# Data

The data for this study come from the USDA's 1994-96 Continuing Survey of Food Intakes by Individuals (CSFII) and the companion Diet and Health Knowledge Survey (DHKS). Through dietary recalls, the CSFII contains information on the foods and nutrients consumed over 2 nonconsecutive days. The DHKS provides information on peoples' attitudes and knowledge about diet, health, and nutrition. In each CSFII household, the DHKS was administered to only one adult over 20 years old who reported at least 1 day of food intake. To maintain a clear link among food consumption, visceral influences, and dietary awareness, this analysis includes only individuals who were also given the DHKS. In total, we have information on 5,645 individuals who reported an average of 6.4 eating occasions, for a total sample of 36,312 observations. Descriptions of variables used in the empirical analysis and their summary statistics are in table 1.

## *Dependent Variables*

**Calories.** For the empirical analysis, the first step is to provide a suitable and measurable definition of unhealthful food choices made at each eating occasion. As energy imbalance and large portion sizes are often linked to weight gain and ultimately poor health, one measure used in this study is the number of calories consumed at an eating occasion. As such, one dependent variable is the share of recommended calories consumed at a specific eating occasion.[7] It is the number of calories an individual consumes divided by an individual's daily estimated energy requirement (EER), multiplied by 100. An individual's EER is calculated using the equations from the Institute of Medicine (2002), which is a function of an individual's age, gender, height, weight, and physical activity.[8] As an example, if an individual's daily EER is 2,000 calories and he or she consumes 500 calories at breakfast, then the dependent variable at that eating occasion is 25.

**Healthy Meal Index.** Low nutrient intake is also correlated with poor health and weight outcomes. Ideally, we would like a measure of nutrient intake per unit of energy consumed, such as nutrient density relative to energy density. At this time, however, there is no standard definition for this measure. To deal with this issue, we try three separate measures of nutrient quality. The first makes use of calculations from an individual's Healthy Eating Index for 1994-96. We refer to this measure as the HEI score. This index, which ranges from 0 to 100, summarizes how well an individual's daily food intake conforms to the 1995 *Dietary Guidelines for Americans*. The components of this index are an individual's intake of fruits, vegetables, grains, dairy products, meat, fat, saturated fat, sodium, and cholesterol (Bowman et al., 1998). To make this measure usable on a per-meal basis, we use a SAS program (available upon request from the USDA) to calculate an individual's daily HEI score.[9] We then augment the program to calculate what each person's daily score would have been in the absence of each eating occasion on a specific day.

The dependent variable at each eating occasion is the difference between what that individual's HEI score actually was on that day of intake and what it would have been in the absence of that specific eating occasion. For example,

[7]A potential weakness of looking at the relationship between meal timing and calories consumed on a per meal basis is that calories consumed at each meal may not adequately reflect total calories consumed over a day. Someone who frequently eats small meals could ultimately consume more total calories over the day. Using the same CSFII data and looking at calories consumed throughout the day, however, Mancino and Kinsey (2004) also found a positive and significant correlation between time between meals and total calories consumed.

[8]In our calculations, we assume that all individuals are inactive because the activity data available in the CSFII are not precise enough to include as part of the EER calculation.

[9]The official HEI score also includes a variety component. However, it is very difficult to calculate the change in daily HEI score while including this component. For that reason, the HEI scores in this study only include 9 components. Thus, the range of scores is from 0 to 90. We feel this does not alter results dramatically, especially since the HEI-2005 score does not include a variety component.

Table 1
**Variables, definitions and summary statistics**

| Variable | Definition | Mean (standard deviation) |
|---|---|---|
| Calories | Percent of daily energy requirements consumed at eating occasion | 22.26 (19.68) |
| Healthy Meal Index (HMI) | How much higher HEI score (0-100 points) would have been without eating occasion | 3.44 (6.21) |
| HMI2005 | HEI 2005 density score at eating occasion ( 0-100 points) | 37.70 (13.22) |
| SOFAAS per meal | Calories from solid fats, alcohol, and added sugar (0-20 points) | 10.29 (9.30) |
| Interval | Hours elapsed between current and previous eating occasion | 3.66 (2.22) |
| Share restaurant food | Share of food from restaurant (calories) | 13.39 (33.51) |
| Brunch | 1 if classified as brunch; 0 otherwise | 0.01 |
| Lunch | 1 if classified as lunch; 0 otherwise | 0.22 |
| Dinner | 1 if classified as dinner; 0 otherwise | 0.30 |
| Snack | 1 if classified as snack; 0 otherwise | 0.42 |
| Weekend | 1 if dietary recall occurred on a weekend; 0 otherwise | 0.26 |
| Knowledge | Score on dietary knowledge assessment (0-100) | 71.71 (13.35) |
| Work hours | Number of hours worked in previous week | 23.46 (23.39) |
| Interaction: Interval* knowledge | Interval*score on dietary knowledge assessment (0-100) | 260.51 (164.64) |
| Interaction: Interval*work hours | Interval*Number of hours worked in previous week | 86.40 (117.23) |
| Interaction: Sharefafh*knowledge | Share of restaurant food*score on dietary knowledge assessment | 954.42 (2,434.42) |
| Interaction: Sharefafh*work hours | Share of restaurant food*Number of hours worked in previous week | 400.55 (1,314.51) |

