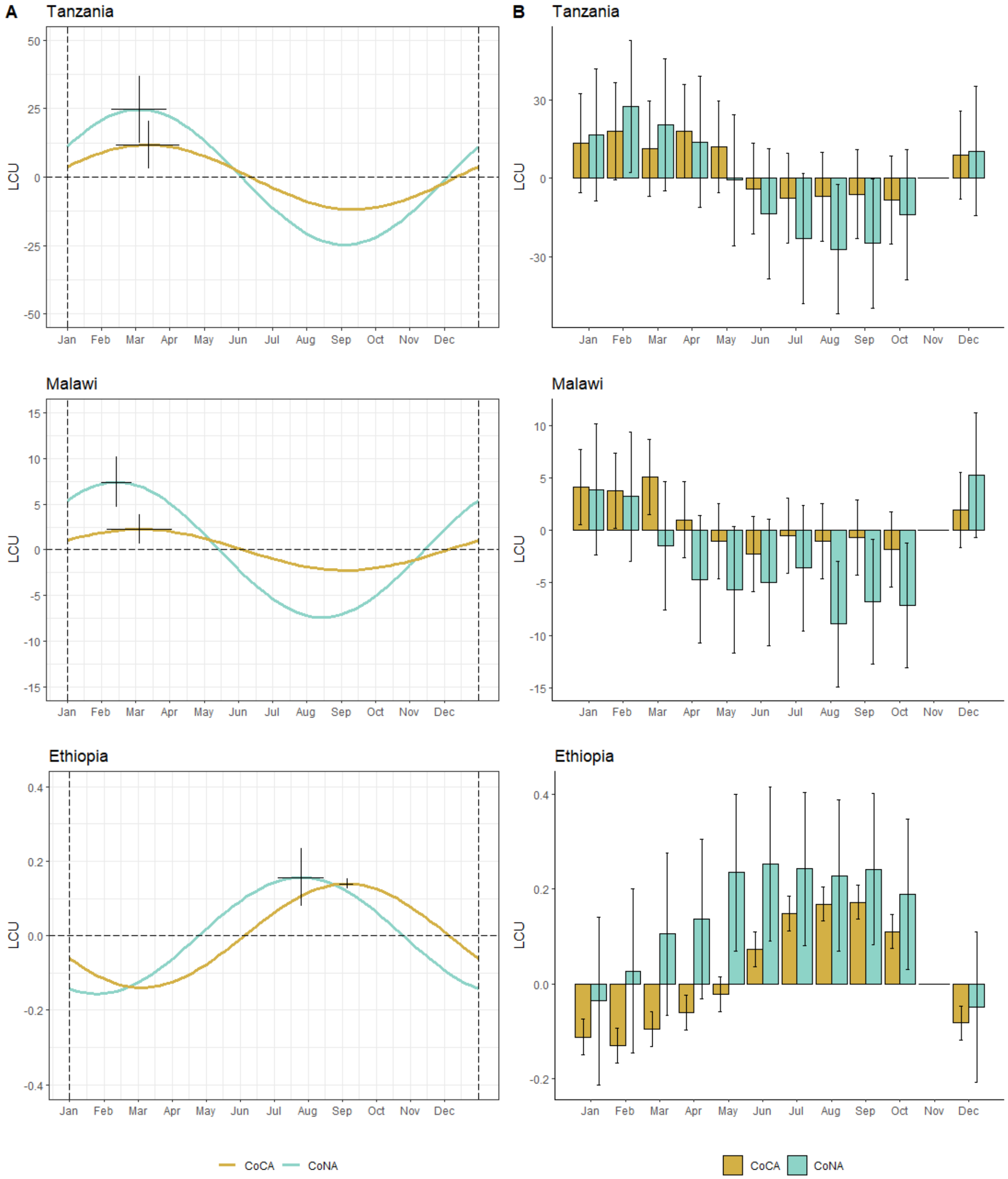


Fig. S1. Seasonality in diet costs for nutrient adequacy (CoNA) and caloric adequacy (CoCA): comparison between harmonic models and monthly indicators (LCU/day)

(**A**) displays estimated harmonic seasonality over a one-year cycle, with error bars showing 95% confidence intervals around the magnitude of seasonal intensity along the vertical axis and peak month along the horizontal axis. (**B**) shows seasonal variation as differences by calendar month, using November as the base period for end of the dry season in Tanzania and Malawi, and end of rainy season in Ethiopia, with error bars showing 95% confidence intervals around monthly differences in LCU/day.

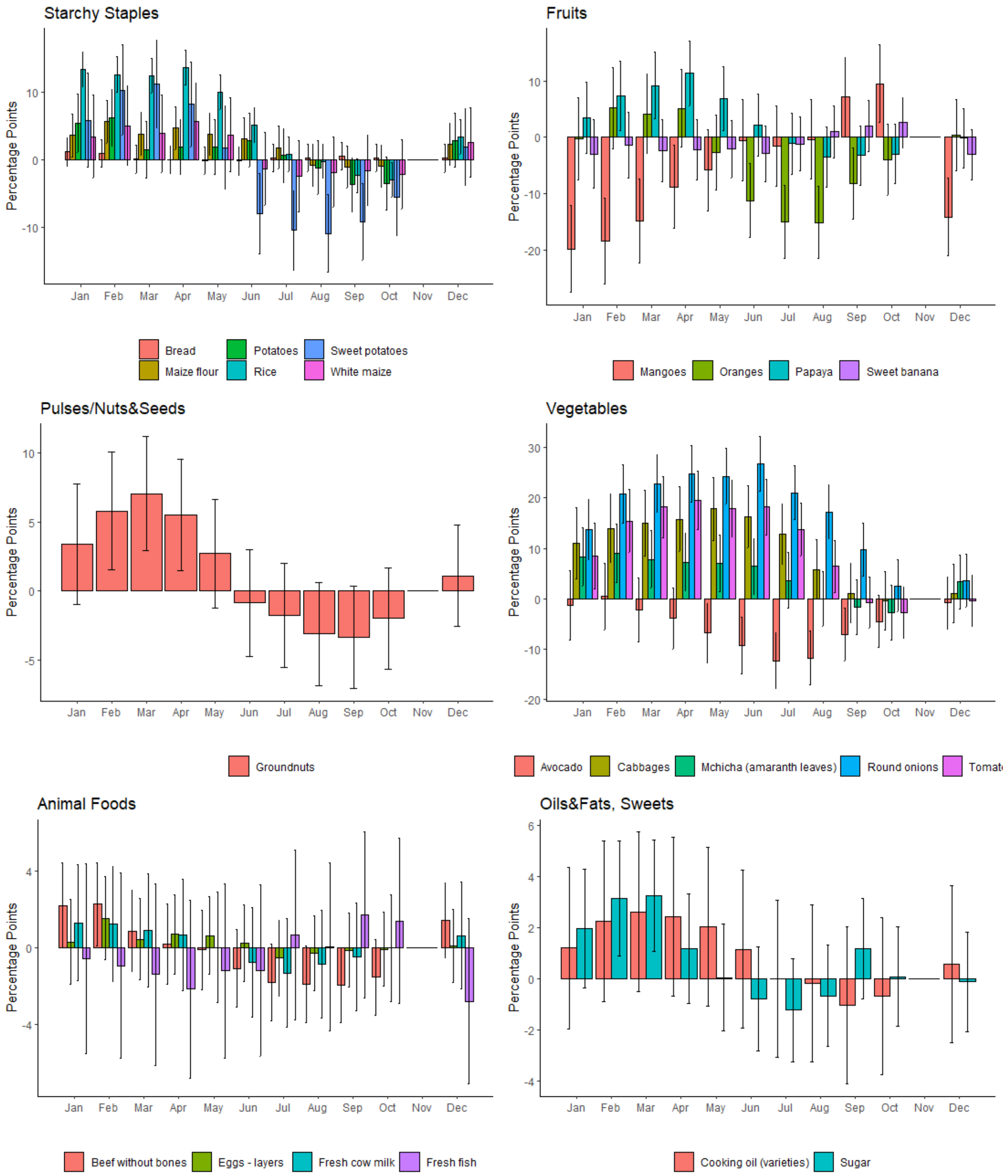
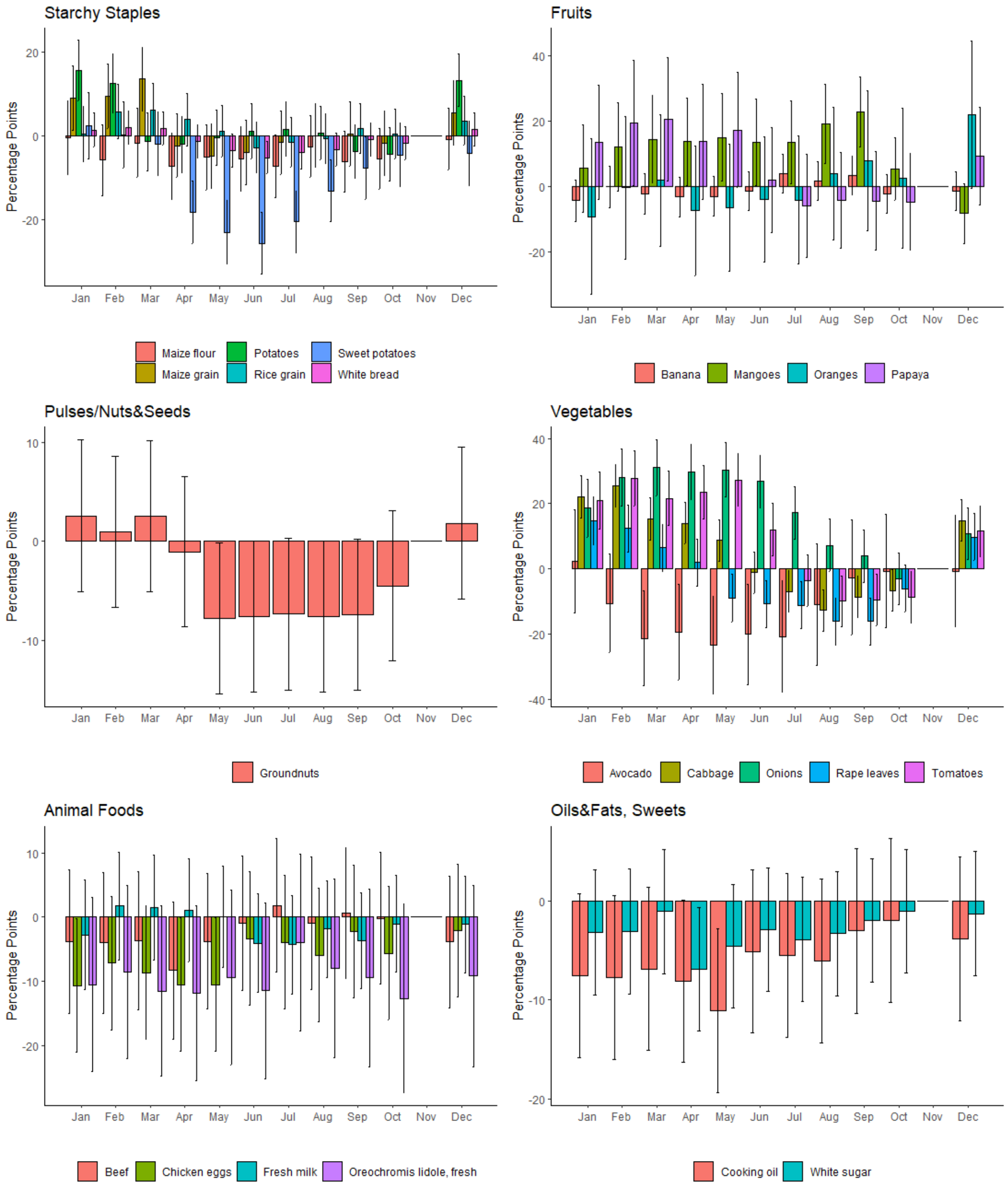


Fig. S2. Seasonality in retail prices of common food items in Tanzania

Data shown are percent differences by calendar month for selected items in each food group, relative to its price in November. Error bars show 95% confidence intervals.


Starchy Staples

Fruits

Pulses/Nuts&Seeds

Vegetables

Animal Foods

Oils&Fats, Sweets

Fig. S3. Seasonality in retail prices of common food items in Malawi

Data shown are percent differences by calendar month for selected items in each food group, relative to its price in November. Error bars show 95% confidence intervals.

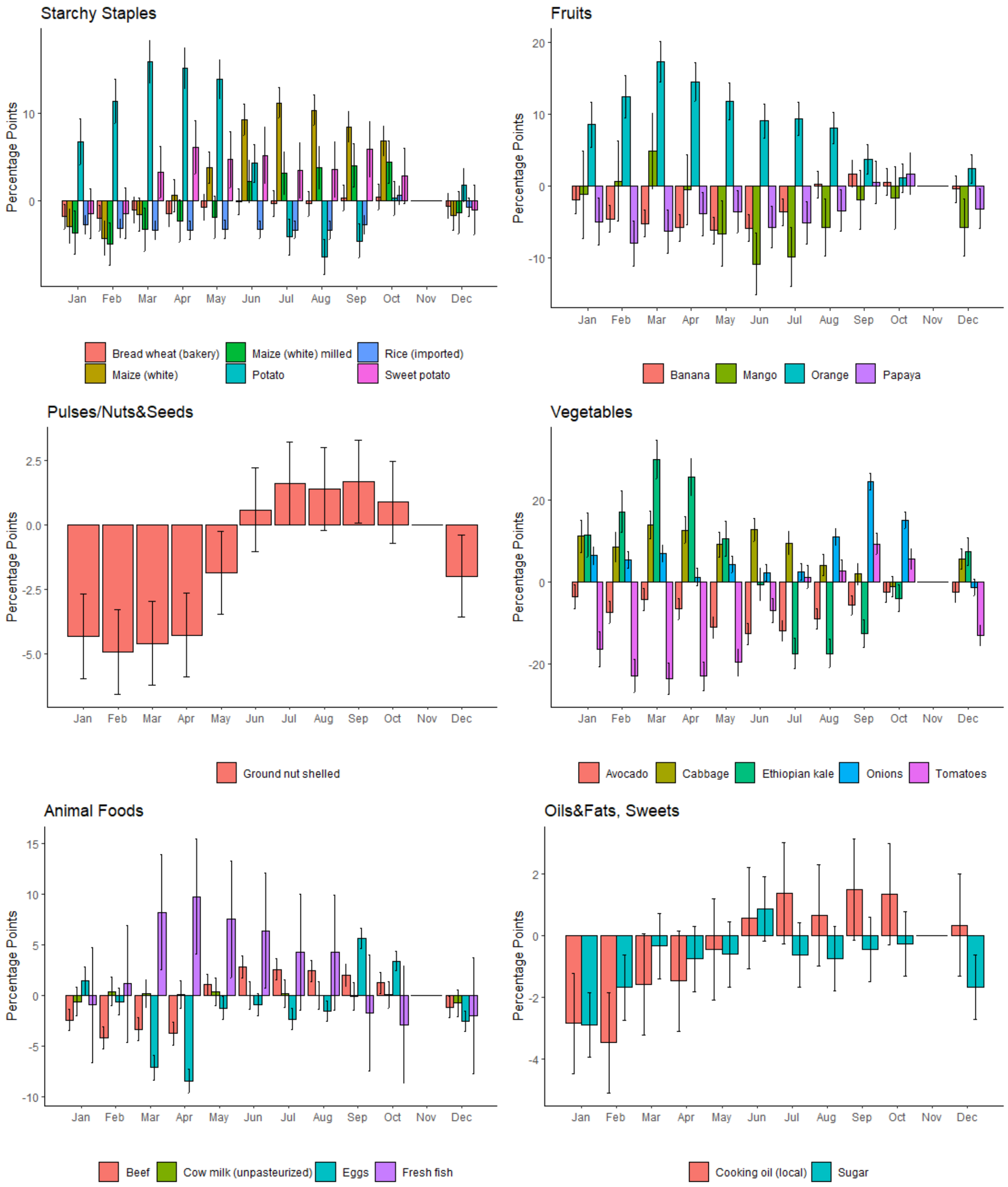


Fig. S4. Seasonality in retail prices of common food items in Ethiopia
Data shown are percent differences by calendar month for selected items in each food group, relative to its price in November. Error bars show 95% confidence intervals.

Table S1. Descriptive statistics and seasonality for monthly prices of all foods in Tanzania between 2011-2015

Food Groups	No	Food Items	Food Prices (LCU per 100kcal)				Peak Time			Seasonal Intensity		
			Obs.	Mean	SD	CV	Estimate	95% CI		Estimate	95% CI	
Starchy Staples	1	Biscuit	1,236	105.4	100.5	1.0	8.6	1.4	15.8	1.3%	(5.0%)	8.1%
	2	Bread	1,236	82.3	23.0	0.3	0.5	(3.3)	4.3	0.7%	(1.1%)	2.5%
	3	Bun	1,236	64.0	23.0	0.4	4.4	0.4	8.4	2.7%	(1.2%)	6.7%
	4	Cassava flour	1,063	44.3	20.3	0.5	10.4	6.9	13.9	2.5%	(0.5%)	5.6%
	5	Cassava fresh	1,236	45.6	16.2	0.4	1.1	0.3	1.8	10.1%	4.5%	16.0%
	6	Cooking bananas	1,236	97.3	43.2	0.4	11.6	5.6	17.6	2.2%	(2.3%)	6.9%
	7	Finger millet	1,236	46.8	13.3	0.3	4.8	(0.0)	9.7	1.6%	(1.0%)	4.3%
	8	Maize flour	1,236	27.8	5.2	0.2	3.0	2.1	3.9	6.0%	3.5%	8.5%
	9	Pastry	1,236	73.2	19.9	0.3	4.9	1.7	8.1	2.6%	(0.3%)	5.5%
	10	Potatoes	1,236	177.4	49.2	0.3	2.2	1.4	3.0	9.6%	5.3%	14.0%
	11	Rice	1,236	44.4	8.4	0.2	3.0	2.8	3.2	17.6%	15.0%	20.3%
	12	Spaghetti	1,236	37.5	3.8	0.1	11.8	9.4	14.1	0.9%	(0.2%)	1.9%
	13	Sweet potatoes	1,232	84.6	29.2	0.4	2.4	1.9	2.9	21.2%	13.4%	29.4%
	14	Wheat flour	1,236	36.4	3.7	0.1	2.5	1.6	3.3	2.7%	1.0%	4.4%
	15	White maize	1,236	18.3	6.1	0.3	2.5	1.4	3.6	6.5%	0.7%	12.6%
Pulses, Nuts and Seeds	16	Beans (soya)	1,236	41.6	7.4	0.2	2.5	2.1	3.0	9.5%	6.3%	12.8%
	17	Lentils	1,235	72.7	21.1	0.3	3.4	2.5	4.3	8.0%	4.3%	11.9%
	18	Natural groundnuts	1,236	38.8	7.0	0.2	2.8	2.0	3.7	8.8%	4.2%	13.5%
	19	Red dry beans	1,236	47.0	8.7	0.2	3.0	2.1	3.8	6.7%	3.9%	9.6%
Animal Foods	20	Beef Sausage	1,236	280.7	46.5	0.2	1.9	(1.1)	4.9	0.8%	(1.1%)	2.7%
	21	Beef with bones	1,236	555.1	94.0	0.2	1.3	0.8	1.8	6.3%	4.4%	8.3%
	22	Beef without bones	1,236	299.2	50.8	0.2	1.7	1.2	2.2	5.2%	3.1%	7.3%
	23	Dried fish	1,137	666.4	325.3	0.5	0.5	(0.9)	1.9	8.0%	(3.6%)	21.1%
	24	Dried sardines	1,236	255.4	138.9	0.5	3.3	2.0	4.5	6.3%	3.3%	9.3%
	25	Eggs - layers	1,188	478.2	66.0	0.1	2.7	0.6	4.8	1.3%	(0.6%)	3.3%
	26	Eggs - traditional	1,236	649.4	164.8	0.3	5.6	(3.0)	14.1	0.6%	(1.7%)	2.9%
	27	Fresh cow milk	1,236	174.8	39.2	0.2	1.6	0.1	3.1	2.3%	(0.5%)	5.1%
	28	Fresh fish	1,236	1226.5	379.0	0.3	9.1	6.9	11.3	3.0%	(1.0%)	7.3%
	29	Fried fish	1,193	874.6	341.2	0.4	1.1	(1.9)	4.0	2.7%	(3.0%)	8.8%

