Effects of Family, Friends, and Relative Prices on Fruit and Vegetable Consumption by African American Youths

Oleksandr Zhylyevskyy, Helen H. Jensen, Steven B. Garasky,
Carolyn E. Cutrona, and Frederick X. Gibbons

Working Paper 10-WP 515
September 2010

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

Oleksandr Zhylyevskyy is an assistant professor and Helen H. Jensen is a professor in the Department of Economics, Steven B. Garasky is a professor in the Department of Human Development and Family Studies, and Carolyn E. Cutrona is a professor and chair of the Department of Psychology, all at Iowa State University. Frederick X. Gibbons is a research professor in the Department of Psychological and Brain Science at Dartmouth College.

The data for this research were developed with support under grants MH48165, MH62666, and MH62668 from the National Institute of Mental Health and grants 029136-02 and 1-UO1-CE001645 from the Centers for Disease Control and Prevention, with additional funding provided by the National Institute on Drug Abuse, the National Institute on Alcohol Abuse and Alcoholism, and the Iowa Agriculture and Home Economics Experiment Station (Project 3320).

Questions or comments about the contents of this paper should be directed to Oleksandr Zhylyevskyy, 460D Heady Hall, Iowa State University, Ames, IA 50011-1070; Ph: (515) 294-6311; E-mail: oz9a@iastate.edu.

Abstract

Facilitating healthy eating among young people, particularly among minorities who are at high risk for gaining excess weight, is at the forefront of the current policy discussions in the U.S. We investigate the effects of social interactions and relative prices on fruit and vegetable consumption by African American youths. We estimate a simultaneous equation ordered probit model of food intake using rich behavioral data from the Family and Community Health Study and price data from the Economic Research Service's Quarterly Food-at-Home Price Database. We find the presence of endogenous effects between a youth and parent, but not between a youth and friend. Lower relative prices of fruits and vegetables tend to increase intakes. Results suggest that health interventions targeting only one family member may be a cost-effective way to increase fruit and vegetable intake by African Americans because of the existence of "spillover" consumption effects between the youths and their parents.

Keywords: endogenous effects, fruit and vegetable consumption, healthy food choices, simultaneous equation ordered probit model, social interactions.

JEL Codes: I12, J15, C35

1 Introduction

Good nutrition in adolescence is key to positive growth and development early in life. Moreover, since dietary patterns formed during teenage years tend to persist into adulthood, adequate nutritional intake by young people sets the stage for maintaining good health later on.

Presently, eating habits of U.S. youths fall short of the federal Dietary Guidelines for Americans (U.S. Department of Health and Human Services and U.S. Department of Agriculture, 2005). Health professionals are particularly alarmed by adolescents' tendency to consume lower amounts of fruits and vegetables than recommended (Dietary Guidelines Advisory Committee, 2010; Task Force on Childhood Obesity, 2010), since the scientific literature indicates that fruit and vegetable intake may protect against cancer, provide benefits against other illnesses, and reduce the likelihood of gaining excess weight (for reviews of this literature, see Van Duyn and Pivonka, 2000; He et al., 2006; He et al., 2007; World Cancer Research Fund, 2007; Dietary Guidelines Advisory Committee, 2010). Given the staggering cost of treating obesity-related ailments (Finkelstein et al., 2009) and repercussions from the high prevalence of overweight among U.S. youths (Ogden et al., 2010) such as, for example, a shortening of life expectancy (Olshansky et al., 2005) and reduction in the country's military readiness (Christeson et al., 2010), shifting dietary patterns among young people toward "energy light" and "nutrient dense" foods such as fruits and vegetables has moved to the forefront of public policy discussions.

In this paper, we investigate determinants of fruit and vegetable consumption by African American youths. More specifically, we exploit the richness of behavioral data collected by the Family and Community Health Study (FACHS) and area-specific food price data compiled by the Economic Research Service (ERS) of the U.S. Department of Agriculture to estimate the effects of fruit and vegetable consumption by an African American youth's parent and best friend and of relative food prices on the youth's own consumption. By performing this analysis we shed further light on factors underlying dietary choices of young people and provide recommendations for developing policy interventions to facilitate healthy eating. Doing so in the context of food consumption by African American youths is particularly important, because African Americans are at an elevated risk for gaining excess weight and having inadequate nutritional intake (Freedman et al., 2008; Dietary Guidelines Advisory Committee, 2010; Ogden et al., 2010; Task Force on Childhood Obesity, 2010).[1]