Each variable's standard deviation is listed below its mean for all continuous variables.
Sample size is 36,312, with 5,645 individuals who report an average of 6.43 eating occasions.

if an individual's HEI score on the first day of intake was 65 and that score would have been 60 had he or she not eaten breakfast that day, the dependent variable for that individual at that breakfast eating occasion would have a value of 5. Thus, higher values for this variable represent meal choices that are more nutritious. We refer to this variable as the Healthy Meal Index (HMI) to be clear that it is meant to gauge the nutritional quality of each eating occasion, whereas the HEI measures diet quality over an entire day.

Each of the 10 HEI components falls within the range of 1 to 10. Thus, someone who ate more than his or her recommended levels of fruit or vegetables could not score bonus points by doing so. Similarly, there are limits to "penalties" for overconsumption of salt, cholesterol, fat, and saturated fat. This top and bottom coding within components enables the HEI score to be more than a simple linear function of each meal's score. In the absence of such coding, however, the sum of HMI scores over an entire day would equal an individual's daily HEI score.

**Healthy Meal Index, 2005.** In 2005, the USDA and Department of Health and Human Services updated the *Dietary Guidelines for Americans*. More recently, the USDA has created the HEI-2005 (Guenther et al., 2007), which reflects the 2005 *Dietary Guidelines* and addresses some shortcomings of the HEI score. For example, it did not assess caloric intake or excesses of some components such as added sugars or alcohol. Also, the HEI score did not differentiate between types of vegetables, grains, or fats. As such, a serving of french fries and a serving of kale contributed equally to an individual's HEI score.

Intuitively, it seems better to gauge the links between dietary knowledge in 1994-96 and dietary intake in 1994-96 using the HEI score derived from the dietary guidelines promoted during this same time. However, we also run estimates using a Healthy Meal Index based on the HEI-2005 score to see how visceral influences, dietary knowledge, and meal location influence alternative measures of diet quality. The 12 components that make up the HEI-2005 score are total fruit, whole fruit, total vegetables, dark green and orange vegetables and legumes, total grains, whole grains, milk, meat and beans, oils, saturated fat, sodium, and calories from solid fats, alcohol, and added sugar (SoFAAS). Scoring for 10 of the 12 components is based on caloric density—either cups, grams, or ounce equivalents per 1,000 grams. Scoring for sodium and SoFAAS is calculated as a percent of energy. As such, it is easier to translate this daily measure into a per meal measure. We simply calculate each of the 12 components using the calories consumed at that meal instead of the total calories consumed over the entire day. We refer to this measure as the Healthy Meal Index 2005 (HMI-2005).

**Calories from Solid Fats, Alcohol, and Added Sugar (SoFAAS).** The twelfth component—SoFAAS—measures discretionary calories. Excess calories and energy imbalance are the cause of weight gain, and limiting discretionary calories is an effective way to reduce caloric intake without necessarily reducing nutrient intake. We run a final model that estimates how visceral influences and other explanatory variables influence intake of discretionary calories, or SoFAAS, at a meal.

### *Explanatory Variables*

According to the theoretical model, per-meal food demand will be a function of food prices, income, current health, and meal-specific visceral factors such as hunger.[10] Although the CSFII does not explicitly ask individuals how hungry they were at each eating occasion, it does provide information on the time elapsed between eating occasions. We therefore use the interval between meals, measured in hours, as a proxy for an individual's level of hunger at a specific eating occasion. This effect may change after longer periods without food, so we also include a quadratic term for the interval between meals.

We include the type of eating occasion (breakfast, lunch, dinner, or snack) and whether or not the dietary recall occurred on a weekend or weekday, as these variables may also affect individuals' food choices. The data on dietary intake come from 2 nonconsecutive days, so there are two observations each day (the first eating occasion) where it is not possible to calculate the interval between meals. Rather than lose these observations, we assign the mean interval between meals for each individual and use that as the level of hunger for the first eating occasion (Cohen et al., 2003).

[10]Because we do not discuss the results of the random effects (RE) estimates in the body of the paper, we do not describe the time-invariant explanatory variables, such as proxies for health status, income, and prices, that were used in these models.

