AMERICAN ASSOCIATION OF WINE ECONOMISTS

AAWE WORKING PAPER
No. 56

Economics

THE EFFECTS OF RELATIVE FOOD PRICES ON OBESITY

EVIDENCE FROM CHINA: 1991-2006

Yang Lu and Dana Goldman

February 2010

www.wine-economics.org

The Effects of Relative Food Prices on Obesity – Evidence from China: 1991-2006*

Yang Lu
RAND Corporation
1776 Main Street
Santa Monica, CA 90407
yanglu@rand.org

Dana Goldman
School of Policy, Planning and Development
University of Southern California
650 Childs Way, RGL 214
Los Angeles, CA 90089-0626
and NBER
dpgoldma@usc.edu

Abstract

This paper explores the effects of relative food prices on body weight and body fat over time in China. We study a cohort of 15,000 adults from over 200 communities in China, using the longitudinal China Health and Nutrition Survey (1991-2006). While we find that decreases in the price of energy-dense foods have consistently led to elevated body fat, this price effect does not always hold for body weight. These findings suggest that changes in food consumption patterns induced by varying food prices can increase percentage body fat to risky levels even without substantial weight gain. In addition, food prices and subsidies could be used to encourage healthier food consumption patterns and to curb obesity.

* We are especially grateful to Michael Grossman for insightful comments and suggestions. We thank James P. Smith, Darius Lakdawalla, Erik Meijer, Homero Martinez, Jeffrey Wasserman, participants in the 2008 Academy Health and 2008 APPAM annual conferences for helpful inputs. This research is funded by the RAND Center for Chinese Aging Studies.

1 Background

Obesity is a condition under which the accumulation of body fat is high enough to adversely affect health (World Health Organization, 2000). It is a risk factor for an array of non-communicable diseases, including cardiovascular disease (CVD), hypertension, diabetes, and certain types of cancer (World Health Organization, 2000).

In industrialized countries, obesity has already been a serious public health issue. In the United States, two thirds of the adult population are overweight, and half of them are obese[1] in 2004 (Ogden et al., 2006). In fact, obesity and being overweight has become the second leading cause of preventable death in the U. S., accounting for 400,000 deaths in 2000 (Mokdad et al., 2004). Meanwhile in developing countries, an emerging nutrition transition is leading a rapid spread of overweight and obese individuals. China is no exception: between 1992 and 2002, the fraction of Chinese who are either overweight or obese had increased from 14.6% to 21.8% – a growth of 49% (Wang et al., 2007). What is even more worrisome is that, even without a BMI greater than 25 or 30, individuals can still be at risk for CVD and diabetes, if they have a high level of body fat – a situation that applies to a considerable portion of the Asian populations (Deurenberg-Yap et al., 2002; He et al., 2001; World Health Organization, 2004). Thus, the magnitude of obesity as a public health threat might be underestimated if the BMI cut-off points of 25 and 30 are the only standard applied to identify overweight and obese individuals.

This paper models the two main culprits of obesity, energy-dense diets and physical inactivity, with the main focus on the former, and measures the magnitude of their effects on obesity. Specifically, it explores whether variations in food prices can have any effects on obesity by influencing individuals' diet patterns. Given that humans have innate preferences for fat (Drewnowski et al., 1991), we hypothesize that when the price of cooking oil decreases relative to other foods, individuals will consume more of it. And because cooking oil is the most energy-

[1] Defined here by the WHO body mass index (BMI) cut-off points. Overweight is defined as BMI >= 25 and obesity is defined as BMI >= 30.

dense[2] of all foods, such a food substitution can result in a much higher level of calorie intake, which subsequently leads to more body fat when excess calories are stored in the body. Findings in current literature support this hypothesis with empirical evidence. Lakdawalla and Philipson (2002) studied US individuals from 1976 to 1994 and showed that falling food prices (as compared to other commodities) due to technological innovation accounted for 40% of the growth in BMI within that time period. And Chou *et al.* (2004) found that reductions in the real and relative prices of fast foods (which are typically high-fat) and availability of fast-food restaurants had a significant impact on obesity. However, there is little research on comparing effects within food categories (such as oils, meats, grains, and vegetables) and there is a lack of literature on effects of food prices on obesity in developing countries including China, where the growth of obesity prevalence is new and fast-paced.

2 Theoretic framework: how relative prices affect diet choice

2.1 The optimal case

Consider an individual who tries to maximize her utility from food consumption. Suppose she values the amount and taste of food, so high-fat food adds to the joy of eating, and the more calories the better. However, she is also concerned with her physical image and health status. Let's further assume that she is informed that both excess calorie intake and a high percentage of calories from fat will increase her body fat and subsequently the likelihood of developing certain diseases such as hypertension and diabetes. In addition, she understands that although she might damage her physical image by gaining weight from consuming too many calories, she might not gain any noticeable weight by consuming a high percentage of calories from fat, if she can limit

[2] The energy content of fat is ~9kcal/g, compared to ~4kcal/g for lean meat and grains, and ~0.25kcal/g for green vegetables. Source: The National Nutrition Laboratory of USDA.

the total sum of her calorie intake to a sustaining amount, say 1,800 calories/day. If she is informed of all the above, she solves

$$\max_{c,f} U\{V(c), Y(f), W(c - \underline{c} - a), D[W(c - \underline{c} - a), F(f)]\} \tag{1.1}$$

s.t.

$$I = P_f c^f + P_{nf} c^{nf} \tag{1.2}$$

$$c = c^f + c^{nf} \tag{1.3}$$

$$f = \frac{c^f}{c} \tag{1.4}$$

$$c \geq \underline{c} \tag{1.5}$$

In this model, V is the utility from consuming calories, in which c denotes the individual's calorie intake. We assume the more calories she consumes, the happier she is, so $V'(\cdot) > 0$. Here, macro-nutrients such as carbohydrates, protein, and fat all contribute to calorie intake.

f denotes percentage calories from fat. Given humans' innate preferences for fat, we assume the higher f is, the happier the individual will be. So $Y'(\cdot) > 0$, where Y is the utility from consuming fat.

W is body weight and it is determined by the amount of excess calories the individual has, and how physically active she is. The assumption is that a high body weight is undesirable and the individual would aim at maintaining a normal weight by either limiting total calorie intake or exercising adequately. $\underline{c}$ denotes the minimum number of calories one needs to sustain life, so $c - \underline{c}$ is the difference between total calorie intake and minimum amount required. a denotes additional calories expended due to physical activity, so $c - \underline{c} - a$ is the amount of excess calorie intake. According to our assumption, $W_c' > 0$, and $W_a' < 0$.

D denotes obesity-induced diseases. It's a function of both weight and body fat. Body fat, F, is a function of f, so we have $D_W' > 0$ and $D_F' > 0$.

According to the assumptions we also have $\frac{\partial U(\cdot)}{\partial V} > 0$, $\frac{\partial U(\cdot)}{\partial Y} > 0$, and $\frac{\partial U(\cdot)}{\partial W} < 0$. Further, if $D(\cdot)$ is addictively separable in W and F, we also have $\frac{\partial U(\cdot)}{\partial D} < 0$.

In the constraints, I is the individual's total disposable income for food. Consumption of other goods is not modeled here, given the assumption that utilities from consuming food and consuming other goods are additively separable, and we focus only on utilities from food consumption in this model. P_f is the relative price of fat per calorie. We set the per-calorie price of non-fat food (P_{nf}) to be 1. We assume that fat is a normal good so when price is lower the individual will consume more. Because $c^f = cf$, and $c^{nf} = (1 - f)c$, we can rewrite the budget constraint as $I = P_f f c + (1 - f)c$. Then we have $c = \frac{I}{(P_f - 1)f + 1}$, and c is then a function of f:

$c = \gamma(f)$. Substitute this $\gamma(f)$ into the utility function, we have

$$\max_{c,f} U[V(\gamma(f) - \underline{c}), Y(f), W(\gamma(f)), D(W(\gamma(f)), F(f))] \tag{1.6}$$

FOC w.r.t f:

$$\frac{\partial U(\cdot)}{\partial f} = \frac{\partial V(\cdot)}{\partial f} + \frac{\partial Y(\cdot)}{\partial f} - \frac{\partial W(\cdot)}{\partial f} - \frac{\partial D(\cdot)}{\partial f} = 0 \tag{1.7}$$

Let f^* be the optimal value of f that satisfies this FOC condition. Therefore, $U(f^*)$ is the

optimal level of utility for the individual achieved by consuming f^* percent of calories from fat.

And $\forall f \neq f^*; U(f) < U(f^*)$, given there are no multiple equilibria.

2.2 The suboptimal case

But, what would happen if the individual were to only observe weight, but not her body fat? This is quite common because it is much easier to weigh oneself than to get a direct measure of body fat. Thus, it is possible that the individual would not achieve the optimal level of utility, if she did not either observe or care about her body fat. In this case, she is ignorant about the potential harms a high body fat might cause her, and she solves her utility function by assuming $D = \tilde{D}(W(c))$ instead of $D = D(W(c), F(f))$, and she would choose $\hat{f}$ to satisfy the FOC

condition as follows:

$$\frac{\partial U(\cdot)}{\partial f} = \frac{\partial V(\cdot)}{\partial f} + \frac{\partial Y(\cdot)}{\partial f} - \frac{\partial W(\cdot)}{\partial f} - [D'_W(\cdot)\frac{\partial W(\cdot)}{\partial c}\frac{\partial \gamma(\cdot)}{\partial F}] = 0 \quad (1.8)$$

The difference between equation (1.8) and (1.7) is that, here, she sets $D'_F(\cdot)\frac{\partial F(\cdot)}{\partial f} = 0$. In the optimal case, $D'_F(\cdot)\frac{\partial F(\cdot)}{\partial f}$ is likely to be non-zero and positive when $f = f^*$, due to the concave shape of D_f. And also because $D'_f > 0$ and $D''_f < 0$, the $\hat{f}$ the individual chooses in the suboptimal case is likely to be higher than f^*. Subsequently, the disutility $\tilde{D}$ will be greater than D^*, and $\tilde{U} < U^*$. The individual will then suffer from a suboptimal level of utility by failing to link a high-fat diet with obesity-induced diseases.