Food Groups	No	Food Items	Food Prices (LCU per 100kcal)				Peak Time			Seasonal Intensity		
			Obs.	Mean	SD	CV	Estimate	95% CI		Estimate	95% CI	
	30	Goat meat	1,231	229.9	47.3	0.2	1.1	0.0	2.2	2.4%	0.6%	4.3%
	31	Industrially bred chicken	981	523.7	108.1	0.2	0.8	(5.9)	7.4	0.5%	(1.9%)	2.9%
	32	Pork meat	1,234	132.9	30.5	0.2	1.0	(0.3)	2.3	2.0%	0.1%	4.1%
	33	Powdered milk	1,236	474.6	68.3	0.1	1.2	(3.0)	5.4	0.6%	(1.2%)	2.5%
	34	Traditionally bred chicken	1,236	755.3	194.1	0.3	0.2	(1.1)	1.4	3.4%	1.0%	5.8%
Fruits and Vegetables	35	Apples (imported)	1,216	1187.3	308.0	0.3	3.7	1.4	6.0	5.4%	(0.9%)	12.2%
	36	Avocado	1,236	112.8	51.4	0.5	1.1	(0.0)	2.3	12.3%	4.9%	20.2%
	37	Bitter tomatoes	1,236	598.4	270.5	0.5	2.4	0.4	4.3	4.0%	(1.6%)	10.1%
	38	Cabbages	1,236	302.7	166.1	0.6	4.0	3.4	4.7	20.9%	14.1%	28.1%
	39	Carrots	1,236	418.9	222.1	0.5	4.9	3.9	6.0	11.9%	6.8%	17.3%
	40	Coconut mature	1,236	757.7	290.1	0.4	11.5	10.7	12.4	11.3%	6.5%	16.2%
	41	Egg plant	1,236	583.5	249.9	0.4	5.9	(1.5)	13.2	1.7%	(4.1%)	7.9%
	42	Green bell pepper	1,236	1575.8	606.7	0.4	4.4	2.8	5.9	8.8%	3.0%	15.0%
	43	Green peas	1,236	808.6	283.0	0.4	3.6	2.0	5.1	9.1%	3.6%	14.9%
	44	Green pepper (hot)	1,236	1029.4	419.8	0.4	9.3	7.4	11.3	7.0%	1.7%	12.6%
	45	Ladies finger (okra)	1,236	713.9	255.6	0.4	11.8	7.8	15.9	2.7%	(3.2%)	9.1%
	46	Lemons	1,236	695.8	333.7	0.5	1.7	1.2	2.2	35.9%	24.6%	48.3%
	47	Limes	1,024	1120.6	709.7	0.6	0.7	(0.2)	1.6	21.1%	9.6%	33.7%
	48	Mangoes	1,226	247.1	106.9	0.4	8.2	7.8	8.7	29.7%	19.5%	40.8%
	49	Mchicha (amaranth leaves)	1,236	342.8	166.0	0.5	3.1	2.2	3.9	12.8%	7.5%	18.4%
	50	Oranges	1,230	257.7	89.1	0.4	2.0	1.2	2.8	15.1%	6.7%	24.2%
	51	Papaya	1,236	370.9	133.9	0.4	3.6	2.6	4.7	13.4%	7.1%	20.1%
	52	Pineapples	1,236	411.3	137.1	0.3	8.6	8.0	9.1	20.9%	13.5%	28.7%
	53	Round onions	1,236	382.1	115.0	0.3	4.9	4.4	5.4	26.7%	20.2%	33.6%
	54	Sweet banana	1,236	199.6	80.0	0.4	8.8	7.4	10.3	6.3%	0.4%	12.6%
	55	Tomatoes red	1,236	533.5	201.2	0.4	4.4	3.8	4.9	25.8%	18.7%	33.3%
Oils and Fats	56	Cooking fat	1,236	145.1	82.8	0.6	2.3	0.4	4.2	2.4%	(0.8%)	5.8%
	57	Cooking oil	1,236	75.4	51.0	0.7	3.2	1.8	4.7	3.4%	1.0%	5.9%
	58	Margarine	1,236	99.8	25.2	0.3	4.3	0.8	7.8	1.4%	(0.3%)	3.2%
Sweets	59	Honey	1,236	151.9	56.1	0.4	1.3	(0.9)	3.5	1.9%	(1.4%)	5.4%

Food Groups	No	Food Items	Food Prices (LCU per 100kcal)				Peak Time			Seasonal Intensity		
			Obs.	Mean	SD	CV	Estimate	95% CI		Estimate	95% CI	
	60	Mixed fruits jam	1,236	278.3	52.4	0.2	1.8	(1.0)	4.6	1.0%	(1.3%)	3.3%
	61	Sugar	1,236	51.5	5.1	0.1	2.0	0.9	3.1	2.3%	0.4%	4.2%

Note: Peak time shows as a continuous variable from 0-12, where 0 indicates the start and 12 as the end of a calendar year.

Table S2. Descriptive statistics and seasonality for monthly prices of all foods in Malawi between 2007 to 2016

Food Groups	No	Food Items	Food Prices (LCU per 100kcal)				Peak Time			Seasonal Intensity		
			Obs.	Mean	SD	CV	Estimate	95% CI		Estimate	95% CI	
Starchy Staples	1	Biscuits	1,301	12.7	6.4	0.5	2.9	(0.4)	6.2	4.1%	(3.1%)	11.8%
	2	Cassava	2,522	7.1	6.3	0.9	8.0	(6.2)	22.2	0.9%	(6.6%)	9.0%
	3	Maize flour	2,791	7.6	5.5	0.7	0.9	(0.8)	2.5	6.4%	(1.1%)	14.5%
	4	Maize four (whole grain)	1,178	9.0	4.6	0.5	11.9	10.5	13.2	10.8%	2.4%	20.0%
	5	Maize grain	3,419	3.0	2.6	0.9	1.6	0.6	2.6	10.7%	2.9%	19.1%
	6	Maize grain, Admarc	3,048	2.2	1.7	0.8	3.4	1.4	5.3	6.4%	1.2%	11.8%
	7	Mandazi (fried-dough)	1,297	17.8	16.9	1.0	11.2	9.6	12.8	13.3%	4.0%	23.4%
	8	Potatoes	3,380	26.1	21.8	0.8	1.4	0.9	1.9	22.6%	14.5%	31.2%
	9	Rice grain	3,039	8.6	5.8	0.7	3.0	(1.1)	7.0	2.4%	(3.8%)	9.0%
	10	Sweet potatoes	2,897	12.7	10.9	0.9	11.8	11.2	12.4	25.2%	16.2%	35.0%
	11	White bread	3,340	10.8	5.0	0.5	1.1	0.2	2.0	5.3%	1.5%	9.2%
Pulses, Nuts and Seeds	12	Brown beans	3,380	12.8	8.4	0.7	0.6	(0.2)	1.4	13.0%	6.7%	19.7%
	13	Cowpeas	1,025	13.9	7.1	0.5	1.9	1.1	2.7	20.6%	8.9%	33.5%
	14	Groundnuts	3,030	8.4	6.1	0.7	1.2	(0.1)	2.4	9.0%	1.1%	17.5%
	15	Pigeon peas	730	15.1	7.9	0.5	1.9	(0.3)	4.2	6.5%	(4.5%)	18.6%
	16	White beans	3,033	11.6	7.5	0.6	1.1	(0.1)	2.2	8.8%	2.0%	16.0%
Animal Foods	17	Beef	3,248	34.0	22.5	0.7	9.4	5.4	13.4	4.2%	(2.9%)	11.9%
	18	Chicken	1,838	67.9	50.2	0.7	2.6	0.9	4.3	8.8%	(1.5%)	20.2%
	19	Chicken eggs	3,187	65.7	41.8	0.6	9.9	7.1	12.7	6.8%	(0.1%)	14.3%
	20	Cichlid, dried	2,736	280.1	199.5	0.7	9.4	7.5	11.3	8.3%	1.2%	15.8%
	21	Fresh milk	2,828	54.1	70.4	1.3	2.6	(0.5)	5.7	3.3%	(4.0%)	11.1%
	22	Goat	3,276	74.3	44.4	0.6	9.3	7.9	10.7	9.1%	3.0%	15.5%
	23	Oreochromis lidole, dry	1,366	316.1	251.8	0.8	10.1	6.4	13.8	5.1%	(1.3%)	11.9%
	24	Oreochromis lidole, fresh	499	141.4	139.1	1.0	9.7	7.1	12.3	8.4%	(0.4%)	18.0%
	25	Pork	1,538	35.0	16.7	0.5	8.7	5.3	12.0	5.5%	(5.1%)	17.3%

	26	Powdered milk	2,034	27.2	20.9	0.8	11.0	8.3	13.8	11.3%	0.9%	22.6%
	27	Sardine	3,406	91.3	65.6	0.7	6.8	6.0	7.5	15.4%	6.0%	25.7%
Fruits and Vegetables	28	Avocado	663	20.5	12.4	0.6	10.7	9.7	11.6	32.8%	17.7%	49.9%
	29	Banana	3,115	24.5	19.6	0.8	8.1	6.9	9.3	6.3%	(0.0%)	13.1%
	30	Cabbage	3,221	37.5	26.9	0.7	2.1	1.9	2.3	46.9%	37.5%	57.0%
	31	Chinese cabbage	1,263	151.4	102.5	0.7	1.5	0.9	2.2	33.0%	17.7%	50.2%
	32	Cucumber	644	177.4	98.0	0.6	9.5	8.7	10.2	52.7%	23.9%	88.1%
	33	Eggplant	1,018	107.0	56.5	0.5	1.6	0.7	2.4	19.2%	5.9%	34.2%
	34	Green beans	894	155.9	98.9	0.6	2.9	1.1	4.8	15.2%	(4.9%)	39.5%
	35	Guava	415	30.2	16.6	0.6	9.3	7.7	10.9	25.1%	7.2%	46.1%
	36	Mangoes	1,258	19.4	16.0	0.8	6.4	5.5	7.3	38.2%	17.3%	62.9%
	37	Okra	2,823	78.1	77.9	1.0	9.4	8.8	10.0	28.7%	16.9%	41.6%
	38	Onions	3,424	97.8	79.6	0.8	3.8	3.3	4.2	41.8%	32.1%	52.1%
	39	Oranges	546	66.1	50.5	0.8	11.3	9.6	13.0	24.0%	5.7%	45.6%
	40	Papaya	468	67.3	50.6	0.8	2.6	1.7	3.5	32.8%	12.2%	57.2%
	41	Pumpkin	406	67.9	63.5	0.9	11.8	10.4	13.2	52.0%	17.8%	96.1%
	42	Pumpkin leaves	2,986	196.9	162.3	0.8	9.2	8.6	9.9	20.7%	10.7%	31.6%
	43	Rape leaves	3,258	46.0	37.0	0.8	1.5	1.1	2.0	32.7%	22.2%	44.2%
	44	Tomatoes	3,432	119.5	99.2	0.8	2.7	2.3	3.1	60.3%	46.1%	75.9%
Oils and Fats	45	Cooking oil	3,030	12.4	9.4	0.8	9.8	7.5	12.1	6.2%	(0.7%)	13.5%
	46	Cooking oil, refilled	3,202	7.7	4.7	0.6	9.7	7.9	11.6	8.3%	1.6%	15.4%
Sweets	47	Brown sugar	3,392	8.0	5.7	0.7	8.6	6.4	10.9	4.5%	(1.2%)	10.6%
	48	White sugar	2,903	6.9	4.7	0.7	11.6	8.5	14.7	4.1%	(0.3%)	8.8%

Note: Peak time shows as a continuous variable from 0-12, where 0 indicates the start and 12 as the end of a calendar year.

Table S3. Descriptive statistics and seasonality for monthly prices of all foods in Ethiopia between 2002 to 2016

Food Groups	No	Food Items	Food Prices (LCU per 100kcal)				Peak Time			Seasonal Intensity (%)		
			Obs.	Mean	SD	CV	Estimate	95% CI		Estimate	95% CI	
Starchy Staples	1	African millet	11,269	0.2	0.1	0.7	8.4	7.9	8.9	11.1%	9.0%	13.4%
	2	Barley black	12,423	0.2	0.1	0.6	8.2	7.8	8.7	8.8%	7.0%	10.6%
	3	Barley white	17,169	0.2	0.1	0.6	8.3	8.0	8.7	10.0%	8.6%	11.5%
	4	Barley white milled	1,033	0.4	0.3	0.6	11.8	8.2	15.3	1.8%	(2.8%)	6.7%
	5	Bread wheat (bakery)	19,749	0.4	0.2	0.5	7.8	6.6	9.0	1.6%	0.5%	2.7%
	6	Durrah	8,071	0.1	0.1	0.7	8.3	7.8	8.7	17.8%	14.6%	21.1%
	7	Hulled barley	3,747	0.2	0.2	0.7	9.1	8.2	10.0	6.7%	4.0%	9.5%
	8	Kocho (unprocessed)	8,271	0.3	0.3	1.3	11.2	6.9	15.4	1.3%	(1.8%)	4.4%
	9	Maize (white)	19,180	0.1	0.1	0.6	7.3	7.1	7.5	16.2%	13.9%	18.5%
	10	Maize (white) milled	6,849	0.2	0.1	0.6	7.8	7.3	8.2	9.1%	7.0%	11.3%
	11	Oats	4,886	0.2	0.2	0.8	8.4	7.5	9.3	7.0%	4.5%	9.6%
	12	Oats milled	9,544	0.4	0.3	0.7	9.5	7.4	11.5	1.2%	0.4%	2.0%
	13	Potatoes	20,162	0.6	0.4	0.7	3.5	3.3	3.8	23.0%	19.8%	26.3%
	14	Rice (imported)	18,032	0.3	0.2	0.5	7.6	7.2	8.0	1.9%	1.4%	2.3%
	15	Sorghum milled	959	0.2	0.1	0.6	9.7	5.4	14.0	9.2%	0.1%	19.1%
	16	Sorghum red	13,867	0.1	0.1	0.6	8.7	8.5	9.0	16.6%	13.7%	19.6%
	17	Sorghum white	13,531	0.1	0.1	0.7	8.6	8.3	8.9	13.9%	11.8%	16.1%
	18	Spaghetti (local)	17,840	0.5	0.2	0.5	7.4	6.9	7.9	2.3%	1.5%	3.1%
	19	Sweet potatoes	10,985	0.3	0.2	0.9	6.3	5.8	6.9	8.3%	4.5%	12.3%
	20	Teff black (red) milled	1,697	0.2	0.2	0.7	8.6	7.7	9.5	6.1%	3.2%	9.0%
	21	Teff mixed milled	2,685	0.3	0.2	0.6	7.9	7.2	8.6	5.2%	3.1%	7.3%
	22	Teff white milled	2,590	0.3	0.2	0.6	8.3	7.6	9.0	4.4%	2.5%	6.4%
	23	Wheat black (red)	12,893	0.2	0.1	0.6	8.4	8.1	8.8	9.9%	8.2%	11.6%
	24	Wheat mixed milled	824	0.2	0.1	0.7	8.9	7.6	10.2	10.0%	5.7%	14.6%

Food Groups	No	Food Items	Food Prices (LCU per 100kcal)				Peak Time			Seasonal Intensity (%)		
			Obs.	Mean	SD	CV	Estimate	95% CI		Estimate	95% CI	
	25	Wheat white	17,905	0.2	0.1	0.6	8.1	8.0	8.3	12.7%	11.3%	14.0%
	26	Wheat white milled	1,838	0.2	0.1	0.7	7.7	6.8	8.6	6.0%	2.7%	9.4%
Pulses, Nuts and Seeds	27	Chickpeas	19,275	0.2	0.2	0.7	10.0	9.7	10.4	7.1%	6.1%	8.1%
	28	Fenugreek seed	18,027	0.4	0.3	0.7	10.2	9.6	10.8	4.7%	3.7%	5.7%
	29	Ground nut shelled	14,075	0.3	0.2	0.7	8.2	7.8	8.5	7.7%	6.3%	9.0%
	30	Haricot beans	10,816	0.2	0.1	0.7	5.5	5.1	5.9	13.3%	10.3%	16.3%
	31	Horse beans	19,620	0.2	0.2	0.7	8.0	7.9	8.2	15.9%	14.4%	17.5%
	32	Lentils	18,589	0.4	0.3	0.8	7.5	7.2	7.7	8.7%	7.5%	10.0%
	33	Lima beans	6,277	0.2	0.2	1.3	6.0	5.7	6.3	17.6%	13.2%	22.1%
	34	Linseed red	18,124	0.3	0.2	0.8	8.1	7.7	8.4	8.8%	7.2%	10.5%
	35	Linseed white	3,726	0.4	0.4	0.9	7.4	6.7	8.0	7.2%	4.0%	10.5%
	36	Niger seed	14,939	0.3	0.2	0.8	8.0	7.6	8.4	7.8%	6.1%	9.6%
	37	Peas green (dry)	10,640	3.1	1.9	0.6	7.6	7.4	7.8	15.1%	13.3%	17.0%
	38	Peas split	16,427	0.4	0.2	0.7	6.6	6.3	6.9	4.6%	3.7%	5.5%
	39	Sesame seed	9,377	0.3	0.2	0.8	7.0	6.5	7.5	7.5%	4.8%	10.3%
	40	Soya beans	5,191	0.2	0.2	0.7	7.6	7.3	8.0	13.6%	10.5%	16.7%
	41	Sunflower seed	13,463	0.4	0.3	0.7	10.1	8.0	12.2	1.9%	0.3%	3.6%
	42	Vetches	11,125	0.2	0.2	0.7	10.9	10.7	11.1	16.3%	14.6%	18.0%
Animal Foods	43	Beef	18,312	5.4	4.1	0.8	7.8	7.6	8.0	7.8%	6.9%	8.6%
	44	Camel meat	1,599	1.9	1.7	0.9	8.4	6.7	10.1	4.5%	1.1%	8.1%
	45	Camel milk	2,108	1.1	0.8	0.8	3.5	2.0	5.0	6.3%	0.1%	12.8%
	46	Cheese cottage	9,761	1.5	1.3	0.9	1.9	(0.2)	4.1	2.3%	(0.2%)	4.9%
	47	Cow milk (unpasteurized)	17,120	1.0	0.7	0.7	4.6	1.8	7.4	0.9%	(0.3%)	2.1%
	48	Eggs	20,139	2.3	1.7	0.7	10.2	9.9	10.5	8.1%	6.7%	9.6%
	49	Fresh fish	2,352	3.5	3.0	0.9	4.9	4.2	5.6	12.3%	7.6%	17.1%
	50	Goat milk	675	1.6	1.1	0.7	8.9	(0.5)	18.2	1.5%	(3.6%)	7.0%
	51	Sardines (imported)	9,246	6.2	3.9	0.6	4.5	(13.8)	22.9	0.1%	(0.9%)	1.2%
	52	Yoghurt (traditional)	13,753	1.9	1.4	0.7	7.0	4.3	9.6	0.7%	(0.8%)	2.3%

Food Groups	No	Food Items	Food Prices (LCU per 100kcal)				Peak Time			Seasonal Intensity (%)		
			Obs.	Mean	SD	CV	Estimate	95% CI		Estimate	95% CI	
Fruits and Vegetables	53	Avocado	11,877	1.0	0.7	0.7	11.9	11.4	12.4	9.0%	5.8%	12.3%
	54	Banana	19,208	1.2	1.0	0.8	10.7	10.2	11.1	8.6%	6.8%	10.5%
	55	Beet root	17,910	1.5	1.2	0.8	4.4	4.0	4.8	12.6%	10.1%	15.1%
	56	Cabbage	18,281	1.7	1.3	0.8	3.6	3.1	4.1	14.5%	10.4%	18.9%
	57	Cactus	927	1.1	1.1	1.0	1.4	0.9	2.0	55.5%	25.3%	92.9%
	58	Carrot	17,459	1.8	1.4	0.8	5.8	5.6	6.1	25.9%	21.8%	30.1%
	59	Cauliflower	1,143	17.9	14.4	0.8	4.2	0.3	8.2	4.9%	(3.9%)	14.4%
	60	Ethiopian kale	16,608	1.3	1.3	1.0	2.8	2.6	3.0	46.9%	38.0%	56.4%
	61	Garlics	19,322	2.5	2.2	0.9	6.5	6.2	6.7	22.4%	19.5%	25.4%
	62	Grapes	346	22.0	20.6	0.9	4.5	1.8	7.1	14.5%	(3.1%)	35.3%
	63	Green peas	4,071	4.3	3.0	0.7	0.7	(2.1)	3.4	3.0%	(1.3%)	7.5%
	64	Green pepper	19,113	4.3	3.4	0.8	3.5	3.2	3.8	27.3%	23.3%	31.5%
	65	Hot pepper (dried)	15,342	1.0	0.8	0.8	7.5	7.4	7.6	26.6%	24.3%	28.9%
	66	Leaks	5,677	2.0	1.7	0.9	10.5	9.4	11.7	8.0%	3.6%	12.5%
	67	Lemon	17,749	2.5	2.4	1.0	2.7	2.5	2.9	60.7%	52.1%	69.9%
	68	Lettuce	9,060	5.6	5.6	1.0	3.7	2.2	5.2	6.0%	0.9%	11.4%
	69	Mandarine	2,017	2.6	2.1	0.8	4.3	3.4	5.2	12.4%	6.2%	19.1%
	70	Mango	11,235	2.8	2.2	0.8	0.7	0.2	1.2	16.2%	11.1%	21.5%
	71	Onions	19,655	1.1	0.7	0.6	10.7	10.3	11.0	13.6%	11.6%	15.8%
	72	Orange	15,207	5.0	4.6	0.9	3.5	3.1	4.0	12.8%	9.5%	16.3%
	73	Papaya	12,767	2.0	2.1	1.0	8.8	8.0	9.6	6.2%	3.4%	9.1%
	74	Pumpkin	9,599	0.5	0.4	0.9	4.0	3.4	4.5	17.1%	12.6%	21.7%
	75	Spinach	9,859	1.3	1.2	0.9	2.1	1.5	2.8	15.7%	8.7%	23.1%
	76	Tomatoes	18,562	3.1	2.4	0.8	9.0	8.8	9.3	38.7%	31.7%	46.2%
Oils and Fats	77	Butter unrefined	18,461	1.0	0.7	0.7	4.8	4.5	5.1	12.1%	10.1%	14.0%
	78	Cooking oil (imported)	17,971	0.2	0.1	0.4	3.7	2.6	4.9	1.3%	0.8%	1.9%
	79	Cooking oil (local)	12,305	0.3	0.2	0.7	7.9	7.4	8.5	3.5%	2.4%	4.6%
	80	Vegetable butter (imported)	10,047	2.1	0.8	0.4	7.3	6.7	8.0	1.8%	1.0%	2.7%

Food Groups	No	Food Items	Food Prices (LCU per 100kcal)				Peak Time			Seasonal Intensity (%)		
			Obs.	Mean	SD	CV	Estimate	95% CI		Estimate	95% CI	
Sweets	81	Honey	14,745	1.3	1.0	0.8	3.3	2.0	4.7	2.1%	1.0%	3.1%
	82	Sugar	18,651	0.3	0.1	0.5	6.4	6.1	6.7	3.0%	2.2%	3.8%

Note: Peak time shows as a continuous variable from 0-12, where 0 indicates the start and 12 as the end of a calendar year.