This research contributes to the literature along several dimensions. First, we augment a standard economic framework in which individuals engage in various health behaviors (Cawley, 2004) by explicitly allowing for impacts of social interactions (Manski, 1993) on a youth's food choice. Presence of peer social effects is well established in the case of young people's substance use (e.g., Gaviria and Raphael, 2001; Powell et al., 2005; Krauth, 2006; Lundborg, 2006; Clark and Lohéac, 2007; Fletcher, 2010). Supporting evidence is also emerging on the role of social interactions in the spread of obesity (e.g., Christakis and Fowler, 2007; Renna et al., 2008; Trogdon et al., 2008). However, apart from suggestive qualitative evidence from focus group studies (e.g., Neumark-Sztainer et al., 1999), little is

[1] For example, African Americans have the lowest intakes of fruits and vegetables among all main ethnic groups in the U.S. (Dietary Guidelines Advisory Committee, 2010, p. B3-1).

known about the effects of parental and peer eating habits on adolescents' food choices.[2] A better understanding of various factors shaping dietary patterns is critical for developing programs to address poor eating habits, particularly because social interactions can amplify the effectiveness of health policy interventions. What is especially novel about our approach is that the richness of behavioral data in the FACHS allows us to assess the impact of the parent's food consumption on the youth separately from the effect of the best friend's consumption.[3]

Second, we contribute to the literature on the effects of food prices on health and nutrition. To date, much attention has been paid to the impact of food prices on body mass index (e.g., Chou et al., 2004; Auld and Powell, 2009; Powell, 2009). Substantial knowledge has also accumulated on the magnitude of price effects on household food demand (see Dong and Lin, 2009; Unnevehr et al., 2010). We add to this body of research by evaluating whether individual consumption of fruits and vegetables by African Americans is sensitive to changes in the relative prices when controlling for social interactions. Since lowering relative prices of healthy foods is suggested as a means of improving food choices and reducing obesity among U.S. youths (Task Force on Childhood Obesity, 2010), knowledge of whether individual eating habits are significantly affected by relative prices may be more informative to policymakers than the existing household-level price elasticity estimates. A unique aspect of our empirical strategy is the use of the ERS's Quarterly Food-at-Home Price Database (QFAHPD), which is a detailed source of food prices faced by the U.S. population and available at relatively disaggregated geographic areas. This database includes comprehensive information on fruits and vegetables.

Third, we propose and estimate a simultaneous equation ordered probit model, which, to the best of our knowledge, is methodologically novel in the literature on social interactions. The simultaneous equation strategy allows us to incorporate assumptions about the nature of social interactions among a youth, parent, and friend.

Our main estimation results are as follows. In the case of fruit, we detect the presence of statistically significant endogenous consumption effects between an African American youth and his or her parent, but not between the youth and friend. The effect of the relative fruit price on the fruit intake by the parent and friend is estimated to be negative and statistically significant. The price impact on the youth's intake is also negative but not significant. In the case of vegetables, we find a positive statistically significant impact of the parent's consumption on the youth's consumption, but we find no impact in the reverse direction and no apparent endogenous consumption effects between the youth and friend. The relative vegetable price tends to have a statistically significant negative impact on the intake of vegetables by the youth and friend. The estimated price effect on the parent is not

[2] Early experimental studies demonstrated that the amount of food intake by an individual tended to increase with the number of other people present during the meal (e.g., de Castro and Brewer, 1991), but these studies did not attempt to address social interaction effects. More recently, Epstein et al. (2001) showed that modifying parental eating through an intervention might induce a nutritional change in children. We are not aware of any similar studies focusing on the effects of parental and peer eating habits on adolescents' food choices.

[3] It is not possible to perform a similar analysis using the National Longitudinal Study of Adolescent Health, a dataset often employed to study food consumption by adolescents (e.g., Videon and Manning, 2003; Stewart and Menning, 2009), because it does not contain information on parental eating habits.

significant. Overall, the results suggest that among African American youths, eating habits are formed mainly through "mimicking" dietary choices of parents rather than by learning from food habits of friends.