3 Data

We employ the China Health and Nutrition Survey (CHNS), which is an on-going panel of individuals from over 200 communities in 9 provinces.[3] CHNS collects longitudinal data on demographics, anthropometric measurements, health indicators, and community-level commodity prices. It began in 1989 and followed the participants subsequently in 1991, 1993, 1997, 2000, 2004, and 2006. we use data from all the waves except 1989 because most food prices in 1989 were determined by the central government rather than market-driven and only a limited sample of individuals were interviewed for health and nutrition information in that year.

[3] The nine participating provinces include Liaoning, Heilongjiang, and Shandong in the northeast, Jiangsu in the east, Henan, Hubei, and Hunan in the middle, Guizhou in the southwest, and Guangxi in the south. Heilongjiang was not surveyed in 1991 and 1993, and Liaoning was not surveyed in 1997. (Source: http://www.cpc.unc.edu/projects/china/design/sample.html).

For each wave, we limit the sample to adult individuals who are between 18 and 75 years old at the time of interview. For women, we exclude the waves when they were pregnant, to avoid irrelevant weight shocks. This results in 37,816 individual-year observations from 233 communities in total. It is an unbalanced dataset, and there are about 5,000 - 8,000 individuals in each interview year,[4] and 15,649 individuals in total.

The key outcome is obesity, or adiposity. Obesity is defined as the condition of having an abnormally high proportion of body fat (National Institutes of Health, 1998, p174). In order to better capture effects on obesity, we adopt two measures to indicate the extent of body fat: body mass index (BMI) and triceps skinfold thickness (TSF), both of which are commonly used for measuring adiposity (Must et al., 1991).

BMI is traditionally used to classify obesity as a surrogate measure for percentage body fat, mainly because it is convenient to obtain and easy to interpret. It is defined as weight in kilograms divided by the square of height in meters: BMI = weight(kg)/height squared (m^2). An individual with a BMI $\geq$ 25 is classified as overweight and is further considered obese if the BMI $\geq$ 30 (National Institutes of Health, 1998). To construct BMI from the data, we use weight values from all interview waves and the first normal, non-missing height.[5] BMI is easy to calculate and widely used. However, because it is derived from only height and weight, it does not convey information on body composition, which is defined as the ratio of lean body mass to body fat mass. Thus, its accuracy in measuring body fat varies by muscularity, age, gender (Gallagher et al., 1996), and ethnicity (Calle et al., 1999). Specifically, BMI tends to overestimate body fat in muscular subjects and to underestimate percentage body fat of the elderly and of certain ethnicity groups such as Asians (Deurenberg et al., 1998). Due to this limitation, BMI can be a weak indicator for adiposity of certain population sub-groups (Burkhauser and Cawley, 2008; Deurenberg-Yap et al., 2002; Piers et al., 2000). In fact, overweight and mild obesity indicated

[4] Numbers of observations vary by year due to attrition and addition of participating individuals and provinces.
[5] Height is measured in every wave but we found there are more abnormal values in the latest waves of 2004 and 2006. Because the height of an adult is pretty stable (in this sample the decrease in height is about 0.003cm every additional year of age), for each individual, we use the first non-missing height value that is within the normal range.

by BMI alone have shown to not be an independent risk factor for cardiovascular diseases (Romero-Corral et al., 2006).

Concerning the limitation of BMI as a proxy for obesity, we use a second measure for body fat - TSF. TSF is a vertical skinfold measured at the posterior midpoint between the acromion and the olecranon. It directly measures subcutaneous body fat, and is widely used to measure body composition in clinical studies. It is considered a more appropriate measure for obesity and body fat than BMI (Rolland-Cachera et al., 1997). Moreover, unlike BMI, TSF, or the logarithm of TSF, is consistently shown to be a linear function of body fat (Durnin and Womersley, 1974; Lean et al., 1996). For example, Lean et al. (1996) showed that percent body fat increases by 0.76 percentage points for men and 0.67 percentage points for women, when there is a 1mm increase in TSF, holding all else equal. The limitation of TSF is that it does not directly calibrate visceral fat, and thus might underestimate the total body fat mass in obese individuals with high visceral fat (Williams et al., 1992, p604). Therefore we discard the very few records where BMI is higher than 35, and limit the analysis to non-morbidly obese individuals (with a BMI of 35 or higher). In addition, we also refer to the Anthropometric Reference Data from NHANES III (1988-1994)[6] and limit TSF values to a reasonable range of 3 – 40 mm.

The five key independent variables are food prices, in relative or absolute terms. They are: (1) price of staple oil relative to staple food, (2) price of staple oil to lean pork, (3) price of staple oil to commonly consumed vegetables, (4) price of staple oil by itself, and (5) price of staple oil to government-recommended intake of protein and carbohydrate, respectively.

These price variables are mainly relative and absolute prices of cooking oil. We construct the prices by taking into consideration how food is usually prepared and consumed. Grains (staple food), meats, and vegetables are the main food sources, and they are usually prepared with a certain amount of cooking oil. Because cooking oil is the most caloric of all foods, the fraction of oil used in food preparation helps determine the sum of calories a meal contains. And because

---

[6] Source: http://www.cdc.gov/nchs/about/major/nhanes/Anthropometric%20Measures.htm.

the amount of oil to use in a dish is often discretionary, it allows variations in the portion of oil, and therefore variations in percentage calories from oil intake.

We use the five oil price variables separately, each in a different regression model. When the key regressor covers certain food categories, prices of other main food categories are also being controlled for. For example, if the price of staple oil relative to staple food is the key regressor, then the price of lean pork and the price of commonly-consumed vegetables are being controlled for. And if the price of staple oil by itself is the key regressor, then the price of staple food, the price of lean pork, and the price of vegetables are being controlled for.

In these price effects models, oil and other foods are considered complementary to each other. Our hypothesis is that consumption of foods corresponds to how expensive the foods are, and when the relative price of oil decreases, people will choose to use more oil when preparing meals, and vice versa. Subsequently, higher percentage of oil intake leads to more adiposity.

By analyzing five separate models, we can test whether the price effects are persistent and robust, and can obtain more information on how the magnitude of price effects changes when the comparison group is different. Below we describe in details how prices of each food category are coded.

Out of all edible oils where price information is available, three (soybean, rapeseed, and peanut) are identified as staple oils as they account for more than 80% of the edible oil consumption in a certain region of China (Fang and Beghin, 2002). Accordingly, we assign each participating province the price of its staple oil, depending on the region it belongs to. Namely, it is soybean oil for the northeastern provinces, including Liaoning, Heilongjiang, Shandong; rapeseed oil for the provinces in the middle and west, including Jiangsu, Henan, Hubei, Hunan, Guizhou; peanut oil for the southern province Guangxi.

The prices of staple food/grains are also region-specific, because the north and south have different food staples. We use the average price of wheat and noodle for the northern region (Liaoning, Heilongjiang, Shandong, Henan), and the average price of rice and noodle for the south (Jiangsu, Hubei, Hunan, Guangxi, Guizhou).

We use lean pork to represent meats because it is the main source of animal protein in China. For vegetables, we use the average price of all the vegetables available in the survey: cabbage, rape[7], and “other commonly eaten vegetables".

The last price variable is the ratio of the price of staple oil relative to the price of a combination of protein and carbohydrates intake. We construct this key variable by acknowledging protein and carbohydrate as the main contributors of a typical meal in terms of calories. In theory, we can either model intake with real consumption patterns or with ideal proportions. Real consumption data change frequently and are hard to obtain, so we use ideal proportions by referring to the Dietary Guidelines for Chinese Residents 2007 by the Chinese Nutrition Society, in which the recommended daily intake of meats and gains in terms of calories for adults is about 1:2 to 1:3. In the analysis we use three patterns of intake combinations, 1:3, 1:2, and 1:1, mainly to check robustness of the price effects. We will describe in more detail in the results section but the three price ratios return similar and comparable coefficients on obesity. So due to space limitation, we only show the price ratio of 1:2 in this paper.

To summarize, these five food price variables can be written as: $\frac{P_{staple\ oil}}{P_{staple\ food}}$, $\frac{P_{staple\ oil}}{P_{lean\ pork}}$, $\frac{P_{staple\ oil}}{P_{vegetables}}$, $P_{staple\ oil}$, and $\frac{P_{staple\ oil}}{\frac{2}{3}P_{staple\ food}+\frac{1}{3}P_{lean\ pork}}$.

For all food categories, we use free market prices by default, and substitute with either state store market prices (from all interviews occurred in 1991, 1993, and 1997) or large store retail prices (from interviews occurred in 2000, 2004, and 2006) wherever free market prices are missing. We

[7] A leafy green vegetable in the mustard family, similar to Canola.

refer to the national food price data published by the Chinese Department of Agriculture,[8] in order to identify and to exclude abnormal price values as a way of data quality control.

Besides the key variables, we include covariates such as gender, age, age squared, type of residence (urban or rural), education attainment, income levels, physical activity levels, and year dummies.

We construct five levels of education attainment: (1) no education, (2) elementary school, (3) middle school, (4) high school diploma or equivalent (including technological and vocational school education), and (5) bachelor's degree and beyond.

There are two types of income variables in CHNS: personal income and household income. we use household income divided by household size to calculate household income per person, because (1) there are significantly fewer missing data points in the variable of household income than in personal income; (2) food preparation and consumption often do not vary within the household level in China, and per capita household income can better capture the level of disposal income for food than personal income.