Table S4. Descriptive statistics and seasonality for national CoNA, CoCA, CoNA premium, energy composition and cost components of CoNA by food groups, and regional CoNA in Tanzania between 2011-2015

Categories	No	Items	Obs.	Mean	SD	CV	Peak Time			Seasonal Intensity		
							Estimate	95% CI		Estimate	95% CI	
Indicators (LCU)	1	CoNA	1,236	912.1	172.3	0.2	1.9	1.3	2.5	6.3%	3.7%	9.0%
	2	CoCA	1,236	378.4	122.0	0.3	2.4	1.2	3.5	5.9%	0.8%	11.3%
	3	CoNA premium	1,236	533.7	133.5	0.3	1.7	0.9	2.5	6.3%	2.4%	10.4%
CoNA Energy Composition (kcal)	1	Starchy Staples	1,236	1,317.8	78.6	0.1	10.3	4.6	15.9	6.1	(7.5)	19.7
	2	Pulses, Nuts and Seeds	1,236	618.9	146.2	0.2	4.0	(2.0)	10.0	8.2	(11.5)	28.0
	3	Animal Foods	1,236	153.8	115.6	0.8	6.0	(0.8)	12.8	4.6	(12.1)	21.4
	4	Fruits and Vegetables	1,236	13.0	15.6	1.2	8.0	5.7	10.3	2.6	(1.7)	7.0
	5	Oils and Fats	1,236	4.1	31.0	7.5	12.0	10.2	13.8	6.2	(2.1)	14.6
	6	Sweets	1,236	0.0	0.0	.	.	.	.	.	.	.
CoNA and Cost Components (2011USD)	1	Starchy Staples	1,236	0.5	0.1	0.3	2.3	1.3	3.3	0.03	0.01	0.05
	2	Pulses, Nuts and Seeds	1,236	0.4	0.1	0.3	2.8	1.8	3.8	0.03	0.01	0.05
	3	Animal Foods	1,236	0.6	0.2	0.4	0.9	(1.1)	2.9	0.02	(0.01)	0.06
	4	Fruits and Vegetables	1,236	0.1	0.0	0.8	3.6	0.9	6.3	0.01	(0.00)	0.02
	5	Oils and Fats	1,236	0.0	0.0	7.6	0.1	(1.8)	1.9	0.00	(0.00)	0.01
	6	Sweets	1,236	0.0	0.0	.	.	.	.	.	.	.
	7	CoNA	1,236	1.5	0.2	0.2	2.2	1.5	2.9	0.08	0.05	0.12
Regional CoNA (LCU)	1	Dodoma	60	947.6	142.7	0.2	1.4	0.2	2.6	12.9%	(0.2%)	27.7%
	2	Kilimanjaro	60	1048.9	171.9	0.2	6.9	5.3	8.4	4.6%	(0.7%)	10.2%
	3	Arusha	60	1062.2	182.7	0.2	2.9	2.1	3.8	9.4%	5.0%	13.9%
	4	Tanga	60	865.4	111.1	0.1	3.2	1.7	4.7	6.1%	1.5%	10.9%
	5	Morogoro	60	1051.9	117.1	0.1	4.3	3.4	5.2	9.1%	6.2%	12.1%
	6	Pwani	60	945.8	73.4	0.1	2.3	1.3	3.3	7.1%	2.0%	12.3%
	7	Dsm	60	1145.2	164.0	0.1	9.7	1.6	17.9	0.9%	(1.9%)	3.8%
	8	Lindi	60	1036.2	155.5	0.2	2.6	1.9	3.3	8.0%	4.5%	11.6%
	9	Mtwara	60	921.0	126.4	0.1	4.1	2.2	5.9	8.0%	2.7%	13.6%
	10	Ruvuma	60	788.1	109.7	0.1	2.6	0.7	4.5	5.7%	(1.1%)	13.1%
	11	Iringa	60	880.8	137.0	0.2	1.9	1.2	2.5	13.5%	6.0%	21.4%

12	Mbeya	60	852.9	83.4	0.1	1.5	0.7	2.3	7.6%	2.3%	13.2%
13	Singida	60	932.5	194.9	0.2	1.1	0.2	1.9	24.7%	8.2%	43.7%
14	Tabora	60	891.3	118.2	0.1	2.6	1.9	3.2	9.5%	5.6%	13.5%
15	Rukwa	60	765.5	118.2	0.2	0.9	(0.4)	2.3	6.7%	(0.1%)	13.9%
16	Kigoma	60	654.5	68.3	0.1	1.1	0.5	1.6	18.2%	9.9%	27.2%
17	Shinyanga	60	900.0	180.2	0.2	0.5	(0.8)	1.7	8.9%	1.9%	16.4%
18	Kagera	60	796.0	109.2	0.1	0.3	(2.1)	2.6	3.6%	(1.3%)	8.8%
19	Mwanza	60	908.8	77.1	0.1	1.8	0.1	3.5	2.4%	(0.8%)	5.8%
20	Mara	48	830.0	103.7	0.1	12.0	8.6	15.3	3.8%	(2.5%)	10.6%
21	Manyara	48	912.7	78.6	0.1	3.3	1.0	5.6	4.2%	(0.4%)	9.0%

Note: Peak time shows as a continuous variable from 0-12, where 0 indicates the start and 12 as the end of a calendar year.

Table S5. Descriptive statistics and seasonality for national CoNA, CoCA, CoNA premium, energy composition and cost components of CoNA by food groups, and regional CoNA in Malawi between 2007-2016

Categories	No	Items	Obs.	Mean	SD	CV	Peak Time			Seasonal Intensity		
							Estimate	95% CI		Estimate	95% CI	
Indicators (LCU)	1	CoNA	3,480	129.6	73.6	0.6	1.1	0.6	1.7	10.0%	5.7%	14.6%
	2	CoCA	3,480	41.6	30.6	0.7	1.5	0.4	2.6	8.0%	1.5%	14.9%
	3	CoNA premium	3,480	88.0	54.6	0.6	0.8	(0.1)	1.7	9.0%	2.7%	15.6%
CoNA Energy Composition (kcal)	1	Starchy Staples	3,480	1,387.9	261.2	0.2	8.2	7.2	9.2	55.6	19.8	91.4
	2	Pulses, Nuts and Seeds	3,480	372.5	261.6	0.7	7.8	5.9	9.7	34.6	(9.7)	79.0
	3	Animal Foods	3,480	115.4	173.0	1.5	2.8	(0.1)	5.8	12.7	(7.6)	33.0
	4	Fruits and Vegetables	3,480	89.9	127.1	1.4	2.0	1.3	2.7	39.8	21.0	58.7
	5	Oils and Fats	3,480	132.5	150.3	1.1	1.5	0.1	3.0	24.7	(0.4)	49.8
	6	Sweets	3,480	9.4	78.3	8.3	2.4	1.2	3.6	14.7	3.1	26.4
CoNA and Cost Components (2011USD)	1	Starchy Staples	3,480	0.4	0.2	0.5	1.3	(0.5)	3.1	0.02	(0.01)	0.04
	2	Pulses, Nuts and Seeds	3,480	0.2	0.1	0.6	10.9	8.0	13.8	0.02	(0.00)	0.03
	3	Animal Foods	3,480	0.3	0.2	0.5	4.4	2.0	6.8	0.02	0.00	0.03
	4	Fruits and Vegetables	3,480	0.2	0.2	1.0	1.1	0.6	1.6	0.08	0.05	0.11
	5	Oils and Fats	3,480	0.1	0.1	1.1	1.5	0.2	2.8	0.02	0.00	0.04
	6	Sweets	3,480	0.0	0.0	13.4	2.2	0.2	4.1	0.00	(0.00)	0.01
	7	CoNA	3,480	1.2	0.4	0.3	1.3	0.8	1.8	0.13	0.08	0.17
Regional CoNA (LCU)	1	Balaka	120	134.1	91.8	0.7	0.8	(0.1)	1.7	22.4%	6.6%	40.6%
	2	Karonga	240	142.5	75.4	0.5	2.8	1.5	4.1	14.0%	2.2%	27.2%
	3	Mulanje	120	132.1	73.8	0.6	1.7	0.7	2.6	19.9%	3.9%	38.3%
	4	Chitipa	120	125.6	73.9	0.6	10.7	7.0	14.4	7.2%	(1.4%)	16.6%
	5	Dedza	120	121.6	71.0	0.6	5.1	(1.1)	11.3	7.0%	(7.7%)	24.0%
	6	Mzimba	240	145.9	81.3	0.6	2.4	0.9	3.9	14.0%	4.7%	24.2%
	7	Zomba	120	129.2	73.8	0.6	1.9	(0.5)	4.2	9.1%	(7.7%)	28.8%
	8	Kasungu	120	95.1	46.7	0.5	10.0	7.7	12.3	28.4%	4.9%	57.1%
	9	Machinga	120	143.6	86.5	0.6	6.4	(3.9)	16.6	2.0%	(10.1%)	15.8%
	10	Blantyre	120	105.5	59.4	0.6	0.7	(0.1)	1.5	28.5%	11.1%	48.6%

11	Mangochi	120	129.2	73.8	0.6	1.9	(0.5)	4.2	9.1%	(7.7%)	28.8%
12	Chiradzulu	120	113.5	67.4	0.6	1.0	(0.2)	2.2	11.8%	0.6%	24.2%
13	Mchinji	120	96.1	43.9	0.5	2.6	0.4	4.9	14.2%	(5.2%)	37.7%
14	Lilongwe	360	119.4	58.3	0.5	0.3	(1.0)	1.5	17.1%	4.3%	31.5%
15	Dowa	120	138.7	64.9	0.5	11.7	10.5	12.8	35.2%	15.5%	58.2%
16	Mwanza	120	134.5	72.6	0.5	1.5	0.6	2.4	13.0%	2.7%	24.3%
17	Chikwawa	120	138.5	79.2	0.6	1.3	0.4	2.1	16.9%	4.9%	30.2%
18	Nkhata Bay	120	142.1	83.4	0.6	3.6	(1.4)	8.7	4.7%	(4.5%)	14.8%
19	Nkhotakota	120	147.6	72.8	0.5	0.9	(1.0)	2.7	15.6%	(5.5%)	41.4%
20	Nsanje	120	124.7	72.7	0.6	0.8	(0.7)	2.3	7.8%	(1.0%)	17.5%
21	Ntcheu	120	131.5	85.0	0.7	0.7	(0.4)	1.9	24.8%	4.0%	49.7%
22	Phalombe	120	134.3	80.7	0.6	2.8	1.6	3.9	16.2%	5.0%	28.7%
23	Rumphi	120	141.1	79.1	0.6	3.1	(1.4)	7.5	4.4%	(5.4%)	15.1%
24	Salima	120	133.2	73.9	0.6	6.0	3.2	8.8	17.0%	(7.0%)	47.2%
25	Thyolo	120	131.3	63.2	0.5	0.4	(1.1)	2.0	12.3%	0.4%	25.7%

Note: Peak time shows as a continuous variable from 0-12, where 0 indicates the start and 12 as the end of a calendar year.

Table S6. Descriptive statistics and seasonality for national CoNA, CoCA, CoNA premium, energy composition and cost components of CoNA by food groups, and regional CoNA in Ethiopia between 2002-2016

Categories	No	Items	Obs.	Mean	SD	CV	Peak Time			Seasonal Intensity		
							Estimate	95% CI		Estimate	95% CI	
Indicators (LCU)	1	CoNA	20,704	6.7	5.1	0.8	6.6	6.1	7.2	4.0%	2.5%	5.5%
	2	CoCA	20,806	1.9	1.2	0.6	7.9	7.7	8.1	13.9%	12.2%	15.6%
	3	CoNA premium	20,704	4.8	4.3	0.9	3.6	2.9	4.3	5.3%	3.8%	6.8%
CoNA Energy Composition (kcal)	1	Starchy Staples	20,704	1,088.7	371.7	0.3	3.3	2.5	4.2	50.9	31.6	70.2
	2	Pulses, Nuts and Seeds	20,704	521.8	387.9	0.7	9.3	6.1	12.4	14.6	(5.3)	34.6
	3	Animal Foods	20,704	277.0	131.7	0.5	7.2	5.8	8.5	8.1	0.7	15.5
	4	Fruits and Vegetables	20,704	120.4	128.5	1.1	9.8	8.7	11.0	15.8	9.2	22.5
	5	Oils and Fats	20,704	99.1	152.1	1.5	9.8	8.8	10.8	18.1	11.7	24.4
	6	Sweets	20,704	0.7	14.7	22.7	6.7	4.6	8.7	0.5	(0.4)	1.4
CoNA and Cost Components (2011USD)	1	Starchy Staples	20,704	0.23	0.11	0.5	7.1	6.2	7.9	0.01	0.01	0.02
	2	Pulses, Nuts and Seeds	20,704	0.21	0.14	0.7	6.3	5.1	7.4	0.01	0.00	0.02
	3	Animal Foods	20,704	0.67	0.27	0.4	6.3	4.4	8.1	0.01	(0.00)	0.02
	4	Fruits and Vegetables	20,704	0.17	0.26	1.5	5.9	5.1	6.7	0.03	0.01	0.04
	5	Oils and Fats	20,704	0.06	0.13	2.2	9.7	8.4	11.1	0.01	0.01	0.02
	6	Sweets	20,704	0.00	0.01	23.2	6.4	4.4	8.4	0.00	(0.00)	0.00
	7	CoNA	20,704	1.34	0.48	0.4	6.6	6.0	7.2	0.06	0.03	0.08
Regional CoNA (LCU)	1	North Western Tigray	360	6.6	4.0	0.6	8.7	7.3	10.1	7.4%	3.3%	11.7%
	2	Central Tigray	358	6.1	3.6	0.6	9.1	7.8	10.4	6.4%	3.6%	9.2%
	3	Eastern Tigray	357	6.7	4.8	0.7	8.4	6.2	10.6	5.4%	0.2%	10.9%
	4	South Tigray	360	6.9	5.8	0.8	9.3	7.6	10.9	9.4%	(0.3%)	19.9%
	5	Awsi (Zone 1)	360	7.4	4.7	0.6	9.6	7.3	12.0	5.5%	1.5%	9.5%
	6	Gabi (Zone 3)	360	7.3	4.8	0.7	9.0	7.1	10.9	4.3%	1.4%	7.3%
	7	North Gonder	359	6.2	3.7	0.6	6.7	1.6	11.8	2.0%	(6.4%)	11.2%
	8	South Gonder	360	5.2	3.0	0.6	7.5	6.1	8.8	9.1%	1.2%	17.6%
	9	North Wollo	360	5.9	3.7	0.6	10.3	6.1	14.6	2.5%	(1.6%)	6.7%
	10	South Wollo	359	5.5	3.3	0.6	9.6	7.8	11.3	4.3%	1.8%	6.9%

Categories	No	Items	Obs.	Mean	SD	CV	Peak Time			Seasonal Intensity		
							Estimate	95% CI		Estimate	95% CI	
	11	North Shewa (R3)	358	6.9	5.6	0.8	8.1	5.9	10.2	5.4%	(0.5%)	11.7%
	12	East Gojam	359	5.4	3.2	0.6	5.3	2.2	8.3	2.7%	(2.0%)	7.6%
	13	West Gojam	359	5.2	3.3	0.6	8.3	5.9	10.7	5.2%	(0.8%)	11.5%
	14	Wag Himra	330	7.4	4.5	0.6	8.2	6.5	9.9	9.9%	2.3%	18.1%
	15	Awi/Agew	360	5.4	3.4	0.6	8.9	6.1	11.7	4.7%	0.8%	8.7%
	16	Oromia	180	5.4	3.0	0.5	7.1	6.0	8.2	6.0%	0.8%	11.5%
	17	West Wellega	359	6.1	3.8	0.6	7.3	5.7	8.8	4.7%	(0.6%)	10.2%
	18	East Wellega	360	6.3	4.2	0.7	5.2	3.7	6.8	11.7%	3.2%	20.9%
	19	Ilubabor	360	5.6	3.8	0.7	6.7	5.5	7.9	6.6%	0.0%	13.6%
	20	Jimma	360	5.8	4.0	0.7	3.7	(0.1)	7.5	2.1%	(0.5%)	4.9%
	21	West Shewa	358	5.9	3.5	0.6	0.8	(1.1)	2.7	5.2%	(0.3%)	10.9%
	22	North Shewa (R4)	356	7.6	5.9	0.8	7.7	5.4	9.9	6.6%	(2.4%)	16.3%
	23	East Shewa	360	6.0	3.6	0.6	6.3	4.7	7.8	3.1%	(0.9%)	7.4%
	24	Arsi	356	6.4	5.3	0.8	6.1	5.3	6.9	13.0%	3.0%	24.0%
	25	West Harerge	358	6.5	4.3	0.7	5.9	5.3	6.4	13.6%	7.7%	19.9%
	26	East Harerge	360	6.6	4.5	0.7	4.0	2.4	5.5	12.7%	6.4%	19.3%
	27	Bale	355	6.2	4.0	0.6	5.7	4.0	7.4	4.1%	(0.8%)	9.3%
	28	Borena	359	7.4	5.0	0.7	3.2	(0.9)	7.3	3.5%	(1.2%)	8.4%
	29	Siti	360	14.0	14.0	1.0	10.7	7.5	13.9	7.1%	(3.0%)	18.3%
	30	Fafan	360	8.4	5.3	0.6	2.9	1.0	4.8	6.7%	0.2%	13.6%
	31	Liben	359	9.7	7.4	0.8	6.0	4.4	7.6	7.7%	(2.1%)	18.5%
	32	Metekel	359	7.1	5.5	0.8	5.6	4.7	6.6	13.9%	4.8%	23.7%
	33	Asosa	360	6.2	4.0	0.7	7.3	3.9	10.8	3.3%	(4.6%)	11.8%
	34	Kemashi	355	6.8	4.6	0.7	4.7	4.0	5.3	25.2%	16.9%	34.1%
	35	Gurage	359	6.3	4.1	0.7	6.7	2.8	10.6	1.9%	(4.1%)	8.2%
	36	Hadiya	360	6.1	3.7	0.6	4.6	2.7	6.6	7.0%	(0.1%)	14.6%
	37	Kembata Timbaro	359	7.8	6.2	0.8	6.6	2.9	10.2	2.5%	(5.1%)	10.6%
	38	Sidama	356	6.1	3.9	0.6	5.9	4.7	7.1	7.2%	0.5%	14.3%
	39	Gedio	360	7.5	5.5	0.7	0.1	(9.4)	9.6	0.8%	(4.6%)	6.6%

Categories	No	Items	Obs.	Mean	SD	CV	Peak Time			Seasonal Intensity		
							Estimate	95% CI		Estimate	95% CI	
	40	Wolayita	359	6.2	5.1	0.8	5.4	3.2	7.7	5.0%	(2.2%)	12.8%
	41	South Omo	347	5.9	4.5	0.8	6.3	4.5	8.1	7.1%	(3.5%)	18.9%
	42	Sheka	360	5.3	3.3	0.6	6.1	4.0	8.1	3.8%	(2.4%)	10.4%
	43	Keffa	360	5.3	3.6	0.7	4.8	2.2	7.3	6.1%	(0.7%)	13.5%
	44	Gamo Gofa	359	5.6	3.5	0.6	5.4	2.4	8.4	3.5%	(3.0%)	10.4%
	45	Bench Maji	350	5.4	3.3	0.6	2.8	(4.7)	10.3	2.6%	(4.5%)	10.3%
	46	Yem	177	8.5	6.2	0.7	5.5	4.9	6.0	36.7%	22.7%	52.3%
	47	Segen Peoples'	717	6.5	4.8	0.8	6.4	5.5	7.2	10.2%	2.5%	18.4%
	48	Dawro	350	7.8	5.4	0.7	3.7	2.0	5.5	10.7%	(2.9%)	26.3%
	49	Basketo	180	6.3	5.4	0.9	6.1	4.9	7.3	13.8%	0.3%	29.1%
	50	Konta	177	8.3	5.8	0.7	11.5	10.0	13.0	6.7%	0.2%	13.6%
	51	Nuer	175	7.5	4.6	0.6	6.1	5.1	7.1	14.4%	2.4%	27.7%
	52	Agnuak	222	8.1	9.1	1.1	6.2	5.2	7.2	27.7%	3.4%	57.8%
	53	Majang	97	7.6	2.6	0.4	0.8	(1.1)	2.7	8.5%	(4.4%)	23.0%
	54	Hareri	180	6.5	3.8	0.6	0.6	(26.2)	27.4	0.2%	(4.1%)	4.7%
	55	Region 14	1,968	7.2	4.7	0.7	8.9	8.3	9.4	8.0%	5.8%	10.1%
	56	Chirkos	179	6.1	4.2	0.7	8.5	7.5	9.5	11.9%	6.8%	17.3%
	57	Dire Dawa	360	8.7	6.8	0.8	9.5	5.5	13.5	4.2%	(1.8%)	10.6%

Note: Peak time shows as a continuous variable from 0-12, where 0 indicates the start and 12 as the end of a calendar year.