The results imply that designing health policy interventions to increase fruit and vegetable consumption by only one family member such as, in particular, the mother – the most likely primary caregiver – may be a cost-effective way to facilitate healthy food choices among African Americans, because increasing parental consumption of fruits and vegetables tends to have a "spillover" impact in the form of a higher intake of these foods by adolescent children. In contrast, we find little evidence to support that peer-group-based interventions may be associated with similar spillover effects. The estimates also suggest that decreasing the relative price of fruit (by subsidizing fruit or by taxing other foods) may raise the intake of fruit by parents of the youths and, because of the spillover effect, may increase fruit consumption by the youths themselves. In contrast, lowering the relative price of vegetables may increase the intake of vegetables by the youths, but is unlikely to have any additional spillover impact.

The remainder of this paper is organized as follows. In Section 2, we provide details on the data used in the analysis. In Section 3, we describe the theoretical framework, specify the econometric model, and outline the estimation approach. In Section 4, we discuss the estimation results and their implications for policy design. We conclude in Section 5 and relegate additional information to appendices.

2 Data

2.1 Family and Community Health Study (FACHS)

Our main data source is Wave 4 of the FACHS, which is an ongoing panel survey of African American youths, sponsored by the National Institutes of Health, designed to measure youths' health and development. It has been extensively used in psychological and sociological research (e.g., Gibbons et al., 2004; Pomery et al., 2005; Granberg et al., 2009).

The FACHS originated in the mid-1990s as a survey of African American children between the ages of 10 and 12 and their immediate family members in Georgia and Iowa. In Georgia, respondents were recruited by community liaisons, who contacted families with children who met sampling criteria to determine their interest in participating. In Iowa, project staff obtained school rosters of students in grades four through six and invited families with children to participate. Wave 1 of the study, completed between January 1997 and June 1998, covered a sample of 897 families, of which 714 youth respondents were re-interviewed in Wave 4, which started in March 2005 and lasted until June 2007.

In Wave 4, the FACHS introduced a major expansion resulting from a grant from the Centers for Disease Control and Prevention (CDC). In particular, every youth respondent (in what follows, we refer to the youth respondent as the *youth*) was requested to name his or her best same-sex friend (*friend*). This friend participated in the study along with the youth's immediate family members, namely, the youth's primary caregiver (*parent*),[4] second

[4] We use the term "parent" as a shortcut alternative to the FACHS term "primary caregiver," since 98.2% of the primary caregivers in our sample are either natural parents or parent-like figures. More specifically,

caregiver, and older sibling, if any. This feature of Wave 4 makes it particularly suitable to our research, because it provides us with data to disentangle the effect of friends, who are not immediate family members, from the effect of parents on food choices of the youths. In total, we have complete observations for 502 youth-friend-parent *triplets*.[5]

2.2 Fruit and Vegetable Consumption in FACHS

The youth, friend, and parent were asked two questions about their food choices in the week preceding the Wave 4 interview. First, they were asked to identify how often they ate fruit or drank fruit juice: *During the past seven days, how many times did you eat a whole piece of fruit (for example, an apple, orange or banana) or drink a glass of 100% fruit juice (do not count punch, Kool-Aid, or sports drinks)?*[6] Second, everyone reported the frequency of vegetable intake: *During the past seven days, how many times did you eat vegetables like green salad, carrots or potatoes (do not count French fries, fried potatoes, or potato chips)?*

The answer categories for the fruit and vegetable consumption questions were (1) none, (2) less than once a day (1-6 times), (3) once a day, (4) 8-12 times, (5) twice a day (or more). Summary statistics for the answers are provided in Table 1. We use the reported "food frequency" as an indicator of the amount consumed.[7]

As can be inferred from Panel A of Table 1, approximately 61%, 60%, and 65% of the youths, friends, and parents, respectively, report consuming fruit at least once a day in the week preceding the interview, while non-negligible fractions of the youths (13%), friends (15%), and parents (11%) say that they neither ate fruit nor drank fruit juice. Panel B reveals pronounced differences in the reported vegetable consumption between the youths and friends, on the one hand, and parents, on the other. In particular, 60% of the youths and 59% of friends report eating vegetables at least once a day, but the corresponding fraction among parents is much larger, at 76%. The difference is stark when we consider the incidence of no vegetable consumption (except for fried potatoes and potato chips): 14% of the youths and 15% of friends say that they ate no vegetables in the last seven days, but the corresponding fraction of parents is a mere 3%.