Physical activity levels are important factors to include because they directly affect obesity. There are five levels of physical activity available in CHNS, based on how strenuous an individual's job is. We re-categorize them into three levels: light, moderate, and heavy, mainly to avoid a potential lack of power due to a limited number of observations in certain activity levels. "Light" includes very light activities such as sedentary office work, and light activities such as jobs that have to be completed while standing, like those of a salesperson's or of a teacher's. "Moderate" includes occupations such as student, electrician, and metal worker. "Heavy" includes heavy and very heavy activities, covering occupations such as farmer, dancer, athlete, miner, and logger.

---

[8] This dataset collects daily prices of several hundred different food items at large free markets nation-wide, for the past 10 years. All provinces are covered, though prices are not available for all free markets at all times. Source: http://www.agri.gov.cn/vegetable/sort.asp

Summary statistics are shown in Table 1, for the whole sample, and for the female and male subsamples. As we can see, there are about equal numbers of men and women. Rural residents constitute two-thirds of the sample, and urban residents one-third. Obesity-wise, both females and males have an average of BMI around 22.5, while females have a slightly higher index. In contrast, there is a much more distinctive difference between men and women in terms of TSF/body fat. Females have a TSF of 15.91 mm, 37% higher than that of men. The difference in body fat is nonetheless expected to a certain extent, due to biological differences between men and women.

The food prices provide a basic idea on how much each food item costs (note that they are a pooled average over a time period of 15 years). Lean pork is the most expensive of all, with the mean at 7 yuan per 500g. The second most expensive is staple oil, which costs slightly more than half of the price of lean pork. Staple food, representative of flour, rice, and noodles, cost about 1.2 yuan per 500g. The least expensive is vegetables, costing about half a yuan per 500g.

In terms of physical activities, almost half of all the individuals belong to the "heavy" category, mainly because most of the interviewees are rural residents, and among them many are farmers.

Perhaps the most striking demographic is education attainment. Only about 23% of all interviewees have an education attainment higher than or equivalent to a high school diploma, and only one in twenty interviewees have a bachelor's degree or higher. Males are more educated than women. 85% of males have at least some years of education, but only two-thirds of the females do.

Below we describe the time trend of the key regressors and obesity outcomes of interest, in addition to the mean values for the whole time period presented in Table 1.

Figure 1 shows the percentage changes of three relative oil prices from 1991 to 2006, including price of oil relative to staple food, price of oil relative to lean pork, and price of oil relative to vegetables. All food prices in nominal terms are increasing every year (not shown in the figure), due to inflation and economic growth, but the rate of increase in oil prices is much lower than in the other three foods that the relative prices of oil unanimously decrease over time. If we use

1991 as the reference year, the percentage decrease in the relative prices of oil range from 24% in the price of oil relative to lean pork to 58% in the price of oil relative to vegetables in 2006. There is also a 35% decrease in the price of oil relative to staple food. The price ratio of oil to vegetables has most prominent variation of all: staple cooking oil was more than 13 times more expensive than commonly-consumed vegetables back in 1991, and that ratio has reduced to a little over 6 in 2006 (numbers not shown in the figure).

Within the same time period, the prevalence of overweight and obesity skyrocketed by the definition of BMI cutoffs: the fraction of overweight ($25 \leq BMI < 30$) individuals increased from 12.4% to 23.5%, and the fraction of obese ($BMI \geq 30$) individuals rose from 1.1% to 3.2%.

However, in continuous terms, the trend does not look as prominent, because one unit change in BMI often means a substantial change in body weight. In Figure 2, it shows the trend of BMI and TSF from 1991 to 2006. (Note: the figure is not age-adjusted but the general increasing trend holds true either way.) The left chart shows changes over time in absolute terms, in which the mean BMI started at around 21.8 in 1991 and grew to 23.2 in 2006 (a1.4-point increase), and the mean TSF increased from 11.4mm in 1991 to almost 16.7mm in 2006. The chart on the right is the percentage changes over time. BMI is 6.5% higher in 2006 than in 1991. The increase in TSF is more substantial: it was 46.7% higher in 2006 compared to the 1991 level. Could this mean the growth of adiposity is faster-paced than that of body weight among Chinese individuals 1991-2006?

Also, was there a connection between the decrease in relative oil prices and the rise in obesity between 1991 and 2006? While the general time trend suggests that relative prices of energy-dense foods are negatively associated with obesity, the correlation could well have been due to chance. So Figure 3 further looks at how changes in BMI and TSF are associated with relative oil prices at the community level. It tests whether the correlation between food prices and obesity would still remain strong, with the assumption that communities are heterogeneous enough that if “chance” had only occurred at the national level it would not have influenced on both prices and obesity at a finer level.

Figure 3 shows two scatter plots with a common x-axis – percentage change in the price of oil relative to recommended protein and carbohydrate intake (or recommended meal in short). The plot on the left depicts the relationship of price with percentage change in BMI on the y-axis. Every plus sign in the plot represents one community. The plot on the right shows the correlation between percentage change in TSF and percentage change in the relative price of oil. Similarly, every circle represents one community. The majority of the communities experienced a decrease in the relative oil price and a rise in BMI and TSF, resembling the general trend. However a few of them witnessed an increase in the relative oil price and a drop in BMI and TSF, which is the opposite of the bigger picture, suggesting heterogeneity at the community level. In both scatter plots, the relationship between the relative oil price and obesity is negative and statistically significant, but the downward slope is steeper for TSF, or body fat, than for body weight.

4 Methods

In this section, we will describe the analyses we employed to explore a causal relationship between food prices and obesity, which involve a series of food price effects regressions at the individual person level. To test whether the effects are persistent and robust, we use all the possible model specifications including pooled ordinary least squares regressions (OLS), random effects (RE) models, and fixed effects (FE) models. For all the regression models, obesity is a function of community-level relative oil prices, and individual person-level characteristics that can be time-invariant or time-varying, such as age, gender, education attainment, type of residence, physical activity levels, and year fixed effects.

"Obesity" as the outcome is measured separately by continuous variables BMI and TSF, in levels and logs, for each individual i in community j at time t. Taking the FE model as an example,

"price" is the price of oil relative to staple food, to lean pork, to vegetables, to recommended meal, or by itself, respectively. X_{ijt} is the set of individual characteristics. τ_t is the time fixed effect. u_i is the individual fixed effect, and ε_{ijt} is the error term. Notation-wise, please see Equation (1.9) with BMI being the outcome of interest.

$$BMI_{ijt} = \beta_1 price_{jt} + \beta_k X_{ijt} + \tau_t + u_i + \varepsilon_{ijt} \tag{1.9}$$

When TSF is used as the outcome of obesity, for each model specification we estimate two regressions, one with BMI being a covariate, and one without. Please see Equations (1.10) and (1.11). This is because in the data, although BMI and TSF are both measures for obesity, they are not perfect substitutes for each other. In fact, the correlation between the two is only 0.45. Therefore, by controlling for BMI when the outcome is TSF, we are able to see whether body fat would still be affected by food prices when body weight is held constant.

$$TSF_{ijt} = \beta_1 price_{jt} + \beta_k X_{ijt} + \tau_t + u_i + \varepsilon_{ijt} \tag{1.10}$$

$$TSF_{ijt} = \beta_1 price_{jt} + \beta_k X_{ijt} + \beta_2 BMI_{ijt} + \tau_t + u_i + \varepsilon_{ijt} \tag{1.11}$$

In all models we estimate the coefficients with heteroscedasticity-robust standard errors clustered at the community level. This is the strictest assumption one can make with the dataset. Bertrand *et al.* (2004) suggested that even in differences-in-differences models (FE models with time fixed effects in this case), the standard deviation of the estimated treatment effects might be underestimated in the presence of serial correlation, when conventional standard errors are used. To be most cautious in interpreting the price effects, we use heteroscedasticity-robust standard errors clustered at the community level. Compared to the conventional standard errors, these clustered robust standard errors do not change point estimates of the coefficients, but often produce much larger standard errors and therefore decrease the likelihood for a coefficient to be statistically significant. However, this is necessary for correcting heteroscedasticity and serial correlation in at the community level where variation in the key regressors – food prices occurs.

5 Results

Table 2, Table 3, Table 4, and Table 5 show the effects on BMI and TSF by the price of staple oil relative to staple food, price of staple oil relative to lean pork, price of staple oil relative to vegetables, price of staple oil, and price of staple oil relative to recommended meal, respectively. All the regressions shown are log-log models, partly because it is easier to interpret the coefficients as price elasticities. In each table, there are six regressions. They are listed side by side for easy comparisons between treatment effects on body weight and that on body fat. The first three on the left are on the outcome of BMI, and the other three are on the outcome of TSF with BMI being controlled for. Regressions with the outcome of TSF without BMI being included are analyzed but not included in the tables because (1) the price effects on body fat do not differ significantly when BMI is added as a control variable, and (2) space is limited.

For each outcome, an OLS, a RE, and a FE model is shown. Regression models in levels are also analyzed. The coefficients estimated have the same expected sign and the statistical significance is comparable to that in the log-log models. Those results are not shown due to space limitation, but are available upon request.

Across all model specifications, relative oil prices are generally shown to be negatively associated with TSF, and thus body fat, except in the FE model of oil price relative to vegetables. This confirms the hypothesis that when oil prices are lower, body fat will increase. We deduce that consumption of staple oil and energy-dense foods goes up as a response to changes in oil prices. This middle step, changes in food consumption, is not demonstrated in the reduced-form regressions, but it is reasonable to assume so because it is well documented elsewhere. Ng et al. (2008) found that increased consumption of cooking oil at the individual-person level was associated with decreasing prices – empirical evidence from the same CHNS dataset but with a shorter panel, from 1991 to 2000.