Table S7. Example least-cost diets in Tanzania, Malawi and Ethiopia, by month (g/day)

Tanzania (Kigoma, 2015)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
White Maize grains	352	352	352	352	352	358	358	358	358	358	358	335
Cassava flour												23
Sweet Potatoes	59	59	59	59	59	64	64	64	64	64	64	69
Soybeans	172	172	172	172	172	164	164	164	164	164	164	168
Dried fish	17	17	17	17	17	17	17	17	17	17	17	17
Green pepper	18	18	18	18	18	14	14	14	14	14	14	10
Mchicha (amaranth leaves)						19	19	19	19	19	19	

Malawi (Kasungu, 2016)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Maize grain	355	311	305	252	175	175	281	132	334	299	334	272
White bread	104	78	92	188	143	143		249	57	68	57	192
Sweet potatoes				69	2	2			51	67	51	
Brown beans							30					
Cowpeas	24	16	13									
Groundnuts									114	130	114	
Chicken eggs		232	225				356	405				
Powdered milk					164	164	58	61				
Beef	75			75								
Sardine									10	10	10	10
Mangoes					116	116						
Pumpkin	116		31									112
Rape leaves		85							84	3	84	
Avocado					35	35						
Cabbage				159								136
Guava	22		25				23	24		25		
Cooking oil	25	43	44	21	21	21	19	15				63
Cooking oil, refill				27								

Ethiopia (Mecha, West Gojam Zone, 2016)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Maize (white)	244	254	292	279	269	309	176	175	326	175	185	295
Wheat white							172	141		135	67	
African millet	32	50	43	29	48	39		37	21	40	47	28
Sweet potato				16								
Chick peas	25	14		15	10							21
Horse beans	21	13			6							
Lima beans						47					41	
Ground nut shelled				36								
Cow milk (unpasteurized)	444	444	444	444	444	444	444	444	444	444	444	444
Sunflower seeds						18						
Ethiopian kale	35	2		12		30	54		7	4	17	7
Pumpkin	421	320	197	807	271	279			692	47	238	774
Spinach		58	137		90	14	132	146	96	123	38	30
Cooking oil (imported)	44	44	43		43	11	46	46	15	46	47	17

Note: The study as a whole uses a total of 25,522 least-cost diets, drawing from a potential list of 71, 55 and 97 distinct food items in Tanzania, Malawi and Ethiopia respectively. To illustrate the composition of an individual diet, and the degree of substitution among items and between food groups over the course of a year, we selected the one market location with the lowest average CoNA in each country, and show the composition of its least-cost diets each month over the most recent calendar year for which prices were available. This reveals which specific food items are selected each month at comparable locations in the three countries, in terms of grams per day. Of the three locations shown, the simplest example is from Tanzania, where people in the lowest-cost district (Kigoma) can meet their requirements most affordably by including 19 g/day of a green leafy vegetable (mchicha) during the June -November period, with a corresponding reduction in the need for green pepper and soybean.

ARTICLE 3: A NOVEL DIET-HEALTH INDEX REVEALS WORSENING TRENDS AND WIDENING DISPARITIES IN THE HEALTH IMPACTS OF DIET IN THE UNITED STATES

SUMMARY

Background Poor diets are associated with poor health outcomes, but existing metrics of diet quality do not directly include health effects of diet. Using a novel international diet-health index (IDHI), we can measure diet-related health impacts from multiple dietary factor simultaneously, given a population's health status and most prevalent causes of death and disability.

Methods We obtained individual-level data on intake of 12 dietary factors and exposure to 2 metabolic risk factors from the National Health and Nutrition Examination Survey (NHANES), 2003-2014, and computed the IDHI for 12 cardio-metabolic diseases and 15 cancers in the U.S. by sex, race, education and income. We then compared IDHI to a modified Alternative Healthy Eating Index (mAHEI) using 10 of the 12 dietary factors, and validated the indices using the National Center for Health Statistics (NCHS) linked dataset for total mortality through 2015.

Findings IDHI declined from -0.314 (95% CI: -0.323, -0.305) in 2003/04 to -0.325 (-0.334, -0.316) in 2013/14 (P=0.007 for trend). Non-Hispanic Black Americans have persistently lower IDHI than other groups, and disparities in IDHI have widened over time by level of income (P=0.004 for interaction), and education (P=0.047 for interaction). IDHI was more closely correlated with the mAHEI at higher levels of diet quality, and both indices were strongly associated with total mortality.

Interpretation The IDHI is a valid tool for measuring diet-related health impacts in the context of a population's most prevalent diseases, potentially offering tailored guidance regarding how best to reduce diet-related health disparities.

INTRODUCTION

Dietary risk factors are the leading avoidable cause of death and disability, accounting for 11 million deaths including 22% of global adult mortality in 2017.[104] In the United States, poor diets were linked to more than 500,000 deaths in 2016, including 5 of the top 10 leading causes of death namely ischemic heart disease, colon and rectum cancer, diabetes, intracerebral hemorrhage and ischemic stroke.[105] To monitor dietary quality using standardized methods, several diet quality indices (DQIs) have been developed and applied, such as the Healthy Eating Index (HEI), the Alternative Healthy Eating Index (AHEI), and the Mediterranean Diet Score (MDS).[42,106,107] These and other DQIs are defined in terms of minimum and maximum levels of dietary components such as servings of fruit from 0 to 4, grams of whole gains from 0 to 75, or the percent of energy from *trans* fats from 0.5 to 4. Intake of each component is rescaled to a score such as 0 to 10, from lowest to highest diet quality, and the scores are summed to an overall DQI. By definition, variation in intake outside the scored range does not affect the DQI, and the importance of each component does not depend on differences in mediating conditions or disease prevalence that might affect diet-related health risks.[45]

This study uses a new kind of international diet-health index (IDHI) that scores the full range of intake for each risk factor, adding up their impacts based on previously estimated relative risks for each disease, plus the mediating effect of individual health status, in proportion to the prevalence of each disease in each sub-population. The mathematical formula for IDHI is designed so that each parameter is drawn directly from meta-analyses of relative risks or nationally representative survey data on the prevalence of each condition, and the index can be updated or expanded to include additional risk factors and disease outcomes as new data become available.

The IDHI was originally developed to compare diet quality across countries, using dietary intakes aggregated by country, age, and sex for a single year.[10] Building on and extending that earlier work, this study computes change over time and different sub-populations in the United States, using individual-level data for 12 dietary factors and 2 mediators affecting risk for 27 diseases whose prevalence differs by race/ethnicity, income and education. We then compared IDHI with a modified AHEI (mAHEI) dietary quality score using the same 10 of 12 dietary factors applied to the scoring mechanism of the AHEI,[106] and validated the indices against total mortality.

METHODS

Data inputs and computation of the IDHI

Variation in the IDHI within each subpopulation comes from differences in dietary intake, and from mediation of relative risks through BMI and blood pressure; variation between subpopulations comes from differences in the prevalence of each disease outcome for which diet is a risk factor. (Appendix 1) We used individual-level dietary intake, and metabolic risk factor data from the U.S. National Health and Nutrition Examination Surveys (NHANES), combined with etiologic effects of diet on disease from published meta-analyses,[108] and disability-adjusted life year (DALY) data from Global Health Data Exchange (GHDx), to build on and extend IDHI calculation methods from a previous study[10] to address disparities and time trends in diet-related health impacts on 12 cardio-metabolic diseases (CMD) and 15 diet-related cancers.[109] A full list of the CMD and cancer outcomes included in this analysis is included in Appendix 2, Table A1. NHANES is a nationally-representative survey of demographic, socioeconomic, dietary and health-related information from about 5,000 persons per year of survey.[110] Data utilized in this study were from adult participants aged 25 and over covering a 10-year time period in 6 cycles (from 2003–2004 to 2013–2014).

Dietary intake and metabolic risk factor data: From NHANES we used dietary intake data for 12 dietary risk factors, of which 7 are risk-reducing beneficial factors, and 5 are risk-increasing adverse factors. (Appendix 2, Table A2) Dietary records were collected by 24-hour recall. Since 2003, all participants have been administrated a telephone-based dietary recall after 3-10 days of their first in-person interview.[111] To correct for measurement error, we used the National Cancer Institute method to estimate usual intake estimates for all dietary factors. The NCI method is the preferred method for estimating usual intake distribution from 24-hour diet recalls.[112] All dietary variables were also adjusted for energy to 2000 kcal/d using the residual method.[113] Body mass index (BMI) is calculated from weight and height measured during physical examinations at the Mobile Examination Center, and blood pressure is self-reported through the NHANES questionnaire.

Data on diet-disease relationships: The IDHI reported here is based on relative risks of each dietary factor for 12 cardio-metabolic diseases and 15 cancers. Relative risks (RRs) for each disease, per unit increase in each dietary factor, were obtained from previous studies for each age-sex group.[10,108,109,114]

DALY data: The U.S. population's Disability-Adjusted Life Years (DALY) for each disease were obtained from the GHDx and varied by sex, age and year of survey.[115] (Appendix 2, Table A2.)

In the results section, we present total IDHI and also IDHI subset by two disease types, CMD (IDHI_{cmd}), and cancers ($\text{IDHI}_{\text{cancer}}$) as well as subset by risk-reducing dietary factors ($\text{IDHI}_{\text{beneficial}}$) and all risk-increasing dietary factors $\text{IDHI}_{\text{adverse}}$). (Appendix 2, Table A3)

Data inputs for comparison to mAHEI and validation against mortality

Comparison to mAHEI diet quality score

To identify differences and similarities between IDHI and a standard approach to diet quality measurement, we used the same 12 dietary risk factors in the scoring method of AHEI-2010 which has been widely used and validated against multiple health outcomes in various populations.[106,116,117] Our mAHEI excludes *trans*-fat intake due to the lack of data from NHANES,[118] and also omits alcohol due to lack of relative risk data needed for IDHI. The other modifications of the AHEI formula needed to make the data inputs identical were to reclassify legumes as vegetables rather than nuts/seeds, and fruit juices as fruits rather than SSBs. (Appendix 2, Table A3)

Rescaling of IDHI and mAHEI for comparability

For comparison and validation, we rescaled IDHI and mAHEI to a range of 0-100, where the minimum of the indices was set to be 0 and the maximum value was 100, and rescaled each index to preserve its distribution by applying formula below:

$$(1) \qquad Index\_rescaled = \frac{(\text{Index} - Index_{min}) \times 100}{Index_{max} - Index_{min}}$$

Validation against total mortality

Total mortality was obtained from 2015 Public-Use Linked Mortality data by the National Center for Health Statistics (NCHS), which links mortality status of participants in 1999-2014 NHANES using a probabilistic match method. Each eligible participant is assigned a mortality status code equal to 0 if assumed alive and 1 if assumed deceased, with elapsed time recorded as months from the date of interview to the date of death or the end of the mortality period (December 31, 2015).[119] In regression models, we controlled for individual covariates including age, sex, race/ethnicity, education, income and smoking obtained from NHANES household interviews. Entering these covariates directly in the regression tests whether they have an independent association with mortality, beyond their role in the IDHI itself.

Statistical Analysis

We computed IDHI, and analyzed trends over time and disparities by sex, race, education and income using general linear regression models over 24,839 individuals. To estimate the mean and its 95% CI for IDHI in each group, we treated each 2-year survey cycle as a categorical variable, and also included the participant's racial category (Hispanic, Non-Hispanic White, Non-Hispanic Black, and all other racial categories), sex (dichotomous), age (continuous), education (dichotomous, as some college education or not) and household income (dichotomous, as a household's income is above 200% of the HHS poverty line), and their interaction terms with survey round. To include a linear time trend, we treated survey year as a continuous variable. To evaluate the statistical significance of population differences in trends over time, we introduced interaction terms between time and subpopulation groups as categorical variables. Sampling weights were included to account for the complex design of NHANES. We then examined the correlation between IDHI and mAHEI for a total of 26,790 observations and reported both Pearson's and Spearman's coefficients and the p-values. The sample size for these comparisons is slightly larger than for the analysis of trends by subgroup, due to missing covariates for some respondents.

Finally, we validated the indices against the total mortality in the US using both Cox proportional hazards and Poisson regression models with robust standard errors, with a sample size of 24,810 due to missing covariates. The proportional hazards assumptions passed log-log and Schoenfeld residual tests. (Appendix 3) Three multivariate models were used: Model 1 adjusted for age (continuous and age-squared), sex and race as described above; Model 2 also adjusted for smoking (dichotomous, as having smoked 100 or more cigarettes in lifetime or not), education (dichotomous, as having some college education or not) and household income (dichotomous, as having a household income above 200% of the HHS poverty line or not). All analyses and data visualizations were conducted using Stata/SE version 15.1 and RStudio 1.3.1093, and p-value <0.05 was considered statistically significant.

RESULTS

Distribution of IDHI in the US population between 2003 and 2014

The IDHI in the US population between 2003 and 2014 averaged -0.32 (IQR: -0.46, -0.12) , indicating greater detrimental health impacts from adverse dietary factors than ameliorative health effects from beneficial dietary factors. $IDHI_{adverse}$ is characterized by greater dispersion around its mean of -0.59 (IQR: -0.77, -0.34), compared to $IDHI_{beneficial}$ which has a mean of 0.27 (IQR: 0.20, 0.33). Decomposing the index by type of disease, CMDs accounted for almost all of the overall IDHI, with an average of -0.31 (IQR: -0.44, -0.11), while the index for all cancers averaged -0.01 (IQR: -0.01, 0.00). (Appendix 2, Table A3) These data are reported without rescaling, so as to identify the net direction of risk counting both adverse and beneficial dietary factors.

NHANES respondents with the highest tertile of IDHI were older, had higher household income and education level, and were more often female and non-smokers than those at lower levels of IDHI. (Table 1) Additionally, due to their higher age at the time of data collection between 2003 and 2014, participants in the highest tertile of IDHI were also more likely to have a recorded death by the end of 2015.

Table 1. Characteristics of US adults by IDHI, NHANES, 2003–2014

Characteristics	Total 24,810 (-2.5,-0.5)	Tertile 1 8,270 (-2.5,-0.4)	Tertile 2 8,270 (-0.4, -0.2)	Tertile 3 8,270 (-0.2, 0.5)	P-value
Age (years), mean (SD)	51.8 (16.6)	44.0 (12.0)	53.4 (15.4)	58.1 (18.6)	<0.001
Sex, n (%)					
Male	12,075 (48.7%)	6,588 (79.7%)	3,612 (43.7%)	1,875 (22.7%)	<0.001
Race/Ethnicity, n (%)					
Hispanic	5,709 (23.0%)	1,813 (21.9%)	2,022 (24.4%)	1,874 (22.7%)	<0.001
Non-Hispanic White	12,132 (48.9%)	3,439 (41.6%)	3,860 (46.7%)	4,833 (58.4%)	
Non-Hispanic Black	5,131 (20.7%)	2,635 (31.9%)	1,747 (21.1%)	749 (9.1%)	
Others	1,838 (7.4%)	383 (4.6%)	641 (7.8%)	814 (9.8%)	
Smoking status[1], n (%)					
Less than 100 cigarettes in life	12,997 (52.4%)	3,656 (44.2%)	4,247 (51.4%)	5,094 (61.6%)	<0.001
Income[2], n (%)					
Below 200% of HHS poverty line	11,284 (45.5%)	4,081 (49.3%)	3,809 (46.1%)	3,394 (41.0%)	<0.001
Education[3], n (%)					
No post-sec. (high school or less)	12,157 (49.0%)	4,482 (54.2%)	4,072 (49.2%)	3,603 (43.6%)	<0.001
Final Mortality Status					
Assumed alive	22,275 (89.8%)	7,816 (94.5%)	7,416 (89.7%)	7,043 (85.2%)	<0.001

1. Dichotomous variable: whether the interviewee has smoked fewer than 100 cigarettes in their life (or not);
2. Dichotomous variable: whether a household's income is below 200% of the HHS poverty line (or not);
3. Dichotomous variable: whether the highest education of the interviewee is high school or less (no college).

Trend analysis shows that IDHI declined from -0.314 (95% CI: -0.323, -0.305) in 2003/04 to -0.325 (95% CI: -0.334, -0.316) in 2013/14 (P=0.007 for trend). In subgroup analyses shown in Figure 1, IDHI is lower among men than among women, and the IDHI of Non-Hispanic Black population is significantly lower than for individuals in the other racial categories (P<0.001 for both). However, IDHI among women (P<0.001 for trend), Non-Hispanic Whites (P=0.029 for trend) and Hispanics (P=0.008 for trend) indicate declining trends over time. Additionally, disparities in IDHI between people from higher and lower income households (P=0.004 for interaction), as well as with higher and lower education levels (P=0.047 for interaction), widened over time (P<0.001 for trend of both).

Decomposing the trend analysis into IDHI calculated for adverse vs. beneficial dietary factors revealed no significant time trend for $IDHI_{adverse}$ but a decline in $IDHI_{beneficial}$ over time ($P<0.001$ for trend). In terms of disparities, the Non-Hispanic Black population had a significantly more negative $IDHI_{adverse}$ ($P<0.001$) with higher magnitudes but no difference in $IDHI_{beneficial}$, and there were widening gaps between people with higher and lower socioeconomic status in $IDHI_{adverse}$ ($P=0.009$ for the income interaction, and $P=0.022$ for education interaction), but not in the $IDHI_{beneficial}$. (Appendix 4, Figure A2 and A3) Comparing IDHI to a standard diet quality scoring method, we found a moderate improvement over time in the mAHEI ($P<0.001$ for trend), while disparities by socioeconomic group are generally similar between the two methods. (Appendix 4, Figure A4)

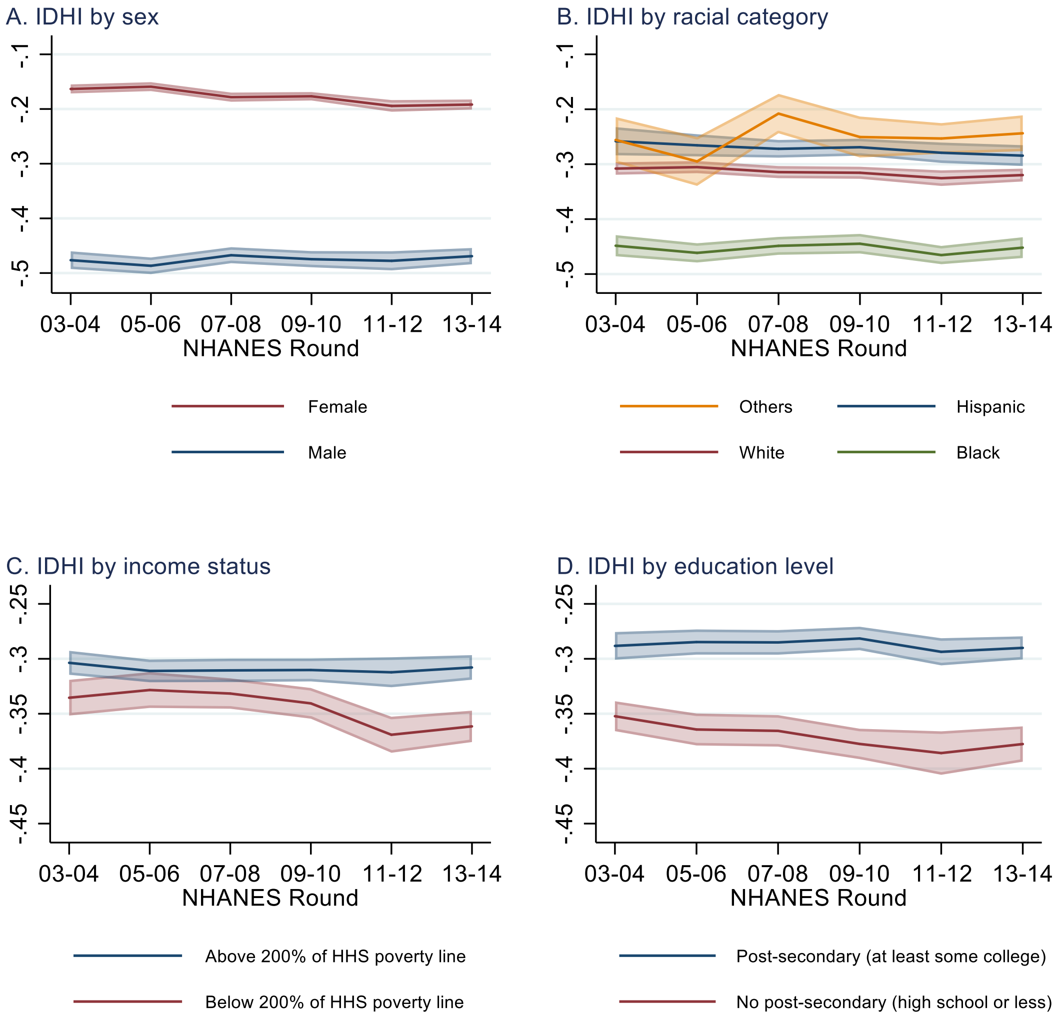


Fig 1. Trends in IDHI across sex, race, income and education groups in the US, 2003-2014 (N=24,839) Shaded areas indicate the 95% CI. Estimates were based on general linear models controlling for survey round as a categorical variable, participant's race, sex, age, education and family income and their interaction terms with survey round. In the trend estimation we assigned mean values for all covariates. In (B), White and Black refer to the non-Hispanic populations.

Comparison of IDHI with the mAHEI

Over the entire US population, the rescaled IDHI and its corresponding mAHEI metric are strongly correlated, with a Pearson's correlation coefficient of 0.705 ($P<0.001$) and a larger Spearman's rho of 0.783 ($P<0.001$). The nonlinear pattern of correlation is shown in Figure 2, revealing the wider range of IDHI at lower levels of mAHEI. This asymmetry arises in part because each component score of mAHEI is truncated at zero when intake of harmful dietary risk factors exceeds some threshold, whereas the IDHI varies with intake of harmful factors even beyond AHEI thresholds leading to a long-left tail of high-risk diets at the bottom of Figure 2.