Since the FACHS is not intended to be nationally representative, it is important to explore whether conclusions of our analysis using this study of African Americans from, primarily, Georgia and Iowa may apply to a broader population of African Americans in the U.S. In Appendix A, we investigate comparability of the reported fruit and vegetable consumption patterns of the FACHS participants to consumption patterns in relevant samples from the National Health and Nutrition Examination Survey (NHANES) 2005-2006. The comparison

---

the primary caregiver is the mother of the youth in 86.4% of the cases, while in 3.8%, 3.8%, 2.8%, and 1.4% of the cases the primary caregiver is the father, grandmother, adoptive parent, and some other parent-like figure (e.g., a step-parent, aunt, or uncle), respectively. Only in 1.8% of the cases the primary caregiver is the sister, brother, cousin, or some other relative or non-relative of the youth.

[5] We do not use data on second caregivers and siblings, because the corresponding sample sizes are small.

[6] The data do not allow us to distinguish between consumption of whole fruit and consumption of 100% fruit juice. The definition we use is consistent with the one employed in a recent CDC report (Centers for Disease Control and Prevention, 2010).

[7] Although the measurement error associated with reported food frequency may be substantial when the frequency is used to estimate usual dietary intake, Subar et al. (2006) find a positive and significant correlation between the food frequency measures and mean 24-hour intake, especially for food groups.

Table 1: Fruit and vegetable consumption in FACHS

Answer	Youth, %	Friend, %	Parent,%
Panel A: During the past 7 days, how many times did you eat a whole piece of fruit or drink a glass of 100% fruit juice?			
(1) none	12.75	14.94	10.96
(2) less than once a day (1-6 times)	26.49	24.70	23.71
(3) once a day	30.48	30.88	40.24
(4) 8-12 times	11.55	8.37	6.77
(5) twice a day (or more)	18.73	21.12	18.33
Total	100.00	100.00	100.00
Panel B: During the past 7 days, how many times did you eat vegetables like green salad, carrots or potatoes?			
(1) none	13.75	14.94	3.19
(2) less than once a day (1-6 times)	26.10	26.29	20.52
(3) once a day	37.85	35.86	43.82
(4) 8-12 times	8.17	7.77	8.96
(5) twice a day (or more)	14.14	15.14	23.51
Total	100.00	100.00	100.00

Notes:
The number of youth-friend-parent triplets is 502.
Fractions may not add up to 100% because of rounding.

indicates that the difference between the patterns in the FACHS and NHANES is small. Thus, the food consumption habits of the FACHS participants appear to be in line with the habits of the corresponding U.S. population.[8]

2.3 Characteristics of FACHS Participants

Summary statistics for selected demographic and socioeconomic characteristics of individuals in our FACHS sample are provided in Table 2. As can be seen in Panel A, youths are between the ages of 17 and 22 and, on average, 19.3 years old. Forty two percent are male, and 96% identify themselves as African American (the rest mostly identify themselves as biracial).

In Panel B, we report characteristics of friends. In comparison to the youths, the age of the friends shows more variation, as they are between 14 and 52 years old. However, their average age is 19.9 years, which is only slightly higher than the average youth's age. Because of the Wave 4 restriction on the sex of the friends, the proportion of males among them is identical to the proportion of males among the youths (42%). In contrast, there is no restriction on the race of friends: 84% of them are African American, which is a lower proportion than among the youths.

[8]This comparison is only suggestive rather than definitive because it is not possible to perform a formal statistical test of whether the consumption patterns in the FACHS and NHANES are identical.

Table 2: Demographic and socioeconomic characteristics of FACHS participants

Characteristic	Mean	Std. Dev.	Min	Max
Panel A: Youth				
Age in years	19.28	(0.83)	16.85	21.89
Indicator of male sex[a]	0.42	(0.49)	0	1
Indicator of African American race[b]	0.96	(0.20)	0	1
Panel B: Friend				
Age in years	19.87	(3.34)	13.54	51.59
Indicator of male sex[a]	0.42	(0.49)	0	1
Indicator of African American race[c]	0.84	(0.36)	0	1
Panel C: Parent				
Age in years[d]	45.06	(7.68)	32.56	88.87
Indicator of male sex	0.05	(0.22)	0	1
Indicator of African American race[e]	0.92	(0.27)	0	1
Indicator of no high school degree[f]	0.18	(0.38)	0	1
Indicator of high school degree[f]	0.34	(0.47)	0	1
Indicator of some college education[f]	0.35	(0.48)	0	1
Indicator of BA or higher degree[f]	0.14	(0.35)	0	1
Indicator of married parent	0.36	(0.48)	0	1
Indicator of poverty[g]	0.28	(0.45)	0	1

Notes:

The number of youth-friend-parent triplets is 502.