In contrast, the expected negative treatment effects on BMI only appear in some model specifications, all of which OLS regressions. Those effects vanish in the RE and FE regressions. However, the key regressor, price of oil relative to vegetables, appears to be an outlier with a positive and significant effect on body weight, which we will discuss in more detail when we describe results from Table 4.

In Table 2, there are six log-log models with one key regressor: the relative price of staple oil to staple food. In the three models on the left with BMI being the outcome, we see that the key regressor does not appear to have any treatment effects on body weight. The sign is negative in the OLS model, which is expected, but it flips in RE and FE models, and is never statistically significant. However, the effects of non-food price control variables are reasonably in accordance with common beliefs. A moderate or heavy physical activity level is associated with less body weight as compared to light. Age and age squared are strong predictors for body weight, suggesting non-linear, quadratic age effects. Income is also positively associated with body weight, except the effects fade away in the FE model. In the OLS and RE model we are also able to see the effects of individual-specific, time-invariant variables such as type of residence, and gender. Both being an urban resident and a female are positively associated with body weight. Education attainment levels are included in the regressions but the coefficients are not shown here because most times they do not appear to be statistically significant, or interesting. Year fixed effects are considerable and statistically significant in all regressions. They are not shown here either because without further efforts we are not able to isolate and specify effects of individual events happened in those years that had an influence on obesity.

Still in Table 2, the three regressions on the right are on the outcome of TSF, or body fat. There appears to be significant price effects on body fat in all three models, although the magnitude wanes when model specifications make more stringent assumptions about the data. The price coefficient is -0.15 in OLS, -0.13 in RE, and -.08 in FE. But what do price effects imply?

Suppose in one community, the price of staple food is 4 per 500 gram, and the unit price of staple oil is 8. The relative price of staple oil to staple food is 8÷4= 2. Then, if the price of staple oil increases from 8 to 9, while the price of staple food increases from 4 to 5, the price of staple oil to staple food will change to 9÷5= 1.8. And that is a (1.8-2) ÷ 2 = -10% change from the original relative price. In the OLS model, holding all else equal, a 10% decrease in the price of staple oil relative to staple food leads to a 1.5% increase in TSF. With the same scenario, the treatment effect is reduced to a 1.3% increase in TSF in the RE model, and a 0.8% increase in the FE model.

Although the treatment effects are smaller in the FE model than in the OLS and RE models, they are in general much more reliable point estimates. There are a few reasons for that. First, in order for OLS estimators to be consistent, the error term has to be i.i.d. (independent and identically distributed), or in other words, random, and uncorrelated with the control variables, which is an ideal case that often does not exist in the real world. Second, RE models a composite error term $(u_i + \varepsilon_{ijt})$ instead of identifying each individual effect like FE model does, assuming orthogonality between individual effect u_i and control variables X_{ijt}.

However, unobservable individual characteristics absorbed by u_i can be correlated with observed variables and can lead to biased point estimates. There are at least two ways to test whether a RE model produces consistent results. One is to run a FE model where u_i will be estimated, and check the correlation between u_i and Xb. If the correlation is strong, it means that the orthogonality assumption a RE model makes is invalid, and therefore the RE estimator is inconsistent. If the correlation is weak or close to zero, then both the FE and the RE estimators are consistent, while the RE model is more efficient. The other way is to run both RE and FE models and conduct a Hausman test to see whether we can reject the null hypothesis that the RE estimator is consistent. We use the first method to evaluate whether the RE estimator is appropriate, because the statistic is automatically calculated for the regression model by Stata when a FE model is being estimated. In this specific FE model with price of staple oil relative to staple food being the key regressor, the correlation between individual fixed effects and control variables is -0.9988. The absolute value is very close to 1, suggesting the orthogonality assumption is invalid, and that the FE model should be more appropriate in estimating the price effects. In fact, the correlation between u_i and control variables X_{ijt} in all the other FE models estimated is much close to 1 in absolute value, similar to the current example, so from now on, we will focus mainly on interpreting the FE estimators for all key regressors.

In the last column of Table 2 where TSF is the outcome of interest in a FE model, the price of lean pork is negatively associated with body fat, while the price of vegetables has little, if any, effects on body fat at all. BMI is found to be a strong predictor for body fat, as expected. The estimated coefficient says that when BMI increases by 10%, body fat will increase by 12.6%, holding all else equal. Because the point estimate/elasticity is greater than one, it suggests that within individuals, body fat may be growing at a faster rate than is body weight, holding all else equal. Interestingly, within the same individual, changes in physical activity levels do not seem to affect body fat significantly, holding all else equal.

Lastly, by comparing the FE estimators between the two obesity outcomes, we find that effects of staple oil price relative to staple food on obesity are manifested in TSF, a more direct measure of body fat, but are not captured by BMI, an equivalent to body weight.

Table 3 shows treatment effects of oil price relative to lean pork. The six regressions here explore to what extent individuals would substitute animal protein with cooking oil when price differences are noticeable, and how this would affect their body weight and body fat. The first three models model prices effects on the outcome of BMI. The relative oil price appears to be negatively associated with BMI, suggesting that when the relative price of oil becomes lower than that of lean pork, there would be a higher percentage of calories from oil consumption and a higher calorie intake in total, which eventually lead to a higher body weight. However, the price effect loses its statistical significance in the FE model.

In contrast, when TSF is the outcome of interest, all three regression models indicate negative and significant treatment effects: the estimation of price effects varies from -0.16 in OLS to -0.14 in RE to -0.08 in FE. If we take the FE estimation at face value, holding all else equal, a 10%

decrease in the price of staple oil relative to lean pork will result in a 0.8% increase in TSF, or body fat. The effects are comparable to that from Table 2.

Results represented in Table 4 are six regression models with the price of staple oil relative to vegetables as the key regressor. In the third model, the FE model on BMI, holding all else equal, a 10% decrease in the price of staple oil relative to vegetables results in a 0.04% decrease in BMI ($p<0.01$). Such a positive yet very small price effect is perplexing because it is contradictory to the common belief that lower oil price induces more oil consumption, and obesity. This effect is also quite different from what is presented in Table 2 and Table 3, where little or no relative oil price effects are shown on BMI. This positive effect is also contradictory to the general trend of food prices and BMI. The price of oil relative to vegetables has gone down between 1991 and 2006, as shown before, and at the same time BMI has increased, suggesting a negative correlation.

One possible explanation for this is that vegetables are a Giffen good, and when they are priced higher there will be more consumption. But there is little literature supporting this point. Alternatively, it could be price endogeneity – healthier, light-weighted individuals have a strong preference for vegetables, driving both prices and consumption up. A third possibility is that it is a type we error. None of these hypotheses can be tested here, however. But it would be easier to at least explore the first hypothesis, when prices of each food category are being controlled for individually, as presented in Table 5. we will discuss this further after we describe the TSF part of Table 5.

Results on the outcome of TSF are also different from that of earlier model specifications. The price effects are negative and significant in OLS and RE but that statistical significance is lost in the FE model, suggesting that we cannot reject the null hypothesis the price of oil relative to vegetables has no effects on body fat.

Table 5 shows results from six regressions with the key regressor being the absolute price of staple oil. Prices of staple food, lean pork, and vegetables are also being controlled for individually.

When BMI is the outcome of interest, in the FE model, it appears that the price of staple oil has no treatment effects on BMI. Actually, the only statistically significant price effect ($p<0.05$) here is from lean pork. When the price of staple oil and everything else is held constant, a higher price of lean pork predicts a higher body weight. The effect is small (point estimate 0.006), but significant. The price of vegetables is negative but only border-line significant ($p<0.1$), suggesting little to no effects on BMI, when other food prices are being held constant.

The results suggesting little price effects of vegetables are more reasonable than assuming negative effects or in other words a Giffen-good phenomenon. One good reason is that vegetables contain very few calories. In fact, there is only 0.25 calories per gram of vegetables on average. So even if the consumption of vegetables doubled within the same time period, it wouldn't add enough calories to significant increase calorie intake, or increase body weight, not to mention the fact that many fiber-rich vegetables are filling, so by consuming more of them people might eat less in general.

Also there is empirical evidence that consumption of vegetables do not increase when their prices go up. Aggregate-level data provided in the 2007 Agricultural Development Report by the Ministry of Agriculture in China show that annual consumption of vegetables per capita has been decreasing over time: from 127kg in 1991 to 101kg in 2006 for urban residents, and from 132kg in 1991 to 118kg in 2006 for rural residents.[9] Within the same time frame, prices of vegetables have increased. Connecting the two, we exclude the possibility that vegetables are a Giffen good. Price endogeneity or type we error are more plausible candidates.

Still in the same table, when TSF is the outcome of interest, prices effects are all statistically significant with an expected sign. The point estimate for oil price effect is -0.09 in the FE model ($p<0.01$), the magnitude quite comparable to that of the relate oil prices compared to staple food, or to lean pork. The price effects of staple food and lean pork on TSF are positive, when the pricing of oil is being held constant, which is in accordance with my hypothesis. The price of vegetables does not appear to have any treatment effects on TSF, or body fat.

---

[9] Source: http://www.agri.gov.cn/sjzl/baipsh/WB2007.htm#24.

Table 6 is the last table displaying price effects regressions for the whole sample. The key regressor shown here is the price of oil relative to recommended meal. Again, in the FE models, the relative oil price appears to have no effects on body weight, and is shown to have a negative and significant ($p<0.01$) effect on body fat (TSF), even when BMI is being controlled for. The magnitude of the effect is even greater than earlier models: a 10% decrease in the price of oil relative to recommended meal will increase TSF by 1.2%, holding all else equal.

By look at results from all model specifications, it suggests that effects of both relative and absolute oil prices on TSF, or body fat, are persistent and robust. However, BMI does not seem to capture these price effects. The only exception to the generalization is the price of staple oil relative to vegetables.