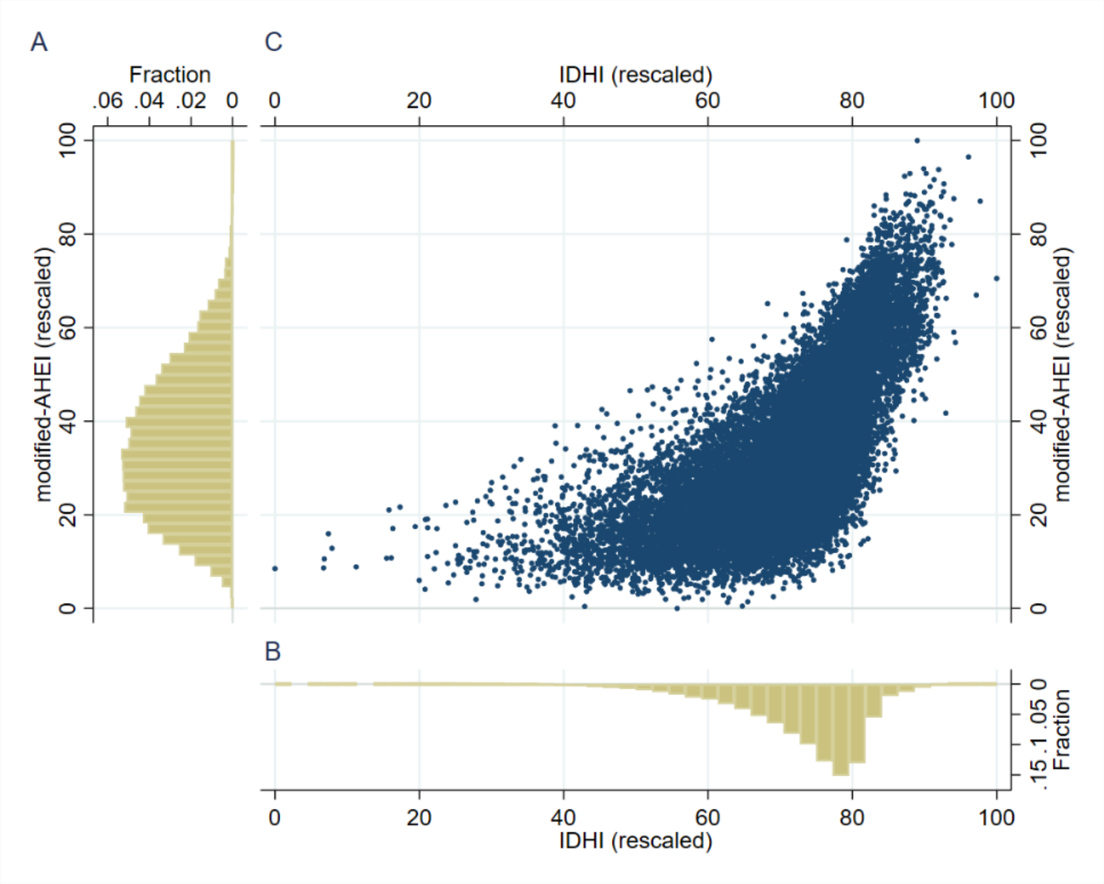


Fig 2. Rescaled IDHI and mAHEI in the US population, 2003-2014 (N=26,790)
(A) and (B) are the histogram distribution plots of the rescaled mAHEI and IDHI. Both indices were proportionally rescaled to a range of 0-100. (C) presents the relation of the two indices.

The IDHI approach differs from standard diet quality metrics such as AHEI and HEI not only in using the full range of dietary intake levels for each risk factor, but also using individual variation in health status (BMI and blood pressure) as mediators, and using variation of relative risks and disease prevalence across age-sex groups and over time. The influence of this difference on the correlation between IDHI and the corresponding mAHEI can be seen by comparing scatter plots like Figure 2 for people of different ages. Relative risks from dietary factors are smaller for older than younger people, due to the presence of competing risks, so the IDHI in older populations is lower[120] and more closely correlated with a standard diet quality metric such as our mAHEI. (Appendix 5, Figure A5)

Validation of IDHI against total mortality in the US population

We assessed the association between IDHI and all-cause mortality to validate the strength of its relationship with this key predictor of population health. We utilized the same transformed scale from 0 to 100 for IDHI for ease of comparison with standard diet quality metrics. We first visualized the dose-response relation of IDHI and Hazard Ratio (HR) from Cox models in Figure 3, which indicates findings from both the minimally adjusted model (Model 1) and fully adjusted model (Model 2). Regression coefficients of all co-variates from Cox and Poisson models are indicated in Figure 4. The HR from the Cox model indicated a significant inverse association between IDHI and mortality (HR 0.77, 95%CI: 0.72, 0.83) and was attenuated but maintained statistical significance in the fully adjusted model (HR 0.82, 95%CI: 0.77, 0.88) by 10 points increase of rescaled IDHI. Incidence Rate Ratio (IRR) from Poisson models presented very similar results, which are 0.81 (95%CI: 0.76, 0.86) in Model 1 and 0.87 (95%CI: 0.81, 0.93) in Model 2.

In addition, smoking was associated with higher mortality, while higher income and education level were negatively associated with mortality rates. Age was positively associated with mortality, revealed by the significant coefficient of the age squared term. Detailed regression results can be found in Appendix 6, Table A6-A7.

We also ran the validation models using the rescaled mAHEI and found similar results. The magnitude of correlation between mAHEI and mortality is generally smaller than between IDHI and mortality, but we do not attempt to text for a significant difference in their predictive power. (Appendix 6, Table A8-A9)

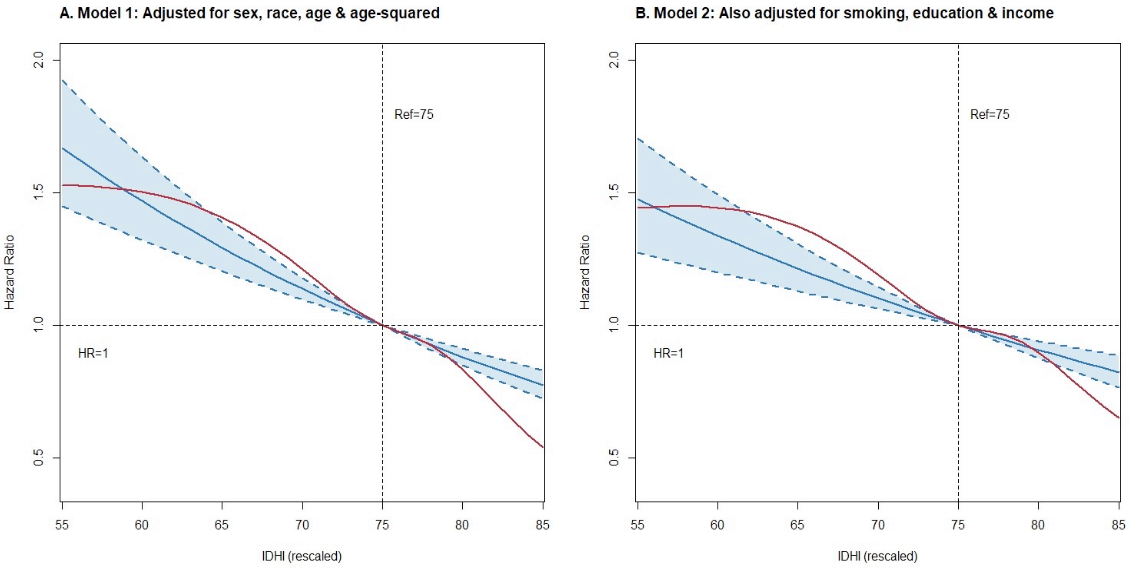


Fig 3. Dose response of IDHI for all-cause mortality

Adjusted hazard ratios (blue lines) with 95% CIs (light blue areas) and restricted cubic spline (red lines) for all-cause mortality. (A) and (B) presents the regression results by the Cox model. Model 1 controlled covariates of sex, race, and age, and Model 2 in addition adjusted smoking, education, and income levels. We show the range of rescaled IDHI from 55 to 85, which includes 90% of the U.S. population.

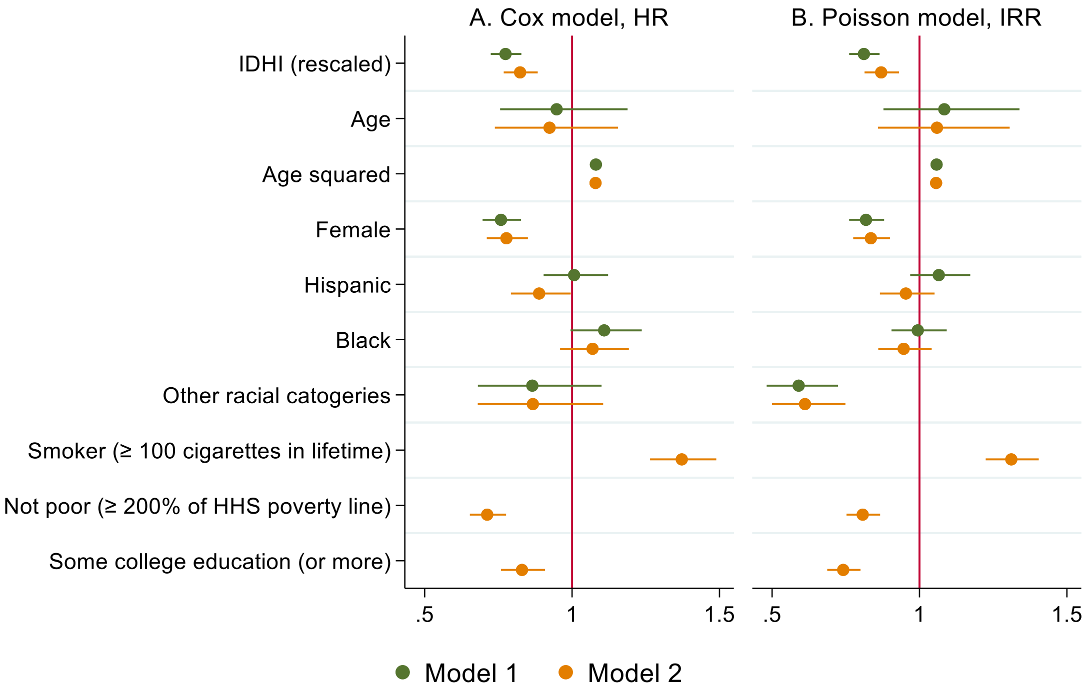


Fig 4. Coefficients of other risk factors for all-cause mortality

(A) and (B) presents the regression results by Cox model and Poisson model. Model 1 controlled covariates of sex, race and age and Model 2 in addition adjusted smoking, education and income levels. The risk ratio is based on 10 points increase in rescaled IDHI and 10 years increase in age.

DISCUSSION

In this study, we generated and reported a new diet-health index, the IDHI, in the US, using 11 years (2003-2014) of nationally representative data from NHANES. We identified worsening trends in the already negative health impacts of diet in the US population, as well as persistent disparities across sex and race/ethnicity and widening gaps between socioeconomic groups. In the US population, the IDHI is directly associated with both an mAHEI score, a standard measure of diet quality, and inversely related to total mortality, a strong metric of population health, indicating that the IDHI provides a valid measure of diet quality for health in a single index. In contrast to existing DQIs, IDHI simultaneously captures trends in both dietary quality and its relative impacts on diet-related chronic diseases across age, sex, race/ethnicity and socioeconomic status, while existing DQIs neither explicitly identify disease outcomes nor could be generalized beyond particular populations and dietary patterns.[10]

Previous studies found that the dietary quality of American people has been steadily improving over time, reflected in increasing trends in HEI and AHEI.[121–123] Meanwhile, in the US, dietary risks became the leading risk factor for deaths and the third leading risk factor for DALYs in 2016 due to increasing burdens of diet-related chronic diseases such as cardiovascular disease, diabetes and certain cancers.[105] Here IDHI presents a different perspective than DQIs, where improving dietary quality may not translate into improved population health impact of diets in the US, and the worsening trends in IDHI over the 11-year study period was largely due to the declining trends in $\text{IDHI}_{\text{benefical}}$ for the protective dietary factors. (Appendix 4, Figure A3) Different trends in IDHI v.s. in HEI or AHEI infer that although US population in general were eating healthier driven by higher intakes of whole grain, nuts and seeds, and reduction in SSBs[124] (Appendix 2, Table A2), the relative burdens of the diseases which are more related to harmful dietary factors (or less related to protective factors) were getting even larger. For example, the average DALY portion of type 2 diabetes increased from 10.6% to 13.9%, while ischemic heart disease decreased from 31.9% to 27.5%. (Appendix 2, Table A1)

Prior studies documented widening disparities in dietary quality among subgroups of the US population,[123,125] and also indicated enduring disparities in cardiovascular disease and diabetes across race and socioeconomic status in the US.[126,127] Our analysis of IDHI supports these prior findings and suggests possible reasons for such trends. For example, our analysis indicates persisting disparities in diet-related health impacts among Non-Hispanic Black Americans in comparison to other race/ethnic groups, primarily due to greater health impacts of adverse dietary factors, as measured by $\text{IDHI}_{\text{adverse}}$. Our analysis also reveals widening disparities over time between lower and higher socioeconomic groups that were also related to a widening gap in the $\text{IDHI}_{\text{adverse}}$, and a declining trend in IDHI for lower socioeconomic groups. (Appendix 4, Figure A2)

Our validation work confirmed a positive and nonlinear correlation between IDHI and the mAHEI, as well as their strong associations with total mortality in the US. IDHI may capture more sensitivity in the health impacts from harmful dietary factors and also presented an even stronger association with total mortality, comparing with the mAHEI. Our findings suggest IDHI an valid and efficient means for policymakers to evaluate the health effects of overall diet, dietary factors and components, on diet-related chronic diseases at population level.

There are several strengths of this study. First, as a scalable and decomposable index, we significantly expanded the IDHI to include both cardio-metabolic diseases and diet-related cancers in the analysis. Second, the nationally representative NHANES surveys allow us to construct the IDHI at individual level and to evaluate the IDHI trend at the population level stratified by sex, race/ethnicity and socioeconomic status, while prior study was on aggregate age/sex grouped data at national level. Third, the NHANES mortality linked data is considered as a large-scale prospective cohort study, leading to robust results in the validation work. Fourth, in the validation work, we applied various models and reach to robust and consistent findings.

Limitations of our work should also be noted. First, self-reported dietary information is subject to random and systematic error, which is only partly overcome by the high quality of NHANES data based on 2 standardized 24-hour diet recalls per person, adjusted to usual intake by the NCI method and for energy to 2000 kcal/d. Second, the IDHI only included the subset of all dietary factors and all diet-related diseases for which high-quality epidemiological data are available, which by definition does not include the totality of human diets and their impacts on health. This study extended coverage beyond the initial IDHI to include one more dietary factor (dairy) and 15 additional disease outcomes (cancers), but future applications of the IDHI can and should consider additional dietary factors and disease outcomes as well as updated parameters to the extent that new data become available over time. The same formula can accommodate any number of dietary factors, mediators or diseases, and moving beyond CMD and cancers to other health outcomes will be especially important for the use of IDHI in developing countries.[10] Third, we used homogeneous relative risk estimates in dietary factor-cancer pairs due to the lack of sufficient evidence to differentiate risks, but future applications of IDHI could use risks estimated from studies that address possible heterogeneity within each demographic group.[109] Fourth, we slightly modified the scoring mechanism in AHEI-2010 to make the dietary risk factors in line with the ones for IDHI. However, we can see that the magnitude of mAHEI is close to the AHEI without trans-fat component reported in the previous study[123] and the mAHEI is also significantly associated with mortality in the validation work. Finally, there are residual confounding factors that were not fully adjusted by our validation models due to missing data, such as physical activity level and alcoholic beverage intakes.

In conclusion, our findings reveal the usefulness of IDHI to measure and compare the health effects of individuals' dietary intake in the United States. Adding up the estimated effects of 12 dietary components (7 risk-reducing beneficial factors and 5 risk-increasing adverse factors) for a total of 27 disease outcomes (12 cardio-metabolic diseases and 15 cancers) reveals that their nationally representative total effect on disability-adjusted life years lost was negative and worsening between 2003 and 2014. Unlike previous diet quality indices, the IDHI takes account of disparities in the prevalence of each disease outcome affected by dietary intake, and also accounts for pre-existing mediators (BMI and blood pressure). After adjusting for sex, age, education and income, Black (non-Hispanic) respondents were found to bear a larger burden of diet-related disease than other racial categories, and the gap between socioeconomic groups was found to be widening over time. Interventions to improve treatment of diet-related diseases, as well as improvements in diet quality itself, would be needed to reverse those trends, eliminate disparities and reduce the high mortality rates currently experienced in the United States.

APPENDICES

Appendix 1. Mathematical formula and definition of IDHI

The IDHI is defined for a population as the sum of individuals' diet-related health risks associated with their intake of each dietary factor and its associated disease outcome. In this analysis, the IDHI for each individual can be described mathematically as follows:

$$(1) \qquad IDHI_{it} = \sum_{k=1}^{27} \left\{ \sum_{j=1}^{12} \left[Intake_{ijt} \times (-lnRR_{jkas}) \times \left(\frac{DALY_{kast}}{\sum_{k=1}^{27} DALY_{kast}} \right) \right] + \sum_{m=1}^{2} \sum_{j=1}^{12} \left[Intake_{ijt} \times \beta_{jm} \right] \times (-lnRR_{mkas}) \times \left(\frac{DALY_{kast}}{\sum_{k=1}^{27} DALY_{kast}} \right) \right\}$$

i = individual
t = time of survey round
j = dietary factor
k = disease outcome
m = mediator
a = age
s = sex
Intake = dietary intake
lnRR = natural log of relative risk
DALY = Disability-Adjusted Life Years
β = linear effect of dietary factor on mediator

In equation (1), subscripts denote the dimensions of variation for each component of IDHI, which starts with reported $Intake_{ijt}$ for each individual *i* of dietary factor *j* at the time of survey *t*; then $lnRR_{jkas}$ is the natural logarithm of relative risk per unit of intake for dietary factor *j* on the incidence of disease *k* in people of that individual's age *a* and sex *s*; $DALY_{kast}$ is the total DALYs associated with disease *k* among people of that individual's age and sex at that time. In this application of IDHI, we also consider individual variation in two health-status measures, BMI and blood pressure, whose mediating effect is denoted β_{jm} as the linear effect between mediator *m* and dietary factor *j*, which is also dependent on the age, sex, race and health conditions of BMI and blood pressure, and $lnRR_{mkas}$ is the natural logarithm of relative risk per unit of mediator factor *m* on the incidence of disease *k* in people of age *a* and sex *s*. In this formulation, mediation is linear, and the utilization of negative natural logarithm of the RR allows that a positive value indicates ameliorative health effects and a negative value indicate detrimental health effects.