[a]Youth and friend are always of the same sex by the FACHS Wave 4 design.

[b]Twenty two youths report a race other than African American: 18 identify themselves as biracial, 3 as Caucasian, and 1 as "other."

[c]Seventy eight friends report a race other than African American: 41 identify themselves as Caucasian, 24 as biracial, 4 as Asian, 4 as Latino, 3 as American Indian, and 2 as "other."

[d]We report minimum age excluding 6 observations when the age difference between the primary caregiver and youth is less than 10 years. In these 6 cases, the primary caregiver is the youth's older sibling or cousin.

[e]Forty parents report a race other than African American: 31 identify themselves as Caucasian, 4 as Latino, 3 as biracial, and 2 as American Indian.

[f]Educational categories represent the highest level of educational attainment.

[g]Poverty status is imputed using household composition, income of family members, and official poverty thresholds. In 5.8% of the cases, we are unable to impute it because of missing income data.

In Panel C, we summarize characteristics of parents. They are between the ages of 33 and 89 and, on average, 45.1 years old.[9] Eighteen percent have no high school degree, 34% have a high school degree or GED, 35% report one to three years of college education or technical training but no bachelor's or higher degree ("some college"), and 14% have a bachelor's or higher degree. Of parents, most are females (95%) and African American (92%); only 36% are married. We impute the poverty status of the parent's household using official poverty thresholds from the U.S. Bureau of Census and information on the household composition and income and the date of data collection.[10] The resulting incidence of poverty in our sample is 28%.[11]

To investigate whether demographic and socioeconomic characteristics of our sample are in line with characteristics of the corresponding U.S. population, we performed a comparison of parents in the sample to a relevant subsample from the Current Population Survey (CPS). Details are presented in Appendix B. On the basis of the comparison, we conclude that basic demographic characteristics are virtually identical in both cases, except that the proportion of married parents is lower in the FACHS. Also, parents in the FACHS tend to have less income and a somewhat lower educational attainment than parents in the CPS.

2.4 Food Price Measures

We construct measures for fruit and vegetable prices using data from the QFAHPD, which was released recently for public use by the ERS. This database contains quarterly prices (in dollars per 100 grams of food as purchased) for 52 separate food groups between 1999 and 2006 for 35 geographical market areas that cover the contiguous U.S. It is based on the Nielsen Homescan survey data, which include detailed information on purchases of barcoded and random-weight food items by a demographically balanced panel of metropolitan and nonmetropolitan households. The ERS aggregated the Homescan data into household-level quarterly prices for the food groups and then aggregated the household-level prices into quarterly market-area food-group prices (Todd et al., 2010).

The QFAHPD is well suited for studying determinants of fruit and vegetable consumption. In particular, it allows us to exploit geographical and time variation in food prices. Our ability to control for quarterly prices over time is especially useful, since relative prices of perishable goods such as fruits and vegetables are likely to fluctuate across seasons.[12]

[9] We report the minimum parent's age excluding six observations when the age difference between the primary caregiver and youth is less than 10 years. In these six cases, the primary caregivers are older siblings or cousins of the youth and are between 20 and 30 years old.

[10] To reduce the incidence of non-reporting of income, which plagued previous FACHS waves, parents in Wave 4 were requested to provide only a category for total family income, namely, less than $10,000, between $10,000 and $14,999, between $15,000 and $19,999, and so forth. We constructed the poverty indicator by comparing the midpoint of the reported income range to the appropriate Census threshold.

[11] Because of missing income data, we are unable to impute poverty status in 5.8% of the cases. To investigate the sensitivity of our results to this data limitation, we estimated empirical models while additionally controlling for missing poverty status with a separate indicator variable. The variable turned out to have no statistically significant impact.

[12] We acknowledge that the prices do not reflect all costs associated with consumption of fruits and vegetables. For example, they do not account for the opportunity cost of time spent shopping and preparing meals. Focus group studies (e.g., Neumark-Sztainer et al., 1999) indicate that the relatively high time cost of fruit and vegetable meals may partly explain why many youths choose to eat fast food instead.