This implies that individuals consume more edible oils when the prices are lower, and additional oil consumption subsequently leads to a higher level of body fat. At the same time, individuals are likely to keep total calories consumed per day to a certain number, and reduce the consumption of other foods, such as rice, flour, meats, and vegetables, to compensate for more oil intake, such that the total calorie intake will not change substantially, but a higher fraction of calories will come from fat.

5.1 Extended Results by Gender

Although FE models are shown to be superior to RE in this paper (because observable individual characteristic are highly correlated with the control variables), the limitation of FE models remains – FE models do not estimate effects of individual characteristics that are invariant over time because those effects are modeled as part of individual fixed effects by definition. One way to overcome that limitation is to analyze fully-interacted models, or in other words, to run regressions using subsamples divided by a variable of interest. We employ this technique and estimate price effects within subsamples by gender to see if and how treatment effects differ between females and males.

Table 7 shows the subsample results (log-log models only). The first column of results shows the regression on the outcome of BMI, and the second is on TSF. Each block of a coefficient and a standard error represents one regression. There are altogether 20 regressions in the table. They are all FE models with robust standard errors clustered at the community level.

When BMI is the outcome, all the price effects are not statistically significant except the price of oil relative to vegetables in the male subsample. When TSF is the outcome, mostly the price effects are comparable between females and males, except for the price of staple oil by itself. In the male subsample, the absolute oil price effect is not statistically significant, but in the female subsample, the p-value is less than 0.01.

6 Endogeneity

Price effects shown earlier cannot be considered causal if there is price endogeneity. In a free and competitive market, prices are co-determined by supply and demand. Self selection from the demand side and reverse causality are two legitimate concerns regarding the estimated price effects from earlier regression models. Therefore we use instrumental variables (IV) to test if price endogeneity is present and to further estimate causal price effects on obesity.

An instrument is considered valid and relevant in this particular case if (1) the instrument is associated with food prices from the supply side, such as a production cost, or a transportation cost, and (2) the instrument itself does not in any other way affect body weight or body fat except through food prices. The price of gasoline is available from the CHNS community survey and meets the two above requirements. First, gasoline is directly used for vehicles that transport and distribute food, so it is part of the transportation cost of food, from the supply side. The relationship between the price of gasoline and the price of food prices can be tested in the first stage of IV. Second, gasoline is unlikely to be directly linked with body weight and body fat. This condition might be violated in developed countries where car ownership is so common that when the price of gasoline is lower, individuals might choose to substitute walking with driving, and store more calories and fat in the body. But in China, very few people have a car at their disposal, and most people commute by foot, bike, or bus. So, changes in gasoline prices supposedly would not change individuals' physical activity levels, or other behaviors related to obesity.

For all the IV estimates, only FE-IV models are shown in Table 8 and Table 9 due to space limitation. There are altogether ten shown regressions, five on the outcome of BMI, five on TSF.

The first stage of these IV regressions is not shown, but for each regression we run a weak identification test and an under-identification test on the validity of the instrument. The weak identification test uses Kleibergen-Paap rk Wald F statistic and the under-identification test uses Kleibergen-Paap rk LM statistic (Baum et al., 2007). The statistics are shown near the bottom of the tables. In all models these statistics are large enough to reject the null hypothesis that instrument is not valid. WE also perform endogeneity tests of the price variables in OLS-IV models, which suggest that the price of staple oil is endogenous.

We find that in the FE-IV models, prices effects remain all negative and statistically significant on TSF, echoing what we find from earlier FE models. Even the price of oil relative to vegetables is shown to influence TSF. And the magnitude of effects is now larger. For example, the effect of staple oil price relative to staple food has risen from -0.08 in FE to -0.214 in the FE-IV model. The coefficient of -0.214 implies if the price of oil relative to staple food decreases by 10%, TSF will increase by about 2.1%, holding all else equal. However, does this instrumented point estimate sound too large to be reasonable? From 1991 to 2006, the price of oil relative to staple food decreased by 35%. Applying the instrumented point estimate, this price change led to a $35\% \times 0.214 \approx 7.5\%$ increase in TSF. Within the same time period, TSF increased by 46.7%. So, the price of staple oil relative to staple food explains $7.5\% \div 46.7\% = 16\%$ of the total increase in TSF. Given that excess food intake and physical inactivity are the two main contributors of high body fat, a point estimate indicating that food price accounts for 16% of the changes in TSF hardly sounds absurd.

On the other hand, the price effects on BMI are all shown to be negative and statistically significant as well, which is aligned with the findings on the outcome of TSF, but different from what we find in the FE models on BMI earlier. In the same example of the price of staple oil relative to staple food, the earlier FE model gives us a coefficient of -0.005 which is negative yet insignificant, but once instrumented, the coefficient is changed to -0.039 and is statistically significant ($p < 0.01$). It says, holding all else equal, if the price of staple oil relative to staple food increases by 10%, an average person's BMI will decrease by about 0.4%. If we take the results at face value, it implies that when relative oil prices are lower, individuals are likely to gain body weight and increase percentage body fat. However, price effects are much larger on body fat than body weight, suggesting body fat might have been growing at a faster rate than body weight.

We do not have enough information to conclude whether regression results from one model are definitively more reliable than those from another, but at least the IV method reconfirms the hypothesis that relative oil prices have negative treatment effects on body weight and body fat. Oil prices are demonstrated to have a greater impact on body fat than on body weight. And the results are consistently shown to be robust.

7 Conclusion

Results from the reduced-form price effects models suggest that food prices affect individuals' body fat levels, probably by changing the composition of meals they consume – when the relative price of staple oil is lower, obesity increases, captured by changes in TSF. In addition, by using any constructed relative oil price, across all model specifications, price effects on body fat are significant and persistent. However, such price effects appear to be more subtle on BMI and cannot always be captured, especially in the FE models (though reflected in FE-IV models). This is partly because BMI is not a perfect measure of body fat. It is by definition the height-adjusted body weight and does not convey information on body composition such as percentage body fat and muscularity.

To summarize, all regressions results show that oil consumption can correspond to relative oil prices and it is shown to subsequently increase individuals' body fat at a faster rate than it affects body weight. The policy implications are two-fold. First, by using different pricing policies, it is possible to effectively induce healthier food consumption patterns and thus to control the growth of obesity. Second, BMI contains limited information on body fat and is a less accurate measure for obesity especially when it applies to certain populations, such as Asians.[10] It is important to raise public awareness that individuals need to be cautious when interpreting their degree of adiposity by using BMI alone, and that it goes beyond maintaining a normal body weight to ensure a healthy level of percentage body fat. If more direct measures of body fat such as TSF are too costly or simply unavailable, it would still help to watch and control daily oil consumption levels.

[10] The WHO and NIH definition of obesity indicated by a BMI between 25 and 30 is derived by measuring mostly young white males only.

References

Baum CF, Schaffer ME, Stillman S. Enhanced routines for instrumental variables/generalized method of moments estimation and testing. Stata Journal 2007;7; 465-506.

Bertrand M, Duflo E, Mullainathan S. How Much Should We Trust Differences-in-Differences Estimates? Quarterly Journal of Economics 2004;119; 249-275.

Burkhauser RV, Cawley J. Beyond BMI: The value of more accurate measures of fatness and obesity in social science research. Journal of Health Economics 2008;27; 519-529.

Calle EE, Thun MJ, Petrelli JM, Rodriguez C, Heath CW. Body-Mass Index and Mortality in a Prospective Cohort of U.S. Adults. New England Journal of Medicine 1999;341; 1097-1105.

Chou S-Y, Grossman M, Saffer H. An economic analysis of adult obesity: results from the Behavioral Risk Factor Surveillance System. Journal of Health Economics 2004;23; 565-587.

Deurenberg-Yap M, Chew SK, Deurenberg P. Elevated body fat percentage and cardiovascular risks at low body mass index levels among Singaporean Chinese, Malays and Indians. Obesity Reviews 2002;3; 209-215.

Deurenberg P, Yap M, Staveren WAv. Body mass index and percent body fat: a meta analysis among different ethnic groups. International Journal of Obesity 1998;22; 1164-1171.

Drewnowski A, Kurth CL, Rahaim JE. Taste preferences in human obesity: environmental and familial factors. American Journal of Clinical Nutrition 1991;54; 635.

Durnin J, Womersley J. Body fat assessed from total body density and its estimation from skinfold thickness: measurements on 481 men and women aged from 16 to 72 Years. British Journal of Nutrition 1974;32; 77-97.

Fang C, Beghin JC. Urban Demand for Edible Oils and Fats in China: Evidence from Household Survey Data. Journal of Comparative Economics 2002;30; 732-753.

Gallagher D, Visser M, Sepulveda D, Pierson RN, Harris T, Heymsfield SB. How Useful Is Body Mass Index for Comparison of Body Fatness across Age, Sex, and Ethnic Groups? American Journal of Epidemiology 1996;143; 228-239.

He M, Tan KCB, Li ETS, Kung AWC. Body fat determination by dual energy X-ray absorptiometry and its relation to body mass index and waist circumference in Hong Kong Chinese. International Journal of Obesity 2001;25; 748-752.

Lakdawalla D, Philipson T. The Growth of Obesity and Technological Change: A Theoretical and Empirical Examination. National Bureau of Economic Research Working Paper Series 2002;No. 8946.

Lean ME, Han TS, Deurenberg P. Predicting body composition by densitometry from simple anthropometric measurements. American Journal of Clinical Nutrition 1996;63; 4-14.

Mokdad AH, Marks JS, Stroup DF, Gerberding JL. Actual Causes of Death in the United States, 2000. Journal of the American Medical Association 2004;291; 1238-1245.

Must A, Dallal GE, Dietz WH. Reference data for obesity: 85th and 95th percentiles of body mass index (wt/ht2) and triceps skinfold thickness. American Journal of Clinical Nutrition 1991;53; 839-846.

National Institutes of Health. Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults: The Evidence Report NIH Publication 1998;No. 98-4083.