To test the proposition that dietary awareness will have less influence on food choice in the presence of strong visceral influences, we include an interactive term to estimate the effects of dietary knowledge as hunger increases. Responses from the DHKS are used to create an index to measure knowledge about health and nutrition for each individual. One-third of this knowledge index is created by summing the number of questions an individual answered correctly about the links between diet and health.[11] Specifically, individuals were asked if they knew how many servings of fruit, vegetables, meat, dairy, and grains they should consume each day. They scored one point for each correct answer. Another third of this index is created from scoring correct answers about the amounts of cholesterol, fat, and saturated fat in specific foods. The remaining third of the information index comes from the DHKS questions on the importance of certain dietary practices, such as eating enough fruits, vegetables, and fiber; limiting intake of fat, cholesterol, saturated fat, salt, and sugar; and maintaining a healthy body weight.

[11]Specific questions and answers used to create the dietary knowledge score are available from the authors upon request.

Where an individual makes his or her food choices may affect the types and amounts of foods consumed. To gauge this effect, we calculate the share of calories consumed at each meal that come from a restaurant (restaurants with table service, fast-food places, pizza places, and bars/taverns).[12] We also estimate if the effect of dietary knowledge wanes when consuming foods away from home because typically there is less health information about foods purchased at restaurants.

[12]We chose this technique over creating a similar dichotomous variable because some eating occasions contain foods from multiple sources.

Finally, stress of time pressures associated with work and family requirements may be a visceral influence that, like hunger, affects food choice. As a proxy for work requirements, we use the number of hours worked in the previous week. Although this specific variable is not in our empirical estimation, we do include an interactive term to determine whether hunger has a stronger effect on individuals who are more time constrained through work. We also interact *hours worked* with food away from home to estimate if time stresses increase an individual's vulnerability to certain pitfalls of eating away from home, such as an expanded array of unhealthy food choices or less information about diet/nutrition.

### *Instrumental Variables*

Using a fixed-effects estimator should drastically reduce the correlation between the disturbance terms and the other explanatory variables. However, it is possible that the variable we use to proxy an individual's level of hunger may be correlated with some unobserved time-varying factor. Obtaining unbiased, efficient estimates requires that the additional variables be both relevant and independent. Because we assume a fixed-effects estimator, the instruments also need to vary across observations for a single individual. Therefore, some time-invariant variables, such as health conditions of other family members, are not viable options. In the fixed-effects model, there are four variables that are created using the interval between meals (hunger, hunger squared, the interaction of hunger with dietary knowledge, and the interaction of hunger with hours worked).[13] Thus, these variables may be correlated with the individual-specific, time-varying error term. One possible instrument is to use the time of day an eating occasion occurred. The time of day is exogenous—while people may choose how long they go between meals, they do not choose the actual time

[13]The share of calories consumed away from home may also be correlated with the time-specific error term. We had originally treated this variable as endogenous, but the resulting instrumental variable (IV) estimator was overidentified. We chose to focus on the endogeneity involved with the interval between meals because of our theoretical model.

of day. However, certain times, such as 4 p.m., may be more correlated with longer intervals between meals compared to 12:30 p.m. Following the technique developed by Arellano and Bond (1991), we also use lags of the endogenous variables as instruments.[14] How long one has gone between meals previously is likely to be correlated with how long one has gone between meals at present. However, feeling hungry at lunch should not influence how hungry one feels at dinner.[15]

In summary, we employ a fixed-effects estimator with instrumental variables (FE-IV) described as follows:

$$(F_{it} - \overline{F}_i) = \beta'(X_{it} - \overline{X}_i) + \gamma'(\alpha_{it} - \overline{\alpha}_i) + (e_{it} - \overline{e}_i) \quad (10a)$$

$$(\alpha_{it} - \overline{\alpha}_i) = \xi' + (Z_{it} - \overline{Z}_i) + (\varepsilon_{it} - \overline{\varepsilon}_i). \quad (10b)$$

We use four different dependent variables: the share of an individual's recommended daily caloric intake consumed at an eating occasion; the difference between what an individual's HEI score would have been in the absence of that specific eating occasion and what it was on that day of intake (HMI); the HEI-2005 density score per eating occasion (HMI-2005); and the share of discretionary calories consumed at each eating occasion (SoFAAS). In each model, we include proxies for visceral factors—stress and hunger—and the interaction of these visceral factors with nutrition information as explanatory variables.

[14]This requires us to drop the first eating occasion of each day.

[15]One alternative adopted by both Park and Davis (2001) and Abdulai and Aubert (2004) is to follow the method developed by Lewbel (1997), where the second and third moments of the endogenous variables are used as additional instruments in the IV estimation. We pursued this strategy as well, but found that in all cases, the models were overidentified.