Appendix 2. Summary of data inputs and components of IDHI and mAHEI

Table A1. DALY portions of diseases included in IDHI over NHNAES survey rounds

Disease	Stats	NHANES Survey Rounds					
		03-04	05-06	07-08	09-10	11-12	13-14
CMDs							
Aortic aneurysm	Mean	1.2%	1.1%	1.0%	1.0%	1.0%	0.9%
	SD	0.5%	0.5%	0.5%	0.5%	0.5%	0.5%
Atrial fibrillation and flutter	Mean	1.4%	1.5%	1.6%	1.7%	1.7%	1.8%
	SD	1.3%	1.4%	1.5%	1.6%	1.7%	1.7%
Cardiomyopathy and myocarditis	Mean	4.6%	4.5%	4.4%	4.3%	4.2%	4.2%
	SD	4.1%	4.0%	3.9%	3.8%	3.7%	3.6%
Endocarditis	Mean	0.8%	0.9%	0.9%	0.9%	0.9%	1.0%
	SD	0.5%	0.5%	0.6%	0.6%	0.7%	0.7%
Hypertensive heart disease	Mean	3.2%	3.4%	3.6%	3.8%	3.9%	4.1%
	SD	1.4%	1.5%	1.6%	1.7%	1.7%	1.8%
Hemorrhagic stroke	Mean	5.2%	5.1%	5.0%	4.9%	4.9%	5.0%
	SD	0.6%	0.6%	0.6%	0.6%	0.7%	0.7%
Ischemic heart disease	Mean	31.9%	30.8%	29.5%	28.3%	27.8%	27.5%
	SD	10.2%	9.7%	9.1%	8.7%	8.4%	8.2%
Ischemic stroke	Mean	5.6%	5.6%	5.6%	5.7%	5.8%	5.8%
	SD	3.1%	3.1%	3.0%	3.0%	3.0%	3.0%
Other cardiovascular and circulatory	Mean	4.0%	4.0%	4.0%	4.0%	4.1%	4.1%
	SD	2.9%	2.8%	2.8%	2.8%	2.8%	2.8%
Peripheral artery disease	Mean	0.4%	0.4%	0.5%	0.5%	0.6%	0.6%
	SD	0.3%	0.4%	0.4%	0.5%	0.5%	0.5%
Rheumatic heart disease	Mean	0.9%	0.8%	0.8%	0.8%	0.8%	0.8%
	SD	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%
Type 2 diabetes mellitus	Mean	10.6%	11.5%	12.7%	13.6%	13.9%	13.9%
	SD	2.1%	2.3%	2.5%	2.6%	2.6%	2.6%
Cancers							
Breast (postmenopausal)	Mean	7.5%	7.3%	7.2%	7.1%	7.0%	6.9%
	SD	8.9%	8.7%	8.4%	8.2%	8.1%	7.9%
Colon and rectum	Mean	6.4%	6.4%	6.4%	6.4%	6.4%	6.4%
	SD	0.7%	0.7%	0.7%	0.7%	0.8%	0.8%
Corpus uteri	Mean	0.6%	0.7%	0.7%	0.7%	0.8%	0.8%
	SD	0.7%	0.8%	0.8%	0.9%	0.9%	1.0%
Esophageal	Mean	1.3%	1.3%	1.3%	1.3%	1.3%	1.3%
	SD	0.9%	0.9%	1.0%	1.0%	1.0%	0.9%
Gallbladder	Mean	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%
	SD	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%
Kidney	Mean	1.3%	1.3%	1.3%	1.3%	1.3%	1.3%
	SD	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%

Liver	Mean	1.6%	1.7%	1.8%	1.9%	2.0%	2.1%
	SD	0.7%	0.8%	0.9%	1.0%	1.0%	1.1%
Multiple myeloma	Mean	0.9%	0.9%	0.9%	0.9%	0.9%	0.9%
	SD	0.4%	0.4%	0.4%	0.4%	0.4%	0.4%
Mouth, pharynx, and larynx	Mean	2.4%	2.4%	2.3%	2.3%	2.3%	2.3%
	SD	1.3%	1.2%	1.2%	1.2%	1.2%	1.2%
Ovary	Mean	1.9%	1.8%	1.8%	1.8%	1.8%	1.7%
	SD	2.1%	2.0%	2.0%	1.9%	1.9%	1.9%
Pancreas	Mean	2.7%	2.8%	2.9%	3.0%	3.0%	3.0%
	SD	1.1%	1.1%	1.2%	1.3%	1.3%	1.3%
Prostate (advanced)	Mean	1.6%	1.6%	1.6%	1.6%	1.6%	1.6%
	SD	2.7%	2.6%	2.6%	2.6%	2.6%	2.6%
Stomach cancer (noncardia)	Mean	1.0%	1.0%	1.0%	1.0%	1.0%	1.0%
	SD	0.3%	0.3%	0.3%	0.3%	0.3%	0.3%
Stomach (cardia)	Mean	0.6%	0.6%	0.6%	0.6%	0.6%	0.6%
	SD	0.2%	0.2%	0.2%	0.2%	0.2%	0.2%
Thyroid	Mean	0.2%	0.2%	0.3%	0.3%	0.3%	0.3%
	SD	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%

Note: CMD outcomes included were: ischemic heart disease (International Classification of Diseases, 10th Revision codes I20–I25), ischemic stroke (I63, I65–I67 (except I67.4) and I69.3), hemorrhagic stroke I60–I62, I69.0–I69.2 and I67.4), atrial fibrillation and flutter (I48), aortic aneurysm (I71), hypertensive heart disease (I11), endocarditis (I33), peripheral artery disease (I73 and I70.2), cardiomyopathy and myocarditis (I42 and I40), rheumatic heart disease (I01, I02.0 and I05–I09), type 2 diabetes mellitus (E11 (except E11.2)) and other cardiovascular and circulatory diseases (I00, I02.9, I27–I28 (except I27.1), I30–I32 (except I31.2 and I31.3), I34–I39, I47, I70.8, I72, I77–I80, I82–I84 and I86–I98). Cases for individual cancer types were obtained by applying the International Classification for Diseases for Oncology, third edition (ICD-O-3) codes corresponding to primary cancer site. Additional specifications on tumor histologic types and anatomic locations were used to obtain the incident cancer cases for esophageal adenocarcinoma and stomach cardia and non-cardia cancers.

Table A2. Intakes of dietary risk factors in the US over NHNAES survey rounds (g/day)

Dietary factors	**Indices**	**Stats**	**NHANES Survey Rounds**					
			03-04	**05-06**	**07-08**	**09-10**	**11-12**	**13-14**
Risk-reducing factors								
Dairy products	IDHI-cancer	Mean	366.4	377.5	360.8	380.8	361.7	369.5
		SD	169.4	164.9	153.8	162.8	159.7	163.4
Fruits*	IDHI-cancer/CMD, AHEI	Mean	106.9	104.7	103.3	106.0	104.0	100.6
		SD	65.0	63.2	62.9	65.9	64.5	64.6
Nuts/Seeds*	IDHI-cancer/CMD, AHEI	Mean	15.7	16.4	15.6	16.3	18.4	18.3
		SD	15.7	15.8	15.1	16.1	18.9	18.5
Seafood-omg3	IDHI-CMD, AHEI	Mean	0.08	0.08	0.08	0.08	0.09	0.09
		SD	0.05	0.05	0.05	0.05	0.05	0.05
Vegetables*	IDHI-cancer/CMD, AHEI	Mean	163.9	164.8	162.6	162.8	163.5	161.5
		SD	44.8	44.2	44.1	45.2	45.7	45.0

Whole Grains	IDHI-cancer/CMD, AHEI	Mean	20.1	20.5	20.7	22.1	24.1	23.8
		SD	14.4	14.3	14.1	15.5	17.8	17.4
PUFA	IDHI-CMD, AHEI	Mean	7.4	7.4	7.5	7.5	7.9	7.9
		SD	1.2	1.2	1.2	1.2	1.3	1.3
Risk-increasing factors								
Processed meats	IDHI-cancer/CMD, AHEI	Mean	39.2	42.1	43.9	43.6	45.8	44.1
		SD	22.0	24.5	22.1	22.3	22.1	22.0
Red meats	IDHI-cancer/CMD, AHEI	Mean	166.9	158.5	166.2	166.0	164.0	161.4
		SD	73.5	77.2	64.9	64.8	63.8	64.4
Sodium	IDHI-CMD, AHEI	Mean	3.3	3.4	3.3	3.4	3.5	3.4
		SD	0.9	0.9	0.9	0.9	0.9	0.9
Saturated fats	IDHI-CMD	Mean	11.0	11.1	11.0	10.9	10.8	11.0
		SD	1.6	1.6	1.6	1.6	1.6	1.6
SSB[*]	IDHI-cancer/CMD, AHEI	Mean	347.9	334.2	326.9	335.7	353.8	327.8
		SD	378.3	364.5	348.5	349.8	362.6	344.5

Note: Legumes are categorized in vegetables, not nuts and seeds. Fruit juices are categorized in fruit group, not SSB.

Table A3. Components in IDHI and mAHEI (N=26,790) in the US, 2003-2014[1]

Component	mean	SD	min	p25	p50	p75	max
IDHI	**-0.32**	**0.29**	**-2.52**	**-0.46**	**-0.25**	**-0.12**	**0.51**
$\text{IDHI}_{\text{adverse}}$	-0.59	0.32	-2.84	-0.77	-0.50	-0.34	-0.07
$\text{IDHI}_{\text{beneficial}}$	0.27	0.10	0.08	0.20	0.25	0.33	1.15
IDHI_{cmd}	-0.31	0.28	-2.47	-0.44	-0.24	-0.11	0.51
$\text{IDHI}_{\text{cancer}}$	-0.01	0.01	-0.05	-0.01	-0.01	0.00	0.04
mAHEI[2]	**30.9**	**8.5**	**11.5**	**24.3**	**30.2**	**36.8**	**66.4**
Fruits[4]	2.57	1.59	0.50	1.24	2.26	3.52	10.00
Nuts/Seeds[4]	5.22	3.03	1.24	2.67	3.76	8.36	10.00
Long-chain (n-3) fats[3]	3.43	1.62	1.57	2.58	2.97	3.63	10.00
PUFA	7.04	1.44	0.19	6.03	6.98	8.01	10.00
Processed/red meats	0.26	1.32	0.00	0.00	0.00	0.00	10.00
Sodium	4.26	2.85	0.00	2.00	4.00	7.00	9.00
SSB[4]	3.47	3.45	0.00	0.00	3.07	6.75	10.00
Vegetables[4]	1.88	0.52	0.63	1.54	1.85	2.19	9.16
Whole Grains	2.80	1.89	0.43	1.17	2.36	3.86	10.00

Note:

1. Data were adjusted for NHANES survey weights to be nationally representative;
2. We do not consider alcohol intakes and trans-fat in the mAHEI due to lack of data;
3. Being consistent with IDHI, include seafood omega-3 only in the analysis;
4. Being consistent with IDHI, legumes are categorized as vegetables, not nuts and seeds; fruit juices are in fruit group, not SSB.

Appendix 3. Proportional Hazard (PH) assumption tests

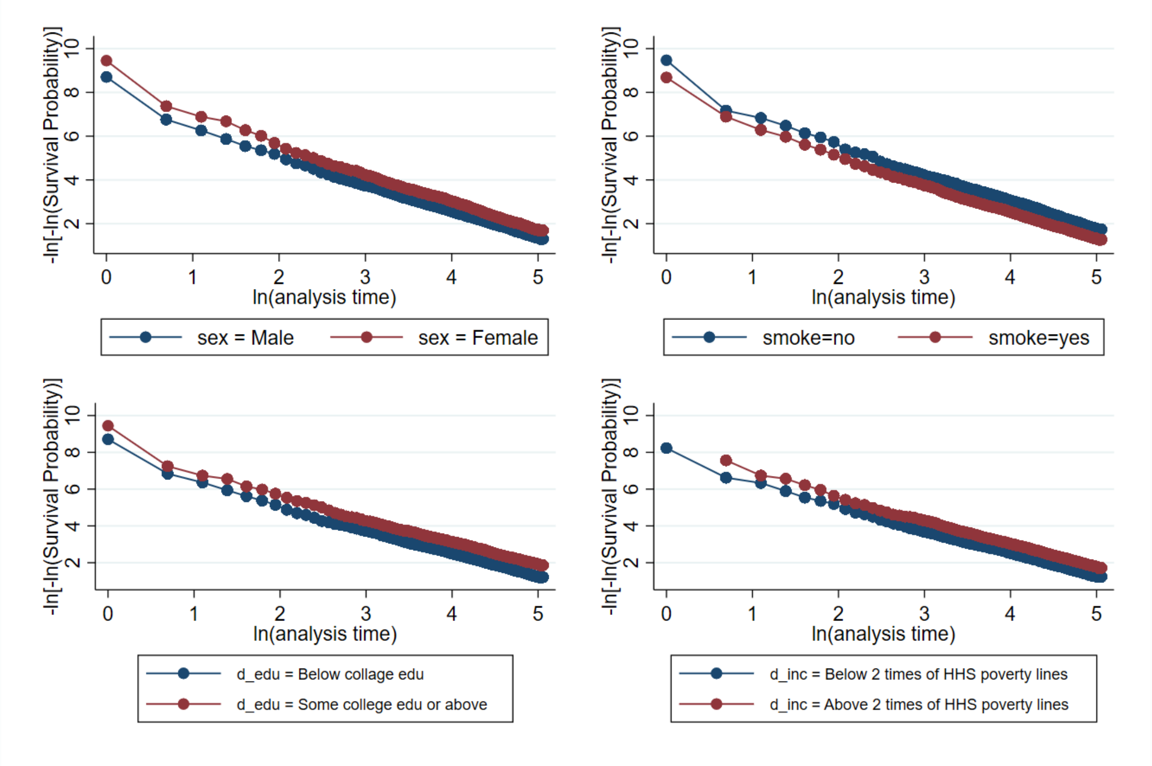


Fig A1. Log-log plots to test PH assumptions for variables of sex, smoke, education and income levels

Table A4. Schoenfeld residual test results for rescaled IDHI and covariates in Model 1 and 2

Covariates	**rho**	**chi2**	**Prob>chi2**	**rho**	**chi2**	**Prob>chi2**
		Model 1			**Model 2**	
rescaled IDHI (by 10 points)	0.005	0.05	0.82	0.000	0.00	0.98
Age (by 10 years)	0.004	0.04	0.83	0.000	0.00	0.99
Age square (by 10 years)	0.004	0.04	0.84	0.010	0.24	0.62
Female	0.012	0.35	0.56	0.013	0.42	0.52
Hispanic	0.026	1.68	0.20	0.027	1.89	0.17
Black	-0.009	0.22	0.64	-0.005	0.08	0.78
Other	-0.018	0.94	0.33	-0.019	0.99	0.32
At least 100 cigarettes				-0.007	0.11	0.74
Above 2 times of HHS poverty lines				0.027	2.10	0.15
Some college edu or above				-0.004	0.05	0.82
Global test		16.85	0.02		18.70	0.04

Table A5. Schoenfeld residual test results for rescaled AHEI and covariates in Model 1 and 2

Covariates	rho	chi2	Prob>chi2	rho	chi2	Prob>chi2
	Model 1			**Model 2**		
rescaled IDHI (by 10 points)	0.032	2.35	0.13	0.024	1.40	0.24
Age (by 10 years)	-0.002	0.01	0.94	-0.005	0.06	0.81
Age square (by 10 years)	0.010	0.23	0.63	0.014	0.47	0.49
Female	0.008	0.16	0.69	0.009	0.18	0.67
Hispanic	0.031	2.52	0.11	0.030	2.46	0.12
Black	-0.009	0.19	0.66	-0.005	0.06	0.81
Other	-0.019	1.03	0.31	-0.020	1.10	0.29
At least 100 cigarettes				-0.006	0.09	0.77
Above 2 times of HHS poverty lines				0.025	1.89	0.17
Some college edu or above				-0.007	0.15	0.70
Global test		20.84	0.00		21.67	0.02

Appendix 4. Trends of $\text{IDHI}_{\text{adverse}}$, $\text{IDHI}_{\text{beneficial}}$ and the mAHEI in subgroups in the US between 2003 and 2014

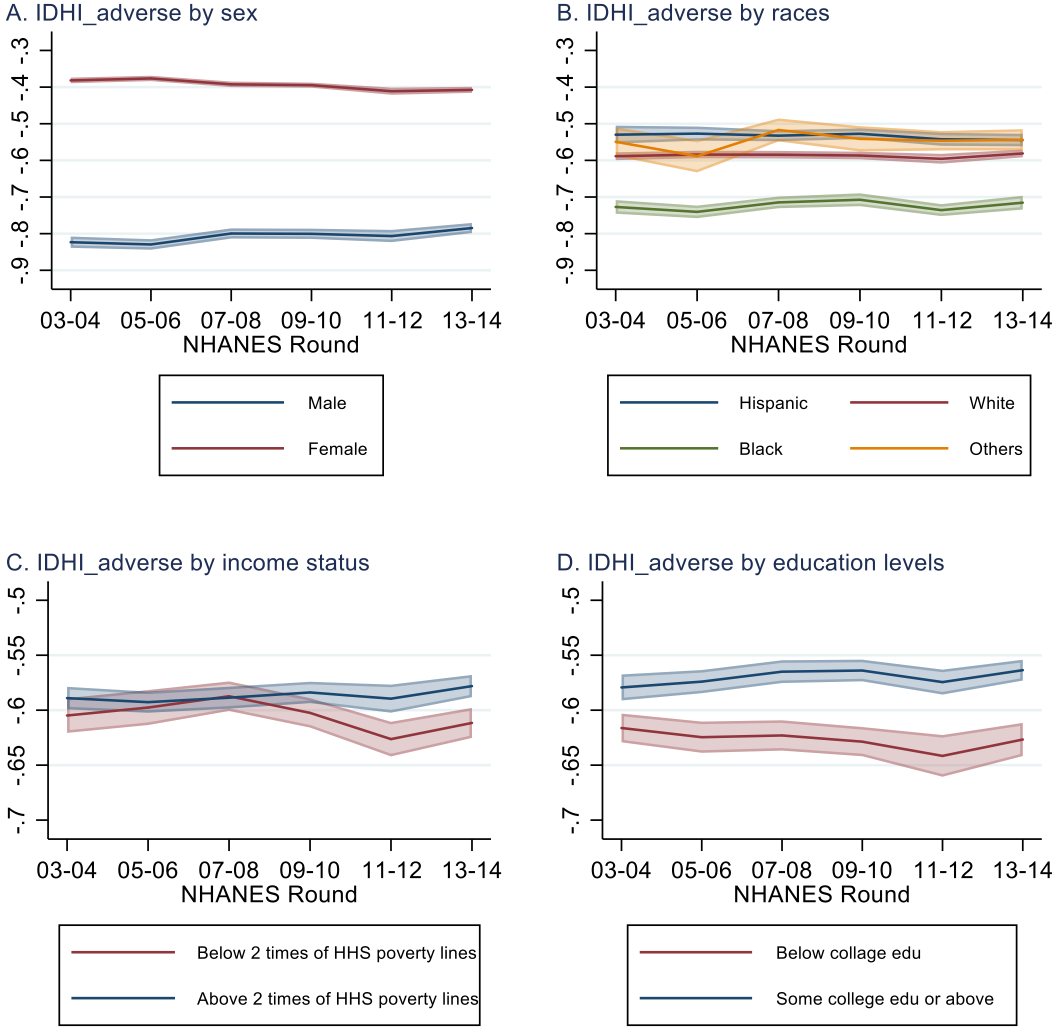


Fig A2. Trends in $\text{IHDI}_{\text{adverse}}$ across sex, race, income and education groups in the US, 2003-2014 (N=24,839)

Area of colors refer to the 95% CI of the indices. Estimates were based on general linear models controlled on survey round as a categorical variable, participant's race, sex, age, education and family income and their interaction terms with survey round. In the trend estimation we assigned mean values for all covariates. In (B), White refers to Non-Hispanic white and Black refers to Non-Hispanic black people.

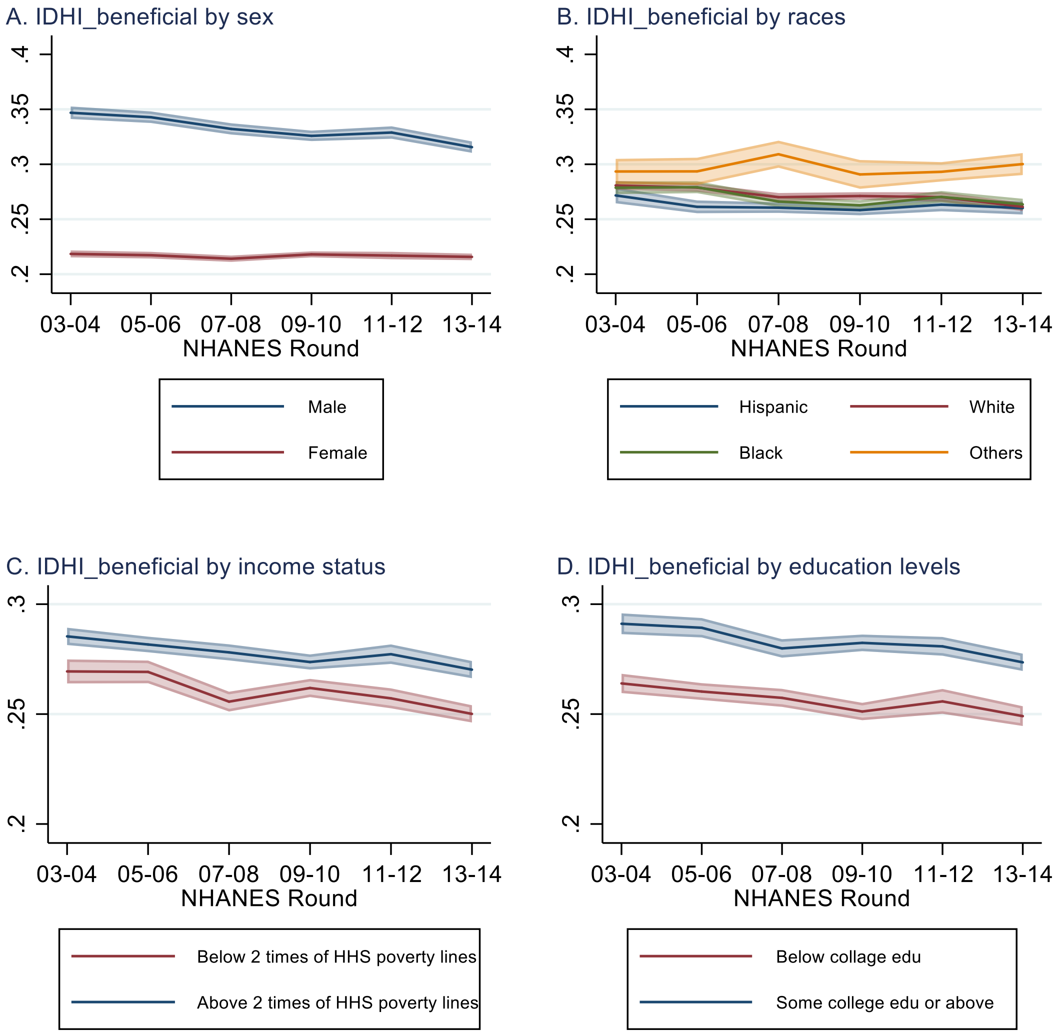


Fig A3. Predicted IHDI$_{\text{beneficial}}$ across sex, race, income and education groups in the US, 2003-2014 (N=24,839)

Area of colors refer to the 95% CI of the indices. Estimates were based on general linear models controlled on survey round as a categorical variable, participant's race, sex, age, education and family income and their interaction terms with survey round. In the trend estimation we assigned mean values for all covariates. In (B), White refers to Non-Hispanic white and Black refers to Non-Hispanic black people.