Ng SW, Zhai F, Popkin BM. Impacts of China's edible oil pricing policy on nutrition. Social Science & Medicine 2008;66; 414-426.

Ogden CL, Carroll MD, Curtin LR, McDowell MA, Tabak CJ, Flegal KM. Prevalence of Overweight and Obesity in the United States, 1999-2004. Journal of the American Medical Association 2006;295; 1549-1555.

Piers L, Soares M, Frandsen S, O'Dea K. Indirect estimates of body composition are useful for groups but unreliable in individuals. International Journal of Obesity 2000;24; 1145-1152.

Rolland-Cachera MF, Brambilla P, Manzoni P, Akrout M, Sironi S, Del Maschio A, Chiumello G. Body composition assessed on the basis of arm circumference and triceps skinfold thickness: a new index validated in children by magnetic resonance imaging. American Journal of Clinical Nutrition 1997;65; 1709-1713.

Romero-Corral A, Montori VM, Somers VK, Korinek J, Thomas RJ, Allison TG, Mookadam F, Lopez-Jimenez F. Association of bodyweight with total mortality and with cardiovascular events in coronary artery disease: a systematic review of cohort studies. The Lancet 2006;368; 666-678.

Wang Y, Mi J, Shan XY, Wang QJ, Ge Ky. Is China facing an obesity epidemic and the consequences? The trends in obesity and chronic disease in China. International Journal of Obesity 2007;31; 177-188.

Williams DP, Going SB, Lohman TG, Hewitt MJ, Haber AE. Estimation of body fat from skinfold thicknesses in middle-aged and older men and women: a multiple component approach. American Journal of Human Biology 1992;4; 595-605.

World Health Organization. Obesity: preventing and managing the global epidemic. WHO Technical Report Series 2000;894.

World Health Organization. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. The Lancet 2004;363; 157-163.