For purposes of this research, the QFAHPD has some advantages over the database maintained by the Council for Community and Economic Research (C2ER, formerly known as ACCRA), a popular source of food price data in the literature. Most notably, the QFAHPD contains separate prices for two groups of fruit (fresh/frozen whole fruit and canned whole fruit), one fruit juice group, and twelve vegetable groups.[13] Each food group price aggregate is based on a range of food items purchased by households. In contrast, the C2ER database includes prices for only three specific fruit items and four vegetable products. Moreover, the QFAHPD is a more accurate source of prices faced by the population at large, because it incorporates transactions from all outlets, including grocery, drug, mass-merchandise, club, supercenter, and convenience stores. In contrast, the C2ER contains prices relevant only for households from the upper quintile of the income distribution and, as a rule, ignores prices at Walmart and other discount stores. The lack of store coverage in the C2ER is a serious limitation because food purchases at discount stores comprise over 30% of consumer food-at-home expenditures (Todd et al., 2010, p. 2).

We construct price measures separately for fruits and vegetables. In the case of fruit, we first compute an index of fruit prices for each market area in every quarter of 2005-2006. It is an expenditure-weighted average of prices of the fruit and fruit juice groups in the QFAHPD.[14] Next, we calculate an index of all non-fruit prices (also specific to market area and quarter). Lastly, we obtain a relative fruit price as the ratio of the fruit index to non-fruit index. Similarly, we obtain a relative vegetable price as the ratio of an expenditure weighted index for the vegetable groups to an index for all non-vegetable groups in the QFAHPD.

Our focus on relative prices deviates from the typical approach in the current literature of utilizing several non-relative prices in one regression specification. In part, our choice is motivated by the need for model parsimony given the modest size of the FACHS sample. Moreover, since our price variables are ratios of indices specific to market area and quarter, they account for market-area-specific price variation over time while eliminating the confounding effects of inflation. In Appendix C, we discuss the robustness of our result to inclusion of "raw" fruit and vegetable price indices in place of the relative ones.

We merged the price variables with the youth, friend, and parent records using information on the ZIP code of residence and date of the interview.[15] Summary statistics are provided in Table 3. As can be seen in Panel A, the FACHS participants face fruit prices with mean values between 0.45 and 0.47 (minimum of 0.38 and maximum of 0.53). The mean value estimates indicate that the cost of 100 grams of fruit constitutes, on average, 45%−47% of the cost of 100 grams of other foods. In turn, Panel B shows that the relative prices of vegetables vary from 0.42 to 0.57 with mean values between 0.48 and 0.49, indicating that

[13] The twelve vegetable groups are as follows: fresh/frozen dark green vegetables, canned fresh/frozen dark green vegetables, fresh/frozen orange vegetables, canned orange vegetables, fresh/frozen starchy vegetables, canned starchy vegetables, fresh/frozen other-nutrient dense vegetables, canned other-nutrient dense vegetables, fresh/frozen other-mostly water vegetables, canned other-mostly water vegetables, fresh/frozen/dried legumes, and canned/processed legumes.

[14] Quarterly expenditures on each food group are available in the QFAHPD.

[15] Out of all 1,506 respondents in our sample, 18 individuals were interviewed in the first and second quarters of 2007 (the rest were interviewed in 2005 or 2006). Because 2007 had no price data in the QFAHPD, we perform all empirical analyses while merging these 18 records with prices for the fourth quarter of 2006. To check the robustness of the results, we additionally estimated empirical models while excluding the corresponding triplets from the sample. The exclusion had no impact on the results (see Appendix C).

Table 3: Relative prices faced by FACHS participants

	Mean	Std. Dev.	Min	Max
Panel A: Relative fruit prices				
Relative price faced by youth	0.466	(0.033)	0.379	0.527
Relative price faced by friend	0.468	(0.033)	0.393	0.527
Relative price faced by parent	0.454	(0.023)	0.407	0.527
Panel B: Relative vegetable prices				
Relative price faced by youth	0.488	(0.025)	0.429	0.566
Relative price faced by friend	0.490	(0.025)	0.424	0.566
Relative price faced by parent	0.476	(0.030)	0.429	0.566

Notes:
The number of youth-friend-parent triplets is 502.
The price variables are specific to place of residence and interview date.

vegetables cost, on average, $48\% - 49\%$ of the price of other foods (by weight).