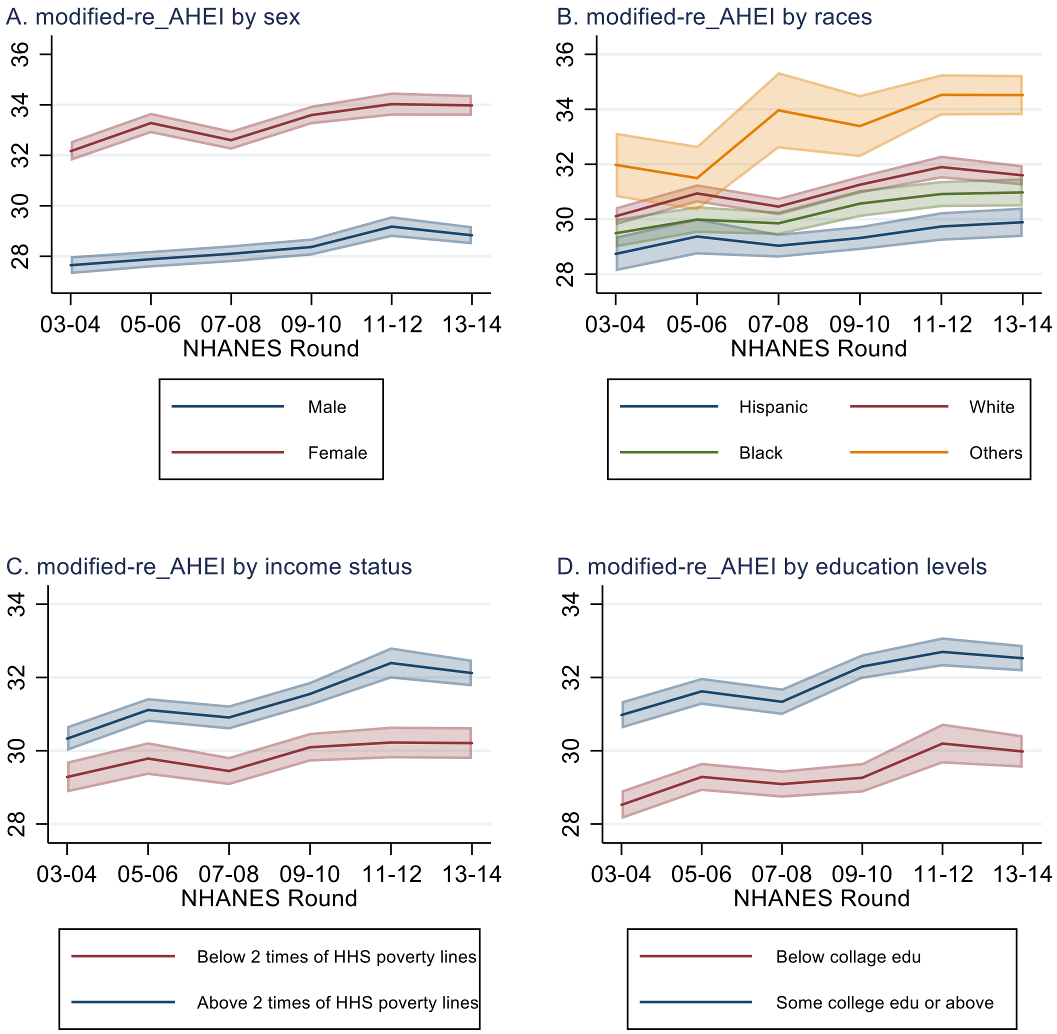


Fig A4. Predicted mAHEI across sex, race, income and education groups in the US, 2003-2014 (N=24,839)

Area of colors refer to the 95% CI of the indices. Estimates were based on general linear models controlled on survey round as a categorical variable, participant's race, sex, age, education and family income and their interaction terms with survey round. In the trend estimation we assigned mean values for all covariates. In (B), White refers to Non-Hispanic white and Black refers to Non-Hispanic black people.

Appendix 5. IDHI and mAHEI by age groups in the US population, 2003-2014

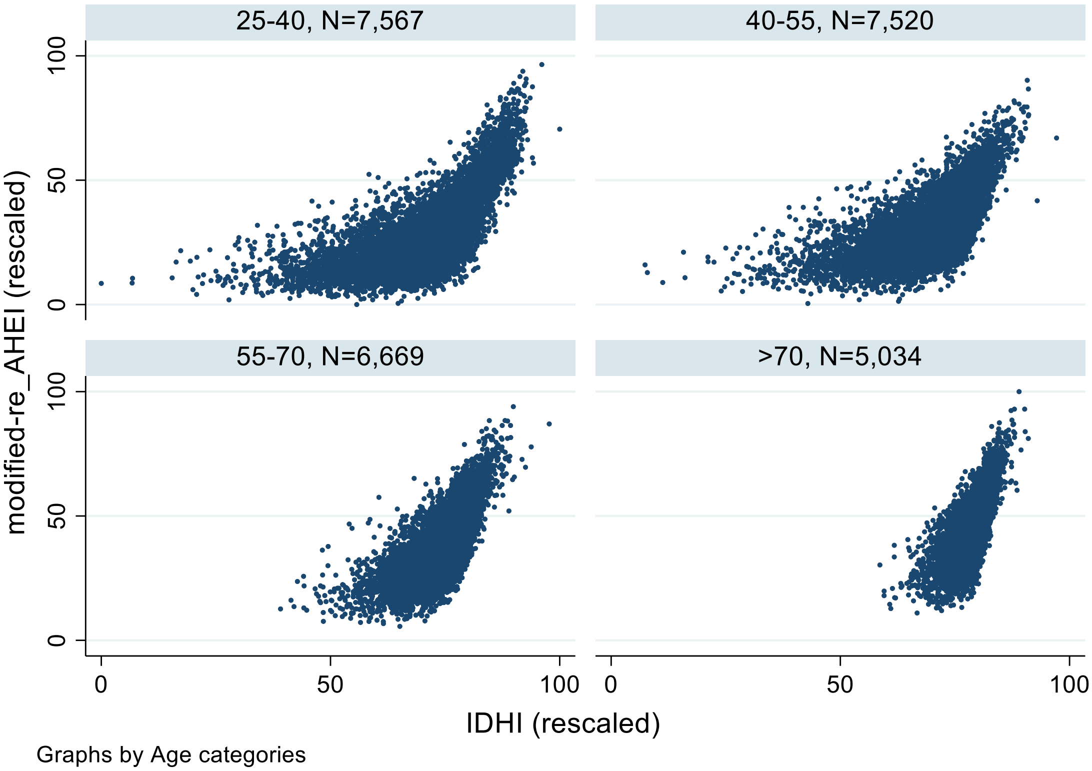


Fig A5. IDHI and mAHEI by age groups in the US population, 2003-2014
Scatter plots present the relations of IDHI and mAHEI on a rescaled basis from 0-100. Pearson's correlation coefficient in four age groups are 0.694, 0.665, 0.729 and 0.703 ($P<0.001$ for all). Spearman's rho in four age groups are 0.739, 0.727, 0.763, and 0.725 ($P<0.001$ for all).

Appendix 6. Regression results on dietary quality indices and total mortality

Table A6. Regression results of IDHI and total mortality by Cox models

Covariates	HR	Robust SE	z	P>z	[95% Conf.	Interval]
Model 1						
IDHI[1] (rescaled, by 10 points)	0.77	0.03	-7.48	0.00	0.72	0.83
Age (by 10 years)	0.95	0.11	-0.46	0.64	0.76	1.19
Age square (by 10 years)	1.08	0.01	8.42	0.00	1.06	1.10
Sex[2]						
Female	0.76	0.03	-6.31	0.00	0.70	0.83
Race[3]						
Hispanic	1.01	0.06	0.13	0.90	0.90	1.12
Black	1.11	0.06	1.86	0.06	0.99	1.24
Other	0.87	0.11	-1.18	0.24	0.68	1.10
Model 2						
IDHI[1] (rescaled, by 10 points)	0.82	0.03	-5.42	0.00	0.77	0.88
Age (by 10 years)	0.92	0.11	-0.69	0.49	0.74	1.16
Age square (by 10 years)	1.08	0.01	8.28	0.00	1.06	1.10
Sex[2]						
Female	0.78	0.04	-5.51	0.00	0.71	0.85
Race[3]						
Hispanic	0.89	0.05	-2.04	0.04	0.79	1.00
Black	1.07	0.06	1.21	0.23	0.96	1.19
Other	0.87	0.11	-1.15	0.25	0.68	1.11
Smoke[4]						
At least 100 cigarettes	1.37	0.06	7.59	0.00	1.26	1.49
Income[5]						
Above 2 times of HHS poverty lines	0.71	0.03	-7.73	0.00	0.65	0.78
Education[6]						
Some college education or above	0.83	0.04	-4.03	0.00	0.76	0.91

Note:

1. IHDI was rescaled to a range from 0-100.
2. Male is the reference group;
3. White American is the reference group;
4. Non-smoker with less than 100 cigarettes in life is the reference group;
5. Household income below 2 times of HHS poverty lines is the reference group;
6. People with no college education is the reference group.

Table A7. Regression results of IDHI and total mortality by Poisson models

Covariates	IRR	Robust SE	z	P>z	[95% Conf.	Interval]
Model 1						
IDHI[1] (rescaled, by 10 points)	0.81	0.03	-6.45	0.00	0.76	0.86
Age (by 10 years)	1.08	0.12	0.75	0.45	0.88	1.34
Age square (by 10 years)	1.06	0.01	6.63	0.00	1.04	1.08
Sex						
Female[2]	0.82	0.03	-5.42	0.00	0.76	0.88
Race						
Hispanic	1.07	0.05	1.30	0.19	0.97	1.17
Black	0.99	0.05	-0.12	0.91	0.91	1.09
Other	0.59	0.06	-5.07	0.00	0.48	0.72
Logt	0.63	0.01	-23.5	0.00	0.61	0.66
Constant	0.21	0.08	-4.30	0.00	0.10	0.43
Model 2						
IDHI[1] (rescaled, by 10 points)	0.87	0.03	-4.06	0.00	0.81	0.93
Age (by 10 years)	1.06	0.11	0.54	0.59	0.86	1.31
Age square (by 10 years)	1.06	0.01	6.47	0.00	1.04	1.07
Sex[2]						
Female	0.84	0.03	-4.74	0.00	0.78	0.90
Race[3]						
Hispanic	0.95	0.05	-0.95	0.34	0.87	1.05
Black	0.95	0.05	-1.13	0.26	0.86	1.04
Other	0.61	0.06	-4.77	0.00	0.50	0.75
Smoke[4]						
At least 100 cigarettes	1.31	0.05	7.76	0.00	1.22	1.40
Income[5]						
Above 2 times of HHS poverty lines	0.81	0.03	-5.94	0.00	0.75	0.87
Education[6]						
Some college education or above	0.74	0.03	-7.72	0.00	0.69	0.80
logt	0.63	0.01	-22.5	0.00	0.61	0.66
Constant	0.17	0.06	-4.89	0.00	0.08	0.34

Note:

1. IHDI was rescaled to a range from 0-100.
2. Male is the reference group;
3. White American is the reference group;
4. Non-smoker with less than 100 cigarettes in life is the reference group;
5. Household income below 2 times of HHS poverty lines is the reference group;
6. People with no college education is the reference group.

Table A8. Regression results of mAHEI and total mortality by Cox models

Covariates	HR	Robust SE	z	P>z	[95% Conf.	Interval]
Model 1						
AHEI[1] (rescaled, by 10 points)	0.88	0.01	-7.97	0.00	0.85	0.91
Age (by 10 years)	0.99	0.11	-0.05	0.96	0.80	1.24
Age square (by 10 years)	1.07	0.01	7.84	0.00	1.06	1.09
Sex[2]						
Female	0.73	0.03	-7.67	0.00	0.67	0.79
Race[3]						
Hispanic	0.91	0.05	-1.60	0.11	0.82	1.02
Black	1.16	0.06	2.76	0.01	1.04	1.29
Other	0.88	0.11	-1.09	0.28	0.69	1.11
Model 2						
AHEI[1] (rescaled, by 10 points)	0.91	0.02	-5.49	0.00	0.88	0.94
Age (by 10 years)	0.95	0.11	-0.41	0.68	0.77	1.19
Age square (by 10 years)	1.07	0.01	7.87	0.00	1.06	1.09
Sex[2]						
Female	0.75	0.03	-6.68	0.00	0.69	0.82
Race[3]						
Hispanic	0.83	0.05	-3.15	0.00	0.74	0.93
Black	1.11	0.06	1.94	0.05	1.00	1.24
Other	0.87	0.11	-1.11	0.27	0.68	1.11
Smoke[4]						
At least 100 cigarettes	1.36	0.06	7.44	0.00	1.26	1.48
Income[5]						
Above 2 times of HHS poverty lines	0.72	0.03	-7.53	0.00	0.66	0.78
Education[6]						
Some college education or above	0.84	0.04	-3.81	0.00	0.77	0.92

Note:

1. The mAHEI was rescaled to a range from 0-100.
2. Male is the reference group;
3. White American is the reference group;
4. Non-smoker with less than 100 cigarettes in life is the reference group;
5. Household income below 2 times of HHS poverty lines is the reference group;
6. People with no college education is the reference group.

Table A9. Regression results of mAHEI and total mortality by Poisson models

Covariates	IRR	Robust SE	z	P>z	[95% Conf.	Interval]
Model1						
AHEI[1] (rescaled, by 10 points)	0.87	0.01	-9.60	0.00	0.85	0.90
Age (by 10 years)	1.17	0.12	1.49	0.14	0.95	1.44
Age square (by 10 years)	1.05	0.01	5.79	0.00	1.03	1.07
Sex						
Female[2]	0.81	0.03	-5.92	0.00	0.76	0.87
Race						
Hispanic	0.97	0.05	-0.53	0.60	0.88	1.07
Black	1.02	0.05	0.52	0.61	0.93	1.12
Other	0.61	0.06	-4.79	0.00	0.50	0.74
Logt	0.63	0.01	-23.7	0.00	0.61	0.65
Constant	0.06	0.02	-8.28	0.00	0.03	0.12
Model 2						
AHEI[1] (rescaled, by 10 points)	0.91	0.01	-6.54	0.00	0.88	0.94
Age (by 10 years)	1.12	0.12	1.06	0.29	0.91	1.38
Age square (by 10 years)	1.05	0.01	5.89	0.00	1.03	1.07
Sex[2]						
Female	0.83	0.03	-5.03	0.00	0.78	0.90
Race[3]						
Hispanic	0.90	0.05	-2.06	0.04	0.82	0.99
Black	0.97	0.05	-0.66	0.51	0.88	1.06
Other	0.62	0.06	-4.55	0.00	0.51	0.77
Smoke[4]						
At least 100 cigarettes	1.30	0.05	7.55	0.00	1.22	1.39
Income[5]						
Above 2 times of HHS poverty lines	0.82	0.03	-5.52	0.00	0.76	0.88
Education[6]						
Some college education or above	0.75	0.03	-7.25	0.00	0.70	0.81
logt	0.63	0.01	-22.7	0.00	0.61	0.66
Constant	0.07	0.02	-7.78	0.00	0.04	0.14

Note:

1. The mAHEI was rescaled to a range from 0-100.
2. Male is the reference group;
3. White American is the reference group;
4. Non-smoker with less than 100 cigarettes in life is the reference group;
5. Household income below 2 times of HHS poverty lines is the reference group;
6. People with no college education is the reference group.

SECTION VI. SUMMARY AND DISCUSSION

Our studies reveal that nutritious diets are not affordable for global poor, and the costs to access to nutritious diets vary substantially across demographic groups and over seasons. The Cost of Nutrient Adequacy (CoNA) could be applied by policymakers as a robust cost-of-living index to target populations at risk of undernutrition and to evaluate performance of food systems in terms of nutritional security over localities and time. Similarly, the International Diet-Health Index (IDHI) is proofed to be a valid novel health index based on methods borrowed from economic approaches, which offers policymakers a useful tool of assessing dietary impacts on health across populations, and tailored guidance regarding how best to reduce disparities.

However, as applying secondary data analysis methods using different models and existing data, this study also has its own limitations in the development of methodologies as well as data quality and availability. First, as we propose innovative applications of existing method, the linear programming, into new research questions, there might be reasonable doubts on the feasibility of such applications. The top challenge for CoNA would be that the diet package selected by the optimization tool is not a real human diet, and it is sensitive to the price change however in the reality people keep their dietary pattern more consistently. Nevertheless, since CoNA was developed to reflect and track the least costs of nutrient adequate diets in a global context, to make this index as standardized and comparable across borders as possible, our strategy is to make minimal subjective and local assumptions. That is, different from the least cost diets applied in the contexts of nutrition interventions, CoNA is by no means to reflect a real human diet, or an average consumption as food CPI applied by the economists but revealing the efficiency of global food systems to deliver nutrients in proportion to human's needs.

Also, the model assumptions of symmetric shape and one year-cycle of seasonality in study 2 may not fit into all food items and regions. However, different from the previous studies which are usually focused on individual foods and markets, CoNA is determined by prices of a bundle of selected foods with different quantities. Therefore, it is very likely that different shapes of seasonal signals of costs in different food items would be harmonized into symmetric seasonal waves. In addition, previous econometric analysis and robustness tests in Study 2 also revealed that harmonic model produces even more robust results than other models with certain or no seasonal shape specifications.[30]

Study 3 explores a new area, trying to combine economic index modelling technique and epidemiological evidence. The value of IDHI is not only linking health outcomes and dietary exposures, but also indexing the risks (impacts) based on a normalized standard, the disease burdens measured by the diseases' DALY contributions.[10] However, as IDHI is a completely new concept, we need to take more efforts on communicating this idea and also demonstrating its potentials by further studies. Therefore, the validation study in the US population of Study 3 is a very important step forward. To promote its global application potentials, future work would involve validating IDHI in other populations, especially in the low- and middle-income countries.

Data quality and availability may limit the analyses as well. The International Comparison Program (ICP) arguably collected the most comprehensive and robust global food price data. However, there is still a lack in prices of some key locally available nutritious items in certain countries and regions, such as dark green leafy vegetables in Sub-Saharan Africa, which may lead to overestimation of the diet costs. Also, ICP may only provide annual average prices in every 5 or 6-year cycle, which may not meet increasing data needs for more timely tracking, especially during the pandemic of COVID-19. In contrast, national data sources may have a more balanced food coverage and monthly data frequency but may only have other issues in data sampling and processing. For example, in Tanzania, the urban bias exists as the food price data were only collected in major regional city centers. In Malawi, food prices only reflect rural areas, and the evidence could be harmed by high missing rates for certain food items. Even for the price data with fewer missing provided by the governmental authorities in Tanzania and Ethiopia, price extrapolation methods are not fully disclosed. For the Study 3, due to the lack of data for key covariates in the regression models, such as physical activity level and alcoholic beverage intakes, we also expect residual confounding issues.

Nevertheless, all such data issues should neither undermine the contributions nor limit the application potentials of these two novel indices proposed, which both follow rigorous economic approaches and incorporate the recent-reviewed scientific evidence as parameters in index formulations, providing valid measurements to assess, evaluate and improve global food systems from the perspectives of human health and nutrition, which is currently facing unprecedented burdens in the midst of long term challenges in nutrition transition[39,64,128], as well as short term shocks from the global pandemic of COVID-19[129,130]. In addition, algorithms of both indices allow them to become even more economically and nutritionally precise when more high-quality and timely data are available and scientific evidence keeps evolving in the future. The last but not the least, to improve availability and quality of nutrition-sensitive data on food prices, food compositions and dietary intakes, there is greater needs nowadays for closer international collaborations among national governments, international organizations, academic research institutions and even players in private sector around the world.

SECTION VII. REFERENCES

1 Brinkman H-J, de Pee S, Sanogo I, Subran L, Bloem MW. High Food Prices and the Global Financial Crisis Have Reduced Access to Nutritious Food and Worsened Nutritional Status and Health. *J Nutr* 2010; **140**: 153S-161S.

2 Bouis HE, Eozenou P, Rahman A. Food prices, household income, and resource allocation: socioeconomic perspectives on their effects on dietary quality and nutritional status. *Food Nutr Bull* 2011; **32**: S14-23.

3 Green R, Cornelsen L, Dangour AD, *et al.* The effect of rising food prices on food consumption: systematic review with meta-regression. *BMJ* 2013; **346**. DOI:10.1136/bmj.f3703.

4 USDA-Food Access. https://www.ers.usda.gov/topics/food-choices-health/food-access/.

5 Dangour AD, Hawkesworth S, Shankar B, *et al.* Can nutrition be promoted through agriculture-led food price policies? A systematic review. *BMJ Open* 2013; **3**: e002937.

6 Duffey KJ, Gordon-Larsen P, Shikany JM, Guilkey D, Jacobs DR, Popkin BM. Food Price and Diet and Health Outcomes: 20 Years of the CARDIA Study. *Arch Intern Med* 2010; **170**: 420–6.

7 Balk BM. A Review of Index Number Theory. In: Wiley StatsRef: Statistics Reference Online. American Cancer Society, 2016: 1–24.

8 Allen RC. Absolute Poverty: When Necessity Displaces Desire. *American Economic Review* 2017; **107**: 3690–721.

9 Masters WA, Bai Y, Herforth A, *et al.* Measuring the Affordability of Nutritious Diets in Africa: Price Indexes for Diet Diversity and the Cost of Nutrient Adequacy. *Am J Agric Econ* 2018; **100**: 1285–301.

10 Wang J, Masters WA, Bai Y, Mozaffarian D, Naumova EN, Singh GM. The International Diet-Health Index: a novel tool to evaluate diet quality for cardiometabolic health across countries. *BMJ Global Health* 2020; **5**: e002120.