Figure 1—Price trends

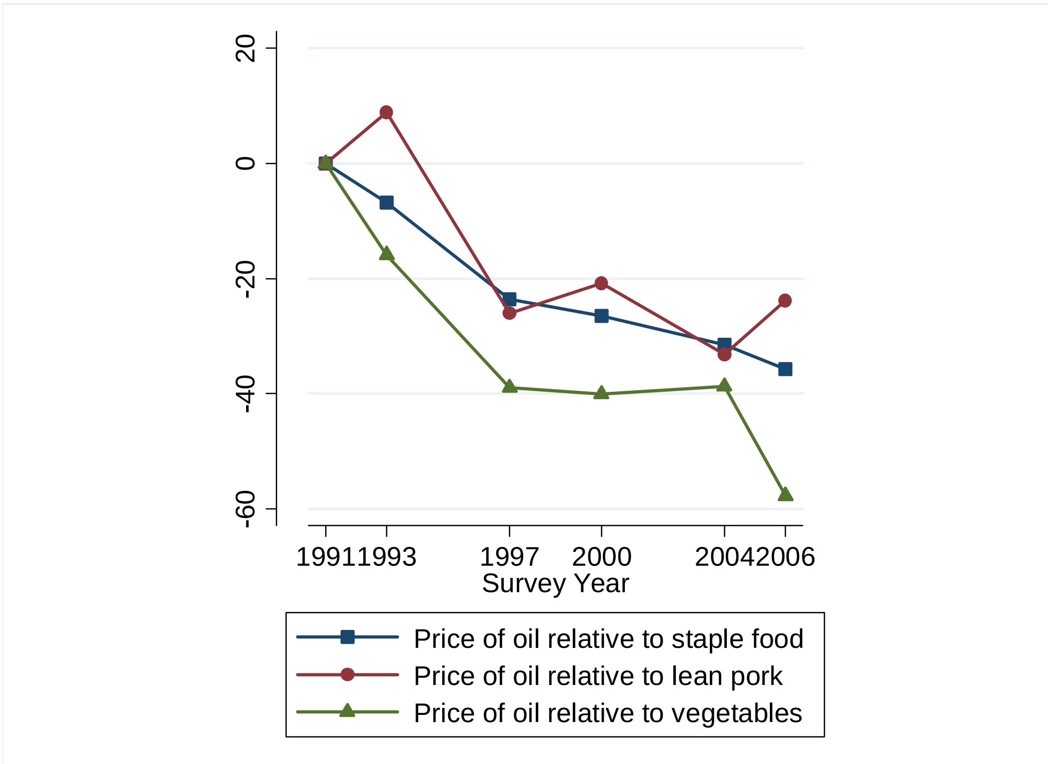

Figure 2—BMI and TSF trend

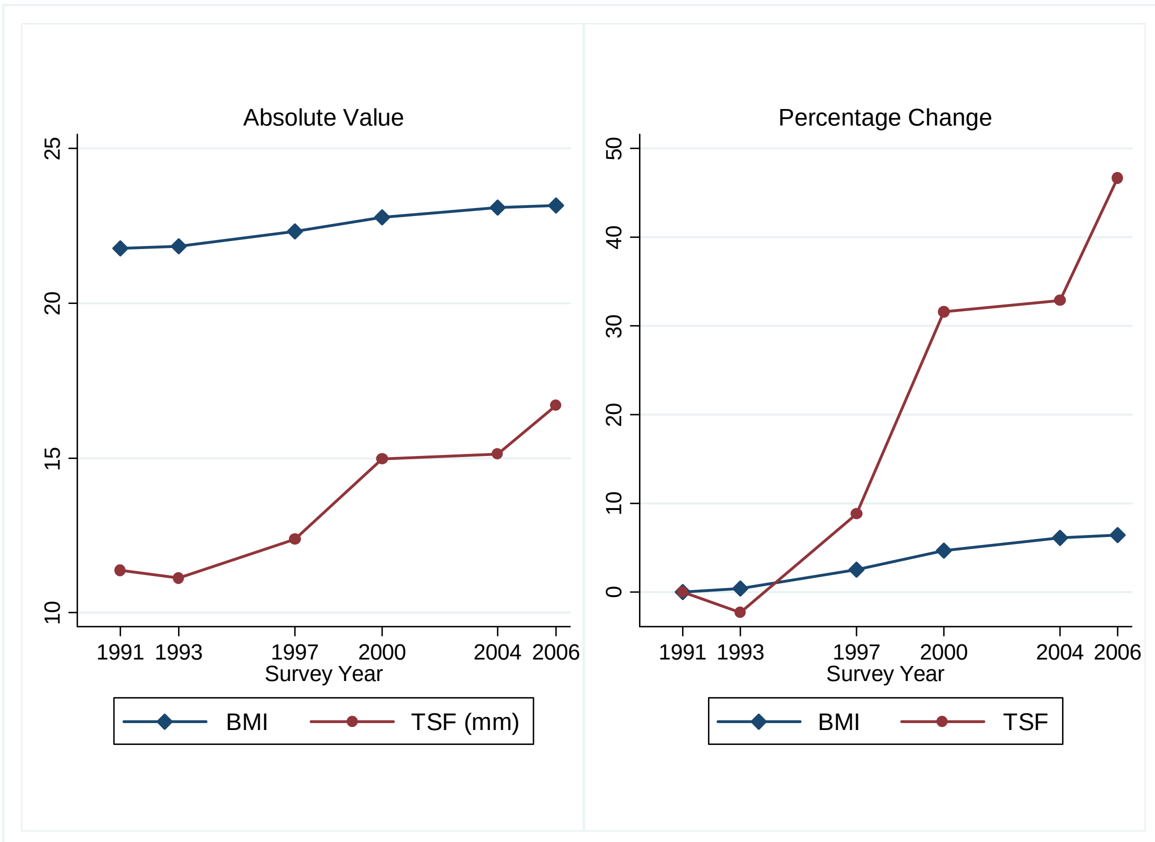

Figure 3—Percentage changes in BMI and TSF vs. in food prices

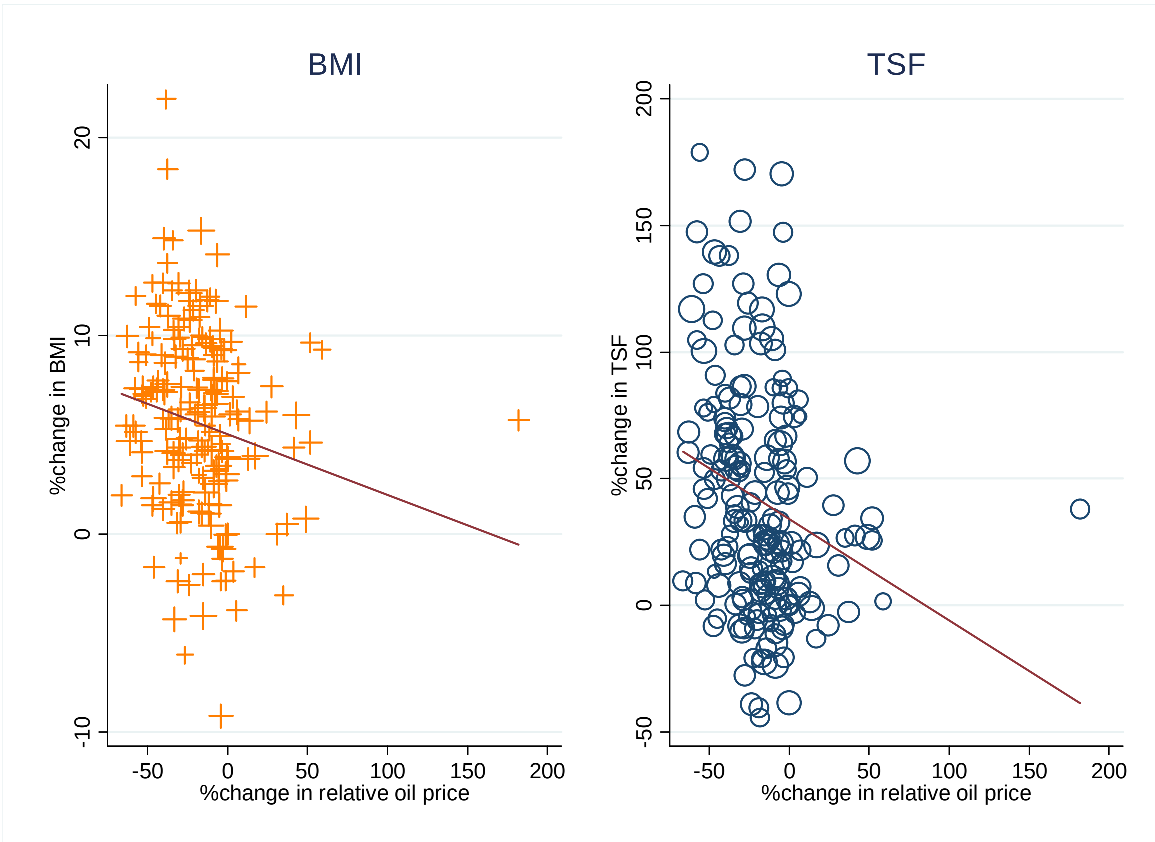

Table 1—Demographics

	All		Female		Male	
Variable List	N	Mean	N	Mean	N	Mean
HH income	40164	13950	20826	13797	19338	14114
HH income per person	40164	3964	20826	3911	19338	4020
Female	40164	0.52	20826	1	19338	0
Urban	40164	0.34	20826	0.34	19338	0.34
Age	40164	44.06	20826	44.05	19338	44.08
Obesity indicators						
BMI	40164	22.55	20826	22.67	19338	22.42
TSF (mm)	40164	13.86	20826	15.91	19338	11.64
Food Prices						
Pr(staple oil)	40164	3.79	20826	3.77	19338	3.80
Pr(staple food)	40164	1.21	20826	1.21	19338	1.21
Pr(vegetables)	39217	0.54	20344	0.55	18873	0.54
Pr(lean pork)	37816	7.01	19610	6.99	18206	7.02
P(oil)/P(staple food)	40164	3.46	20826	3.46	19338	3.46
P(oil)/P(lean pork)	37816	0.59	19610	0.59	18206	0.59
P(oil)/P(vegetables)	39217	9.52	20344	9.47	18873	9.58
Physical Activity Level						
Light	40164	0.39	20826	0.42	19338	0.35
Moderate	40164	0.17	20826	0.14	19338	0.20
Heavy	40164	0.44	20826	0.43	19338	0.46
Education						
No education	40164	0.25	20826	0.34	19338	0.15
Elementary school	40164	0.22	20826	0.21	19338	0.23
Middle School	40164	0.31	20826	0.27	19338	0.35
HS diploma or equiv	40164	0.18	20826	0.15	19338	0.21
Bachelor +	40164	0.05	20826	0.03	19338	0.06
Fuel Prices (IV)						
Gasoline/L	29310	3.09	15178	3.09	14132	3.09

Table 2—Effects of staple oil prices relative to staple food on BMI and TSF

	Outcome: Log(BMI)			Outcome: Log(BMI)		
Variables	OLS	RE	FE	OLS	RE	FE
Log(staple oil/staple food)	-0.005	0.003	0.003	-0.153***	-0.131***	-0.079***
	[0.005]	[0.002]	[0.002]	[0.029]	[0.026]	[0.028]
Log(lean pork)	-0.004	0.005*	0.006**	-0.003	0.027	0.080***
	[0.007]	[0.003]	[0.003]	[0.029]	[0.026]	[0.029]
Log(vegetables)	-0.025***	-0.010***	-0.004*	0.035*	0.029	-0.012
	[0.004]	[0.002]	[0.002]	[0.020]	[0.019]	[0.020]
Log(BMI)				1.707***	1.617***	1.256***
				[0.043]	[0.042]	[0.065]
Moderate PAL	-0.015***	-0.005***	-0.001	-0.015	-0.011	0.010
	[0.003]	[0.002]	[0.002]	[0.012]	[0.011]	[0.011]
Heavy PAL	-0.042***	-0.015***	-0.005***	-0.090***	-0.071***	0.007
	[0.005]	[0.002]	[0.002]	[0.017]	[0.015]	[0.012]
Age	0.011***	0.010***	0.091***	-0.002	0.001	-0.687***
	[0.000]	[0.000]	[0.011]	[0.001]	[0.001]	[0.167]
Age squared	-0.000***	-0.000***	-0.000***	-0.000	-0.000**	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
	[0.007]	[0.006]	[0.008]	[0.028]	[0.027]	[0.049]
Log(income)	0.009***	0.003***	0.001	0.027***	0.023***	0.009
	[0.001]	[0.001]	[0.001]	[0.006]	[0.005]	[0.006]
Urban	0.014**	0.015**		0.036	0.039	
	[0.006]	[0.006]		[0.026]	[0.026]	
Female	0.007**	0.006**		0.365***	0.359***	
	[0.003]	[0.002]		[0.013]	[0.013]	
Education	included	included	included	included	included	included
Year FE	included	included	included	included	included	included
Constant	2.754***	2.791***	-0.124	-2.955***	-2.794***	23.246***
	[0.021]	[0.012]	[0.395]	[0.155]	[0.147]	[5.937]
R-squared	0.108		0.148	0.401		0.177
Observations	36905	36905	36905	36905	36905	36905
Number of persons	15472	15472	15472	15472	15472	15472

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 3—Effects of staple oil prices relative to lean pork on BMI and TSF

	Outcome: Log(BMI)			Outcome: Log(TSF)		
Variables	OLS	RE	FE	OLS	RE	FE
Log(staple oil/lean pork)	-0.021***	-0.005**	-0.002	-0.164***	-0.141***	-0.083***
	[0.005]	[0.002]	[0.002]	[0.029]	[0.027]	[0.026]
Log(staple food)	-0.024***	-0.007**	-0.001	0.005	0.029	0.075**
	[0.005]	[0.003]	[0.003]	[0.028]	[0.028]	[0.032]
Log(vegetables)	-0.021***	-0.009***	-0.003	0.039*	0.030	-0.012
	[0.003]	[0.002]	[0.002]	[0.020]	[0.019]	[0.020]
Log(BMI)				1.689***	1.605***	1.256***
				[0.044]	[0.042]	[0.065]
Moderate PAL	-0.016***	-0.005***	-0.001	-0.021*	-0.015	0.010
	[0.003]	[0.002]	[0.002]	[0.012]	[0.011]	[0.011]
Heavy PAL	-0.042***	-0.015***	-0.005***	-0.087***	-0.070***	0.007
	[0.004]	[0.002]	[0.002]	[0.017]	[0.015]	[0.012]
Age	0.011***	0.010***	0.094***	-0.002	0.001	-0.687***
	[0.000]	[0.000]	[0.011]	[0.001]	[0.001]	[0.165]
Age squared	-0.000***	-0.000***	-0.000***	-0.000	-0.000**	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Log(income)	0.009***	0.003***	0.001	0.028***	0.023***	0.009
	[0.001]	[0.001]	[0.001]	[0.006]	[0.005]	[0.006]
Urban	0.011**	0.014**		0.025	0.032	
	[0.005]	[0.006]		[0.026]	[0.026]	
Female	0.007**	0.006**		0.365***	0.359***	
	[0.003]	[0.002]		[0.013]	[0.013]	
Education	included	included	included	included	included	included
Year FE	included	included	included	included	included	included
Constant	2.731***	2.800***	-0.220	-3.180***	-2.948***	23.244***
	[0.015]	[0.011]	[0.390]	[0.151]	[0.143]	[5.894]
R-squared	0.114		0.147	0.402		0.177
Observations	36905	36905	36905	36905	36905	36905
Number of persons	15472	15472	15472	15472	15472	15472

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 4—Effects of staple oil prices relative to vegetables on BMI and TSF

	Outcome: Log(BMI)			Outcome: Log(TSF)		
Variables	OLS	RE	FE	OLS	RE	FE
Log(staple oil/vegetables)	0.008**	0.007***	0.004**	-0.107***	-0.081***	-0.009
	[0.003]	[0.002]	[0.002]	[0.019]	[0.018]	[0.018]
Log(staple food)	-0.035***	-0.011***	-0.003	-0.010	0.015	0.053
	[0.005]	[0.003]	[0.003]	[0.028]	[0.028]	[0.033]
Log(lean pork)	-0.007	0.004*	0.006**	-0.011	0.017	0.060**
	[0.008]	[0.002]	[0.002]	[0.031]	[0.027]	[0.029]
Log(BMI)				1.710***	1.624***	1.256***
				[0.042]	[0.042]	[0.065]
Moderate PAL	-0.016***	-0.005***	-0.001	-0.018	-0.013	0.010
	[0.003]	[0.002]	[0.002]	[0.012]	[0.011]	[0.011]
Heavy PAL	-0.042***	-0.015***	-0.005***	-0.088***	-0.071***	0.006
	[0.005]	[0.002]	[0.002]	[0.017]	[0.015]	[0.012]
Age	0.011***	0.010***	0.091***	-0.002	0.000	-0.674***
	[0.000]	[0.000]	[0.011]	[0.001]	[0.001]	[0.169]
Age squared	-0.000***	-0.000***	-0.000***	-0.000	-0.000**	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Log(income)	0.009***	0.003***	0.001	0.028***	0.023***	0.010
	[0.001]	[0.001]	[0.001]	[0.006]	[0.005]	[0.006]
Urban	0.011**	0.014**		0.024	0.031	
	[0.006]	[0.006]		[0.026]	[0.026]	
Female	0.007**	0.005**		0.364***	0.358***	
	[0.003]	[0.002]		[0.013]	[0.013]	
Education	included	included	included	included	included	included
Year FE	included	included	included	included	included	included
Constant	2.755***	2.790***	-0.122	-2.958***	-2.819***	22.780***
	[0.021]	[0.012]	[0.395]	[0.153]	[0.146]	[6.034]
R-squared	0.107		0.148	0.401		0.175
Observations	36905	36905	36905	36905	36905	36905
Number of persons	15472	15472	15472	15472	15472	15472