It is important to note that the price variables are specific to individuals rather than to triplets. For instance, the price faced by the youth need not be the same as the price faced by the friend. There are two reasons why the prices may be different. First, members of the same triplet may reside in different market areas. Second, they may have been interviewed during different quarters.[16] We exploit this price specificity in estimation.[17]

3 Empirical Model

3.1 Theoretical Framework

Our empirical analysis is based on a standard economic framework in which an individual maximizes his or her utility by engaging in behaviors related to work, leisure, home production, production of health and body weight and by consuming foods and other goods (Cawley, 2004). Food consumption affects utility through the enjoyment of eating meals and entertainment provided by dining with family and friends (Chou et al., 2004, p. 570). It also has an impact on utility through the effects of the diet on health and body weight. The individual makes his or her decisions subject to a budget constraint, which is affected by income and prices, a time constraint, and constraints imposed by biology. Outcomes such

[16] Out of all 502 triplets, in 63 triplets, at least one triplet member lives in a different market area than the other members. In 241 triplets, at least one member was interviewed during a different quarter (in nearly all such instances the interviews were conducted during adjacent quarters).

[17] After merging the price variables with individual records, we performed a diagnostic analysis of collinearity among the youth-, friend-, and parent-specific price vectors. In the case of fruit, the variance inflation factors range from 1.24 to 6.23. In the case of vegetables, the factors are between 1.19 and 2.94. While no formal, distribution-based cutoff value exists, the literature considers variance inflation factors below ten to be indicative of absence of a serious multicollinearity problem. Thus, the diagnostic analysis suggests that we can, indeed, use the youth-, friend-, and parent-specific relative prices as separate explanatory variables.

as a mix of consumed foods are derived based on marginal costs and benefits. Changes in relative prices of different foods are expected to affect the demand for them.

We augment this framework by allowing for social interactions. Social interactions pertain to the idea that the utility from a given action depends directly on the choices and, possibly, characteristics of others in the individual's reference group (e.g., family, friends, or coworkers), as opposed to the dependence that arises through the intermediation of markets (Brock and Durlauf, 2001). Thus, food consumption by the individual may depend not only on prices and the individual's own characteristics, but also on food choices and, possibly, characteristics of the reference group members (e.g., age, race, and education of friends). Social interactions may be asymmetric (Harris and López-Valcárcel, 2008).

In the literature, the impact of the behavior of others (in the reference group) on the individual's own behavior is known as the *endogenous effect*, while the impact of the characteristics of others is referred to as the *contextual effect* (Manski, 1993). Observable interdependence among the behaviors may also arise because of the *correlated effect*. Moffitt (2001) makes a useful distinction between its two sources. First, sorting may force individuals with similar unobservable preferences to be grouped together. Second, all reference group members may be affected by a common unobservable factor. Distinguishing among these various effects is crucial for designing public health policies, because the endogenous effect is associated with a social multiplier, which can amplify the effectiveness of policy interventions. In contrast, the contextual effect does not generate a multiplier and indicates the need for a different intervention design. In turn, the correlated effect means that neither behaviors nor characteristics of the reference group members have a causal impact on the individual's own behavior.

Identification of social interactions is a challenging econometric problem, and the research in this area is still ongoing (for a survey, see Soetevent, 2006). In our empirical model, we explicitly allow for endogenous effects and, except when indicated otherwise, follow a standard practice of restricting contextual effects (e.g., Gaviria and Raphael, 2001; Powell et al., 2005; Krauth, 2006; Lundborg, 2006; Trogdon et al., 2008). The endogenous effects are allowed to be asymmetric. We attempt to account for the correlated effect by explicitly allowing for unobservable determinants of behaviors to be correlated.

3.2 Econometric Model

We denote a generic youth, friend, and parent by Y, F, and P, respectively, and use these symbols in subscripts and variable names when appropriate. The Y-F-P triplets are indexed by t, $t = 1, 2, ..., T$, where T is the number of triplets in the sample ($T = 502$).

We are interested in explaining *consumption* of fruits and vegetables rather than the *number of times* someone ate them in the past week (Table 1). This number is only a proxy for unobservable consumption, which could potentially be measured in food weight, calories, or other units, reflecting a limitation of the FACHS in collecting food data. To account for the data limitation, we employ latent variables (for a justification of this methodological approach, see Cameron and Trivedi, 1986, p. 49). More specifically, we propose a simultaneous equation model that is an extension of the model of Maddala and Lee (1976) to a setting with ordered responses. The model is described below for the case of fruit consumption.