11 FAO Food Price Index. http://www.fao.org/worldfoodsituation/foodpricesindex/en/.

12 Rao M, Afshin A, Singh G, Mozaffarian D. Do healthier foods and diet patterns cost more than less healthy options? A systematic review and meta-analysis. *BMJ Open* 2013; **3**: e004277.

13 Herforth A, Bai Y, Venkat A, Mahrt K, Ebel A, Masters WA. Cost and affordability of healthy diets across and within countries: Background paper for The State of Food Security and Nutrition in the World 2020. FAO Agricultural Development Economics Technical Study No. 9. Rome, Italy: FAO, 2020 DOI:10.4060/cb2431en.

14 Bai Y, Alemu R, Block SA, Headey D, Masters WA. Cost and affordability of nutritious diets at retail prices: Evidence from 177 countries. *Food Policy* 2020; : 101983.

15 Stigler GJ. The Cost of Subsistence. *Journal of Farm Economics* 1945; **27**: 303.

16 O'Brien-Place PM, Tomek WG. Inflation in food prices as measured by least-cost diets. *American journal of agricultural economics 1983*. *65*(4), 781-784. doi:10.2307/1240466.

17 Håkansson A. Has it become increasingly expensive to follow a nutritious diet? Insights from a new price index for nutritious diets in Sweden 1980-2012. *Food Nutr Res* 2015; **59**: 26932.

18 Omiat G, Shively G. Charting the cost of nutritionally-adequate diets in Uganda, 2000-2011. *African Journal of Food, Agriculture, Nutrition and Development* 2017; **17**: 11571–91.

19 Chastre C, Duffield A, Kindness H, LeJeune S, Taylor A. The Minimum Cost of a Healthy Diet. Save the Children, 2007.

20 Janssen HG, Davies IG, Richardson LD, Stevenson L. Determinants of takeaway and fast food consumption: a narrative review. *Nutrition Research Reviews* 2017; : 1–19.

21 Maillot M, Vieux F, Delaere F, Lluch A, Darmon N. Dietary changes needed to reach nutritional adequacy without increasing diet cost according to income: An analysis among French adults. *PLoS ONE* 2017; **12**: e0174679.

22 USDA. 2017. https://www.cnpp.usda.gov/USDAFoodPlansCostofFood.

23 Parlesak A, Tetens I, Dejgård Jensen J, *et al.* Use of Linear Programming to Develop Cost-Minimized Nutritionally Adequate Health Promoting Food Baskets. *PLoS ONE* 2016; **11**: e0163411.

24 Gerdessen JC, de Vries JHM. Diet models with linear goal programming: impact of achievement functions. *Eur J Clin Nutr* 2015; **69**: 1272–8.

25 Akhter N, Saville N, Shrestha B, *et al.* Change in cost and affordability of a typical and nutritionally adequate diet among socio-economic groups in rural Nepal after the 2008 food price crisis. *Food Sec* 2018; **10**: 615–29.

26 Vossenaar M, Knight FA, Tumilowicz A, Hotz C, Chege P, Ferguson EL. Context-specific complementary feeding recommendations developed using Optifood could improve the diets of breast-fed infants and young children from diverse livelihood groups in northern Kenya. *Public Health Nutr* 2017; **20**: 971–83.

27 Daelmans B, Ferguson E, Lutter CK, *et al.* Designing appropriate complementary feeding recommendations: tools for programmatic action. *Matern Child Nutr* 2013; **9 Suppl 2**: 116–30.

28 National Academies of Sciences E. Global Harmonization of Methodological Approaches to Nutrient Intake Recommendations: Proceedings of a Workshop. 2018 DOI:10.17226/25023.

29 Khandker SR. Seasonality of income and poverty in Bangladesh. *Journal of Development Economics* 2012; **97**: 244–56.

30 Gilbert CL, Christiaensen L, Kaminski J. Food price seasonality in Africa: Measurement and extent. *Food Policy* 2017; **67**: 119–32.

31 Hirvonen K, Taffesse AS, Worku Hassen I. Seasonality and household diets in Ethiopia. *Public Health Nutrition* 2016; **19**: 1723–30.

32 Regolo J, Portugal-Perez A, Brenton P. Food prices, road infrastructure, and market integration in Central and Eastern Africa. The World Bank, 2014.

33 Shively G, Thapa G. Markets, Transportation Infrastructure, and Food Prices in Nepal. *Am J Agric Econ* 2017; **99**: 660–82.

34 Minten B, Kyle S. The effect of distance and road quality on food collection, marketing margins, and traders' wages: evidence from the former Zaire. *Journal of Development Economics* 1999; **60**: 467–95.

35 Gil Á, Martinez de Victoria E, Olza J. Indicators for the evaluation of diet quality. *Nutr Hosp* 2015; **31 Suppl 3**: 128–44.

36 Ruel MT. Is Dietary Diversity an Indicator of Food Security or Dietary Quality? A Review of Measurement Issues and Research Needs. *Food Nutr Bull* 2003; **24**: 231–2.

37 Arimond M, Wiesmann D, Becquey E, *et al.* Simple Food Group Diversity Indicators Predict Micronutrient Adequacy of Women's Diets in 5 Diverse, Resource-Poor Settings. *Journal of Nutrition* 2010; **140**: 2059S-2069S.

38 Timmer CP. Food policy analysis. The World Bank, 1983.

39 Drewnowski A, Popkin BM. The Nutrition Transition: New Trends in the Global Diet. *Nutrition Reviews* 1997; **55**: 31–43.

40 Du S, Mroz TA, Zhai F, Popkin BM. Rapid income growth adversely affects diet quality in China—particularly for the poor! *Social Science & Medicine* 2004; **59**: 1505–15.

41 Tapsell LC, Neale EP, Satija A, Hu FB. Foods, Nutrients, and Dietary Patterns: Interconnections and Implications for Dietary Guidelines12. *Adv Nutr* 2016; **7**: 445–54.

42 Kennedy ET, Ohls J, Carlson S, Fleming K. The Healthy Eating Index: Design and Applications. *Journal of the American Dietetic Association* 1995; **95**: 1103–8.

43 Chiuve SE, Fung TT, Rimm EB, *et al.* Alternative Dietary Indices Both Strongly Predict Risk of Chronic Disease. *J Nutr* 2012; **142**: 1009–18.

44 Fung TT, Rexrode KM, Mantzoros CS, Manson JE, Willett WC, Hu FB. Mediterranean diet and incidence of and mortality from coronary heart disease and stroke in women. *Circulation* 2009; **119**: 1093–100.

45 Waijers PMCM, Feskens EJM, Ocké MC. A critical review of predefined diet quality scores. *British Journal of Nutrition* 2007; **97**: 219–31.

46 Deptford A, Allieri T, Childs R, *et al.* Cost of the Diet: a method and software to calculate the lowest cost of meeting recommended intakes of energy and nutrients from local foods. *BMC Nutrition* 2017; **3**: 26.

47 Gerdessen JC, de Vries JHM. Diet models with linear goal programming: impact of achievement functions. *Eur J Clin Nutr* 2015; **69**: 1272–8.

48 Dietary Reference Intakes: The Essential Guide to Nutrient Requirements. Institute of Medicine, 2006 DOI:10.17226/11537.

49 Institute of Medicine. Dietary Reference Intakes: The Essential Guide to Nutrient Requirements. Washington, DC: The National Academies Press, 2006.

50 FAO, WHO and UNU Expert Consultation. Human energy requirements. Rome, Italy: FAO.

51 Rosner B. Fundamentals of Biostatistics. Cengage Learning, 2010.

52 Cox DR. Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B (Methodological)* 1972; **34**: 187–220.

53 International Comparison Program (ICP). World Bank.

54 Agren G, Gibson RS, Eklund A, *et al.* Food composition table for use in Ethiopia. Stockholm? publisher not identified, 1969.

55 Tanzania Food Composition Tables. The Nutrition Source. 2012; published online Sept 18.

56 Stadlmayr B, Charrondiere UR, Enujiugha V, *et al.* West African food composition table. Rome, Italy: FAO, 2012.

57 SR28: USDA ARS. https://www.ars.usda.gov/northeast-area/beltsville-md-bhnrc/beltsville-human-nutrition-research-center/nutrient-data-laboratory/docs/sr28-download-files/.

58 National Health and Nutrition Examination Survey. 2018; published online Oct 29. https://www.cdc.gov/nchs/nhanes/index.htm.

59 NCHS Data Linkage. 2018; published online June 28. https://www.cdc.gov/nchs/data-linkage/mortality.htm.

60 2020 Global Nutrition Report. Bristol, UK: Development Initiatives.

61 Black RE, Allen LH, Bhutta ZA, *et al.* Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet* 2008; **371**: 243–60.

62 Black RE, Victora CG, Walker SP, *et al.* Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet* 2013; **382**: 427–51.

63 Caleyachetty R, Thomas GN, Kengne AP, *et al.* The double burden of malnutrition among adolescents: analysis of data from the Global School-Based Student Health and Health Behavior in School-Aged Children surveys in 57 low- and middle-income countries. *Am J Clin Nutr* 2018; **108**: 414–24.

64 Roth GA, Abate D, Abate KH, *et al.* Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2018; **392**: 1736–88.

65 Bailey RL, Jr KPW, Black RE. The Epidemiology of Global Micronutrient Deficiencies. *ANM* 2015; **66**: 22–33.

66 Perez-Escamilla R, Bermudez O, Buccini GS, *et al.* Nutrition disparities and the global burden of malnutrition. *BMJ* 2018; **361**. DOI:10.1136/bmj.k2252.

67 Green R, Cornelsen L, Dangour AD, *et al.* The effect of rising food prices on food consumption: systematic review with meta-regression. *BMJ* 2013; **346**: f3703.

68 Herforth A, Bai Y, Venkat A, Ebel A, Masters WA. Cost and affordability of nutritious diets across and within countries. Technical Background Paper for The State of Food Security and Nutrition in the World, 2020. 2020.

69 WHO. The double burden of malnutrition: policy brief. WHO.

70 Masters WA, Bai Y, Herforth A, *et al.* Measuring the Affordability of Nutritious Diets in Africa: Price Indexes for Diet Diversity and the Cost of Nutrient Adequacy. *Am J Agric Econ* 2018; **100**: 1285–301.

71 Fill the Nutrient Gap. World Food Programme, 2020.

72 Bai Y, Alemu R, Block SA, Headey D, Masters WA. Cost and affordability of nutritious diets at retail prices: Evidence from 177 countries. *Food Policy* 2020; : 101983.

73 Hirvonen K, Bai Y, Headey D, Masters WA. Affordability of the EAT–Lancet reference diet: a global analysis. *The Lancet Global Health* 2020; **8**: e59–66.

74 World Bank Country and Lending Groups.

75 FAO/INFOODS Global Food Composition Database for Fish and Shellfish Version 1.0-uFiSh1.0. Rome: Food and Agriculture Organization of the United Nations., 2016.

76 FAO/INFOODS Food Composition Table for Western Africa. Rome: Food and Agriculture Organization of the United Nations., 2019.

77 Shaheen N, Rahim A, Mohiduzzaman M, *et al.* Food Composition Table for Bangladesh. 2013.

78 National Academies of Sciences E. Dietary Reference Intakes for Sodium and Potassium. 2019 DOI:10.17226/25353.

79 Allen LH, Carriquiry AL, Murphy SP. Perspective: Proposed Harmonized Nutrient Reference Values for Populations. *Adv Nutr* 2020. DOI:10.1093/advances/nmz096.

80 Schneider K, Herforth A. Software tools for practical application of human nutrient requirements in food-based social science research. .

81 Michel Berkelaar et al. lpSolve: Interface to 'Lp_solve' v. 5.5 to Solve Linear/Integer Programs. 2015.

82 McGill R, Tukey JW, Larsen WA. Variations of Box Plots. *The American Statistician* 1978; **32**: 12.

83 Atamanov A, Lakner C, Mahler DG, Tetteh Baah SK, Yang J. The Effect of New PPP Estimates on Global Poverty : A First Look (English). Washington, D.C.: World Bank Group, 2020.

84 Bai Y, Masters W. Heterogeneity in the Cost of Nutrient Adequacy by Age, Sex, Pregnancy or Lactation Status and Other Influences on Individual Requirements in Malawi (OR21-04-19). *Curr Dev Nutr* 2019; **3**. DOI:10.1093/cdn/nzz034.OR21-04-19.

85 Brinkman H-J, Pee S de, Sanogo I, Subran L, Bloem MW. High Food Prices and the Global Financial Crisis Have Reduced Access to Nutritious Food and Worsened Nutritional Status and Health. *J Nutr* 2010; **140**: 153S-161S.

86 Bouis HE, Eozenou P, Rahman A. Food Prices, Household Income, and Resource Allocation: Socioeconomic Perspectives on Their Effects on Dietary Quality and Nutritional Status. *Food Nutr Bull* 2011; **32**: S14–23.

87 Naumova EN, MacNeill IB. Seasonality Assessment for Biosurveillance Systems. In: Auget J-L, Balakrishnan N, Mesbah M, Molenberghs G, eds. Advances in Statistical Methods for the Health Sciences: Applications to Cancer and AIDS Studies, Genome Sequence Analysis, and Survival Analysis. Boston, MA: Birkhäuser, 2007: 437–50.

88 Stelmach-Mardas M, Kleiser C, Uzhova I, *et al.* Seasonality of food groups and total energy intake: a systematic review and meta-analysis. *European Journal of Clinical Nutrition* 2016; **70**: 700–8.

89 Kaminski J, Christiaensen L, Gilbert CL. Seasonality in local food markets and consumption: evidence from Tanzania. *Oxf Econ Pap* 2016; **68**: 736–57.

90 Becquey E, Delpeuch F, Konaté AM, *et al.* Seasonality of the dietary dimension of household food security in urban Burkina Faso. *British Journal of Nutrition* 2012; **107**: 1860–70.

91 Martin-Prevel Y, Arimond M, Allemand P, *et al.* Development of a Dichotomous Indicator for Population-Level Assessment of Dietary Diversity in Women of Reproductive Age. *Curr Dev Nutr* 2017; **1**. DOI:10.3945/cdn.117.001701.

92 Hirvonen K, Bai Y, Headey D, Masters WA. Affordability of the EAT–Lancet reference diet: a global analysis. *The Lancet Global Health* 2020; **8**: e59–66.

93 Romanik C. An Urban-Rural Focus on Food Markets in Africa. The Urban Institute, 2016.

94 Sibhatu KT, Qaim M. Rural food security, subsistence agriculture, and seasonality. *PLOS ONE* 2017; **12**: e0186406.

95 Fanzo JC. Decisive Decisions on Production Compared with Market Strategies to Improve Diets in Rural Africa. *J Nutr* 2017; **147**: 1–2.

96 Lukmanji Z, Hertzmark E, Mlingi N, Assey V, Ndossi G, Fawzi W. Tanzania Food Composition Tables. Muhimbili University of Health and Allied Sciences (MUHAS), Dar es Salaam, Tanzania and Tanzania Food and Nutrition Centre (TFNC), Dar es Salaam, Tanzania and Harvard School of Public Health (HSPH), Boston, USA, 2008.

97 Food composition table for use in Ethiopia, Part IV. 1995.

98 Medicine I of. Dietary Reference Intakes: The Essential Guide to Nutrient Requirements. 2006 DOI:10.17226/11537.

99 National Academies of Sciences, Engineering, and Medicine. 2019. Dietary Reference Intakes for Sodium and Potassium. Washington, DC: The National Academies Press. https://doi.org/10.17226/25353.

100 The World Bank: Metadata Glossary. https://databank.worldbank.org/metadataglossary/world-development-indicators/series/PA.NUS.PRVT.PP.

101 Denton FT. Adjustment of Monthly or Quarterly Series to Annual Totals: An Approach Based on Quadratic Minimization. *Journal of the American Statistical Association* 1971; **66**: 99–102.

102 Global Information and Early Warning System. http://www.fao.org/e-agriculture/news/giews-global-information-and-early-warning-system.

103 Wenger JB, Naumova EN. Seasonal Synchronization of Influenza in the United States Older Adult Population. *PLoS One* 2010; **5**. DOI:10.1371/journal.pone.0010187.

104 Afshin A, Sur PJ, Fay KA, *et al.* Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet* 2019; **393**: 1958–72.

105 US Burden of Disease Collaborators, Mokdad AH, Ballestros K, *et al.* The State of US Health, 1990-2016: Burden of Diseases, Injuries, and Risk Factors Among US States. *JAMA* 2018; **319**: 1444–72.

106 Chiuve SE, Fung TT, Rimm EB, *et al.* Alternative Dietary Indices Both Strongly Predict Risk of Chronic Disease. *J Nutr* 2012; **142**: 1009–18.

107 Trichopoulou A, Costacou T, Bamia C, Trichopoulos D. Adherence to a Mediterranean diet and survival in a Greek population. *N Engl J Med* 2003; **348**: 2599–608.

108 Micha R, Shulkin ML, Peñalvo JL, *et al.* Etiologic effects and optimal intakes of foods and nutrients for risk of cardiovascular diseases and diabetes: Systematic reviews and meta-analyses from the Nutrition and Chronic Diseases Expert Group (NutriCoDE). *PLOS ONE* 2017; **12**: e0175149.

109 Zhang FF, Cudhea F, Shan Z, *et al.* Preventable Cancer Burden Associated With Poor Diet in the United States. *JNCI Cancer Spectr* 2019; **3**. DOI:10.1093/jncics/pkz034.

110 NHANES - About the National Health and Nutrition Examination Survey. 2020; published online Jan 8. https://www.cdc.gov/nchs/nhanes/about_nhanes.htm.

111 NHANES - Measuring Guides. 2019; published online May 8. https://www.cdc.gov/nchs/nhanes/measuring_guides_dri/measuringguides.htm.

112 Tooze JA, Midthune D, Dodd KW, *et al.* A new statistical method for estimating the usual intake of episodically consumed foods with application to their distribution. *J Am Diet Assoc* 2006; **106**: 1575–87.

113 Willett W. Nutritional Epidemiology, Third Edition. Oxford, New York: Oxford University Press, 2012.

114 Singh GM, Danaei G, Farzadfar F, *et al.* The age-specific quantitative effects of metabolic risk factors on cardiovascular diseases and diabetes: a pooled analysis. *PLoS ONE* 2013; **8**: e65174.

115 GBD Results Tool | GHDx. http://ghdx.healthdata.org/gbd-results-tool.

116 Schwingshackl L, Hoffmann G. Diet Quality as Assessed by the Healthy Eating Index, the Alternate Healthy Eating Index, the Dietary Approaches to Stop Hypertension Score, and Health Outcomes: A Systematic Review and Meta-Analysis of Cohort Studies. *Journal of the Academy of Nutrition and Dietetics* 2015; **115**: 780-800.e5.

117 Akbaraly TN, Ferrie JE, Berr C, *et al.* Alternative Healthy Eating Index and mortality over 18 y of follow-up: results from the Whitehall II cohort. *Am J Clin Nutr* 2011; **94**: 247–53.

118 Al-Ibrahim AA, Jackson RT. Healthy eating index versus alternate healthy index in relation to diabetes status and health markers in U.S. adults: NHANES 2007–2010. *Nutr J* 2019; **18**. DOI:10.1186/s12937-019-0450-6.

119 NCHS Data Linkage - Mortality Data - Public-Use Files. 2020; published online May 21. https://www.cdc.gov/nchs/data-linkage/mortality-public.htm.

120 Wang J, Masters WA, Bai Y, Mozaffarian D, Naumova EN, Singh GM. The International Diet-Health Index: a novel tool to evaluate diet quality for cardiometabolic health across countries. BMJ Global Health 2020; 5: e002120.

121 Shan Z, Rehm CD, Rogers G, *et al.* Trends in Dietary Carbohydrate, Protein, and Fat Intake and Diet Quality Among US Adults, 1999-2016. *JAMA* 2019; **322**: 1178–87.

122 Dietary Guidelines for Americans, 2015-2020. U.S. Department of Health and Human Services and U.S. Department of Agriculture.

123 Wang DD, Leung CW, Li Y, *et al.* Trends in Dietary Quality Among Adults in the United States, 1999 Through 2010. *JAMA Intern Med* 2014; **174**: 1587–95.

124 Rehm CD, Peñalvo JL, Afshin A, Mozaffarian D. Dietary Intake Among US Adults, 1999-2012. *JAMA* 2016; **315**: 2542–53.

125 Zhang FF, Liu J, Rehm CD, Wilde P, Mande JR, Mozaffarian D. Trends and Disparities in Diet Quality Among US Adults by Supplemental Nutrition Assistance Program Participation Status. *JAMA Netw Open* 2018; **1**: e180237–e180237.

126 Pool LR, Ning H, Lloyd-Jones DM, Allen NB. Trends in Racial/Ethnic Disparities in Cardiovascular Health Among US Adults From 1999–2012. *J Am Heart Assoc* 2017; **6**. DOI:10.1161/JAHA.117.006027.

127 Kris-Etherton Penny M., Petersen Kristina S., Velarde Gladys, *et al.* Barriers, Opportunities, and Challenges in Addressing Disparities in Diet-Related Cardiovascular Disease in the United States. *Journal of the American Heart Association* 2020; **9**: e014433.

128 Popkin BM, Adair LS, Ng SW. Global nutrition transition and the pandemic of obesity in developing countries. *Nutrition Reviews* 2012; **70**: 3–21.

129 Laborde D, Martin W, Swinnen J, Vos R. COVID-19 risks to global food security. *Science* 2020; **369**: 500–2.

130 Headey D, Heidkamp R, Osendarp S, *et al.* Impacts of COVID-19 on childhood malnutrition and nutrition-related mortality. *The Lancet* 2020; **396**: 519–21.