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 5—Effects of absolute staple oil prices on BMI and TSF

	Outcome: Log(BMI)			Outcome: Log(TSF)		
Variables	OLS	RE	FE	OLS	RE	FE
Log(staple oil)	-0.033***	-0.003	0.004	-0.274***	-0.222***	-0.085**
	[0.007]	[0.003]	[0.003]	[0.041]	[0.036]	[0.035]
Log(staple food)	-0.018***	-0.008***	-0.003	0.055*	0.065**	0.075**
	[0.005]	[0.003]	[0.003]	[0.030]	[0.030]	[0.034]
Log(lean pork)	0.009	0.007***	0.006**	0.052*	0.065**	0.082***
	[0.006]	[0.003]	[0.003]	[0.030]	[0.029]	[0.031]
Log(vegetables)	-0.018***	-0.010***	-0.004*	0.063***	0.046**	-0.012
	[0.003]	[0.002]	[0.002]	[0.020]	[0.019]	[0.020]
Log(BMI)				1.675***	1.601***	1.256***
				[0.041]	[0.041]	[0.065]
Moderate PAL	-0.015***	-0.005***	-0.001	-0.016	-0.012	0.010
	[0.003]	[0.002]	[0.002]	[0.012]	[0.011]	[0.011]
Heavy PAL	-0.042***	-0.015***	-0.005***	-0.088***	-0.071***	0.007
	[0.004]	[0.002]	[0.002]	[0.017]	[0.014]	[0.012]
Age	0.011***	0.010***	0.091***	-0.002	0.000	-0.686***
	[0.000]	[0.000]	[0.011]	[0.001]	[0.001]	[0.167]
Age squared	-0.000***	-0.000***	-0.000***	-0.000	-0.000**	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Log(income)	0.009***	0.003***	0.001	0.027***	0.023***	0.009
	[0.001]	[0.001]	[0.001]	[0.006]	[0.005]	[0.006]
Urban	0.010**	0.014**		0.021	0.030	
	[0.005]	[0.006]		[0.025]	[0.025]	
Female	0.007**	0.006**		0.366***	0.359***	
	[0.003]	[0.002]		[0.013]	[0.013]	
Education	included	included	included	included	included	included
Year FE	included	included	included	included	included	included
Constant	2.767***	2.793***	-0.122	-2.816***	-2.711***	23.216***
	[0.021]	[0.012]	[0.395]	[0.156]	[0.146]	[5.953]
R-squared	0.115		0.148	0.408		0.177
Observations	36905	36905	36905	36905	36905	36905
Number of persons	15472	15472	15472	15472	15472	15472

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 6—Effects of staple oil prices relative to recommended protein and carbohydrate intake on BMI and TSF

	Outcome: Log(BMI)			Outcome: Log(TSF)		
Variables	OLS	RE	FE	OLS	RE	FE
Log(staple oil/rec'd intake)	-0.023***	-0.003	0.001	-0.222***	-0.193***	-0.117***
	[0.006]	[0.003]	[0.002]	[0.031]	[0.028]	[0.028]
Log(BMI)				1.670***	1.586***	1.248***
				[0.044]	[0.043]	[0.064]
Moderate PAL	-0.016***	-0.005***	-0.001	-0.018	-0.013	0.010
	[0.003]	[0.002]	[0.002]	[0.012]	[0.010]	[0.011]
Heavy PAL	-0.041***	-0.015***	-0.005***	-0.089***	-0.070***	0.011
	[0.005]	[0.002]	[0.002]	[0.017]	[0.015]	[0.012]
Age	0.011***	0.010***	0.101***	-0.002	0.000	-0.677***
	[0.000]	[0.000]	[0.011]	[0.001]	[0.001]	[0.164]
Age squared	-0.000***	-0.000***	-0.000***	-0.000	-0.000**	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Log(income)	0.007***	0.002***	0.000	0.029***	0.024***	0.009
	[0.002]	[0.001]	[0.001]	[0.006]	[0.005]	[0.006]
Urban	0.005	0.012**		0.039	0.045*	
	[0.006]	[0.006]		[0.026]	[0.026]	
Female	0.007**	0.005**		0.364***	0.357***	
	[0.003]	[0.002]		[0.013]	[0.013]	
Education	included	included	included	included	included	included
Year FE	included	included	included	included	included	included
Constant	2.803***	2.819***	-0.429	-3.033***	-2.816***	22.972***
	[0.016]	[0.010]	[0.389]	[0.147]	[0.139]	[5.856]
R-squared	0.103		0.146	0.403		0.175
Observations	37816	37816	37816	37816	37816	37816
Number of persons	15649	15649	15649	15649	15649	15649

Robust standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 7—Effects of relative oil prices on BMI and TSF by gender

	Log(BMI)	**Log(TSF)**
Female		
Log oil price relative to staple food	0.003	-0.072***
	[0.003]	[0.026]
Log oil price relative to lean pork	-0.004	-0.083***
	[0.003]	[0.030]
Log oil price relative to vegetables	0.003	-0.018
	[0.002]	[0.016]
Log oil price	0.002	-0.098***
	[0.004]	[0.034]
Log oil price relative to recommended intake	-0.001	-0.118***
	[0.003]	[0.030]
Male		
Log oil price relative to staple food	0.003	-0.086**
	[0.003]	[0.036]
Log oil price relative to lean pork	0.001	-0.085***
	[0.002]	[0.033]
Log oil price relative to vegetables	0.005**	0.000
	[0.002]	[0.024]
Log oil price	0.006*	-0.073
	[0.003]	[0.046]
Log oil price relative to recommended intake	0.003	-0.115***
	[0.003]	[0.035]
*** p<0.01, ** p<0.05, * p<0.1		

Table 8—Effects of instrumented staple oil prices on BMI, IV=gasoline prices

	Outcome: Log(BMI)				
Variables					
Log(staple oil/staple food)	-0.039***				
	[0.012]				
Log(staple oil/lean pork)		-0.098***			
		[0.033]			
Log(staple oil/vegetables)			-0.053***		
			[0.017]		
Log(staple oil)				-0.066***	
				[0.020]	
Log(staple oil/rec'd intake)					-0.066***
					[0.021]
Log(staple food)		0.009**	0.000	0.012**	
		[0.004]	[0.002]	[0.005]	
Log(lean pork)	0.011***		0.007***	0.023***	
	[0.002]		[0.002]	[0.005]	
Log(vegetables)	-0.005***	-0.005***		-0.001	
	[0.001]	[0.001]		[0.002]	
Moderate PAL	0.001	-0.001	0.001	0.000	-0.001
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Heavy PAL	-0.004**	-0.002	-0.004**	-0.004**	-0.004**
	[0.002]	[0.002]	[0.002]	[0.002]	[0.002]
Age	0.106***	0.096***	0.155***	0.105***	0.103***
	[0.009]	[0.009]	[0.021]	[0.009]	[0.008]
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Log(income)	0.000	-0.000	0.000	0.000	-0.000
	[0.001]	[0.001]	[0.001]	[0.001]	[0.001]
Education	included	included	included	included	included
Year FE	included	included	included	included	included
Observations	25335	25335	25335	25335	26273
Number of persons	8626	8626	8626	8626	8842
Weak ID test	347.6	54.93	93.98	224.5	150.8
UnID test	340.9	54.81	93.54	221.8	149.7

Standard errors in brackets
*** p<0.01, ** p<0.05, * p<0.1

7.1.1.1.1

Table 9—Effects of instrumented staple oil prices on TSF, IV=gasoline prices

	Outcome: Log(TSF)				
Variables					
Log(staple oil/staple food)	-0.214***				
	[0.070]				
Log(staple oil/lean pork)		-0.383**			
		[0.167]			
Log(staple oil/vegetables)			-0.187**		
			[0.076]		
Log(staple oil)				-0.277***	
				[0.106]	
Log(staple oil/rec'd intake)					-0.325***
					[0.116]
Log(staple food)		0.128***	0.072***	0.137***	
		[0.025]	[0.013]	[0.029]	
Log(lean pork)	0.094***		0.055***	0.124***	
	[0.014]		[0.013]	[0.027]	
Log(vegetables)	-0.024***	-0.028***		-0.013	
	[0.008]	[0.009]		[0.009]	
Log(BMI)	1.263***	1.240***	1.289***	1.261***	1.243***
	[0.048]	[0.050]	[0.051]	[0.048]	[0.048]
Moderate PAL	0.010	0.004	0.009	0.009	0.006
	[0.011]	[0.011]	[0.011]	[0.011]	[0.011]
Heavy PAL	0.016	0.025*	0.012	0.017	0.021*
	[0.012]	[0.013]	[0.012]	[0.012]	[0.012]
Age	-0.750***	-0.792***	-0.576***	-0.754***	-0.784***
	[0.055]	[0.052]	[0.098]	[0.054]	[0.050]
Age squared	-0.000***	-0.000***	-0.000***	-0.000***	-0.000***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Log(income)	0.010**	0.009**	0.011***	0.010**	0.011***
	[0.004]	[0.004]	[0.004]	[0.004]	[0.004]
Education	included	included	included	included	included
Year FE	included	included	included	included	included
Observations	24245	24245	24245	24245	25083
Number of persons	8327	8327	8327	8327	8538
Weak ID test	371.3	68.27	148.1	273.7	175.6
UnID test	363.2	68.05	146.9	269.4	173.9
Standard errors in brackets					
*** p<0.01, ** p<0.05, * p<0.1					