Supplemental Methods

Estimation of SNAP population by strata

Estimates of the SNAP population by age-sex strata are an essential intermediate input in estimating mortality rates among SNAP participants, eligible non-participants and ineligible individuals. Briefly, we used data from the 2014 Current Population Survey (CPS) to estimate the proportion of individuals by age-sex strata who participated in SNAP in the prior 12 months, who were income-eligible (income-to-poverty ratio<1.3) but did not participate,[22-23] and those who were not eligible based on income. CPS data were used as opposed to NHIS, as 2014 NHIS data was not available at the time of project initiation and the large sample size of CPS (n=88,595) was sufficient to precisely estimate SNAP participation rates by age-sex strata. The proportions by age-sex strata were then applied to data from the 2014 American Community Survey population estimates by age-sex-strata to obtain estimates on the three SNAP groups within each strata that add to the 2014 official population estimates.

Current mortality

There are no official estimates of mortality rates due to CHD, stroke and diabetes by age-sex-SNAP strata in the United States. Therefore, we developed an alternative approach to estimating mortality rates among SNAP participants using data from the mortality-linked National Health Interview Survey (NHIS, 2000-2009), the US Death Certificates (2000-2009),[33] and the 2014 CPS and ACS (described above). The first step of this process was to estimate mortality rates

and relative risks of mortality in the mortality-linked NHIS files. We used data from 2000-2009 to create a large sample size so that mortality rates and relative risks were statistically stable. Using data from the persons-level file (~50,000 individuals per year) we identified individuals who participated in SNAP in the prior-year, who were income-eligible but did not participate and who were not income-eligible.[8b] These data were then combined with age-sex strata to create age-sex-SNAP strata (n=36 total strata).

For each stratum, we estimated the survey-weighted mortality rates due to heart disease (the following ICD-10 codes: I00-I09, I11, I13, I20-I51), stroke (I60-I69), and diabetes (E10-E14) with follow-up through December 31, 2011. For each of these causes the ratio of mortality rates comparing SNAP participants and income-eligible non-participants was then compared to the ineligible population. For each age-sex strata, we compared the mortality rates observed in NHIS to those from US death certificates data (which capture all deaths among US residents) to determine if NHIS death rates were comparable to those observed using all available data. We observed that while in most cases death rates were comparable be age-sex strata, there were some differences, particularly among younger (25-34y) and older (≥75y) women, whose mortality experience was under-estimated in NHIS data. This suggests that, while NHIS is nationally representative, there may be some bias in who participates (e.g., in these strata, sicker individuals are less likely to participate). For most groups the difference between NHIS and death certificates was no more than +/-10-15%.

Because of potential biases in NHIS and death certificate estimates of mortality rates, we developed an approach to preserve the relative difference in mortality, but correct the NHIS

mortality estimates to be representative of the mortality rates in the US at-large. Specifically, we estimated the potential bias comparing actual mortality rates to those estimated in NHIS by age-sex strata. This adjustment factor was then applied within each strata to the NHIS mortality estimates by age-sex-strata. This approach makes the assumption that potential biases in NHIS mortality rates are equally distributed by SNAP group within each strata. For example, the NHIS observed stroke mortality rate among men ≥75y was 744.5 per 100,000, but was corrected to be 731.9 per 100,000 given evidence that stroke mortality in this age-sex group was under-estimated by 1.7% in NHIS compared to the death certificate data. This data processing step was essential, as otherwise the NHIS estimates would result in a different number of deaths than actually observed in the population.

To estimate mortality rates by strata these adjusted mortality rates were applied to estimates of the SNAP population previously described (combining data from CPS and ACS). Because NHIS data does not release data on CHD-specific mortality, we adjusted the heart disease mortality rates by estimating the proportion of heart deaths by age-sex strata there were due to CHD (I20-I24). The adjustment factor was smallest for the younger strata (e.g., 0.328 for men 25-34y) and greatest for the older groups (e.g. 0.683 for men 65-74y).

Statistical analysis

For each dietary factor and price change condition *j*, the model calculated the change in the expected relative risk for mortality (M^j) separately for each stratum of American adults

according to age (25-34, 35-44, 45-54, 55-64, 65-74, 75+ y), sex, and SNAP eligibility and participation status (3 categories):

$$M^j = \int_{x=0}^{m} \mathrm{RR}^j(x)\, \mathrm{P}^j(x)\, dx,$$

where $\mathrm{P}^j(x)$ is the stratum-specific distribution of dietary intake of food x in the population at baseline (j=0) or under the policy treatment (j=1); m is the maximum intake level beyond which no further change in risk occurs; and $\mathrm{RR}^j(x)$ is the stratum- and disease-specific relative risk at intake level x.

For each stratum, we computed the potential proportion of disease-specific mortality prevented by the price change intervention (potential impact fraction, PIF), as follows.

$$PIF = \frac{M^0 - M^1}{M^0} \quad .$$

The PIF was used to estimate the resulting deaths prevented or postponed per 10,000 adults in the population and as a percentage of all deaths.

For each food group, the distribution of current dietary intake ($\mathrm{P}^j(x)$) was modeled using a gamma distribution. In contrast with a normal distribution, which is symmetric, the gamma distribution allowed us to describe the right-skewed distribution observed in the NHANES data. For baseline estimates, the mean ($\mathrm{E}[x]=\alpha/\beta$) and variance ($\mathrm{Var}[x]= \alpha/\beta^2$) were obtained from NHANES estimates, using survey sampling weights and adjusting for complex survey design. For estimates under each policy scenario, we multiplied the mean and standard deviation by the

intervention effect. This post-intervention distribution would arise, for example, under the assumption that the intervention effect is constant across individuals.

For food groups subject to the tax (SSBs and processed meats), RR(x) was assumed to increase exponentially as a function of the positive difference between x and a theoretical minimum risk exposure level:

$$RR^j(x) = \exp[\beta(x\text{-}y(x)],$$

where β is the change in log relative risk per unit of exposure and $y(x)$ is the theoretical minimum risk based on dietary recommendations. For food groups subject to the subsidy (fruit, vegetable, nuts, and whole grains), the convention was reversed, so that relative risk increased when consumption fell short of the minimum risk exposure level.

In computation, we used numerical integration to compute integrals in the PIF equation, and monte carlo simulation to quantify uncertainty. For each diet disease pair, stratum, and policy scenario, we drew randomly 1000 times from the normal distribution of the estimate of disease-specific change in the log(RR) corresponding to a one unit increase in intake, the normal distribution of the estimate of the exposure mean, and the normal distribution of the estimate of the intervention effect. Draws of mean intake that were zero or less were changed to 0.00001. Each set of random draws was used to calculate the PIFs and attributable mortality.

For SSBs, we estimated effects mediated by body mass index (BMI). Using the same procedure as Singh et al.,[31] we multiplied the log(RR) per unit increase in BMI by the estimated increase in BMI in response to one unit increase in SSBs. The direct and mediated relative risks used in this computation are reported in Table S4.