Its specification for vegetable consumption is analogous.[18] We do not combine fruit consumption and vegetable consumption together in order to preserve the available variation in the data.[19] To the best of our knowledge, the model is novel in that it considers social interactions in a multivariate ordered probit setting.

Let $w^*_{Y,t}$ be a latent continuous variable that reflects consumption of fruit by Y from triplet t.[20] Instead of $w^*_{Y,t}$, we observe a categorical answer $w_{Y,t}$ about the frequency of Y's consumption in the past week, namely, (1) none, (2) less than once a day (1-6 times), (3) once a day, (4) 8-12 times, or (5) twice a day (or more). For example, if Y reports having consumed fruit once a day, $w_{Y,t} = 3$. We assume that a particular value of $w_{Y,t}$ is observed whenever $w^*_{Y,t}$ falls between corresponding thresholds:

$$w_{Y,t} = j \text{ if and only if } \alpha_Y(j) < w^*_{Y,t} \leq \alpha_Y(j+1) \text{ for } j = 1, 2..., 5,$$

where the thresholds $\alpha_Y(1), \alpha_Y(2) ..., \alpha_Y(6)$ are six real constants such that $-\infty = \alpha_Y(1) \leq \alpha_Y(2) \leq ... \leq \alpha_Y(6) = +\infty$. We define latent variables $w^*_{F,t}$ and $w^*_{P,t}$, observed categorical answers $w_{F,t}$ and $w_{P,t}$, and thresholds $\alpha_F(1), ..., \alpha_F(6)$ and $\alpha_P(1), ..., \alpha_P(6)$ analogously.

Let a $k \times 1$ vector of characteristics of triplet t be denoted by $\mathbf{x}_t$. This vector includes a constant term, variables created from the demographic and socioeconomic characteristics of the FACHS participants (Table 2), as well as relative food price measures (Table 3). We assume that the vector of the observed data $(w_{Y,t}, w_{F,t}, w_{P,t}, \mathbf{x}'_t)'$ is independent and identically distributed (i.i.d.) across t. To facilitate further discussion, let a $k_Y \times 1$ vector $\mathbf{x}_{Y,t}$ be a subset of $\mathbf{x}_t$ specific to Y (in a sense to become self-evident shortly). Similarly, a $k_F \times 1$ vector $\mathbf{x}_{F,t}$ and $k_P \times 1$ vector $\mathbf{x}_{P,t}$ are subsets of $\mathbf{x}_t$ specific to F and P, respectively. In Table 4, we list all variables comprising $\mathbf{x}_t$ and indicate with "$\surd$" which vector – $\mathbf{x}_{Y,t}$, $\mathbf{x}_{F,t}$, or $\mathbf{x}_{P,t}$ – contains a particular variable (an explanation for the specification choice is provided shortly). Note that $\mathbf{x}_t$ does not contain body mass index (BMI) and individual income, since they may be endogenous with respect to food choice behavior.[21] Also, observe that $\mathbf{x}_t$ does not contain place of residence and seasonal indicators, because including them would leave substantially less variation in the prices for us to identify the effects of interest (for a similar approach, see Auld and Powell, 2009).

The model comprises three equations parameterized as follows:

$$\begin{aligned} w^*_{Y,t} &= w^*_{F,t} \cdot \gamma_{FY} + w^*_{P,t} \cdot \gamma_{PY} + \mathbf{x}'_{Y,t} \cdot \boldsymbol{\beta}_Y + \epsilon_{Y,t}, \\ w^*_{F,t} &= w^*_{Y,t} \cdot \gamma_{YF} + \mathbf{x}'_{F,t} \cdot \boldsymbol{\beta}_F + \epsilon_{F,t}, \\ w^*_{P,t} &= w^*_{Y,t} \cdot \gamma_{YP} + \mathbf{x}'_{P,t} \cdot \boldsymbol{\beta}_P + \epsilon_{P,t}. \end{aligned} \tag{1}$$

[18] In a separate analysis not reported here, we uncovered a positive correlation between fruit and vegetable responses by the same individual. This analysis suggests that the two choices are jointly determined and indicates a need for a large model in which fruit consumption and vegetable consumption of the Y-F-P triplet are simultaneously determined. However, this large model would be problematic to identify, since it is difficult to locate an exogenous variable that affects one's fruit but not vegetable consumption (and vice versa). We defer a comprehensive analysis of such a model to future research.

[19] In fact, our estimation results (see Section 4) suggest that it is useful to keep fruit consumption and vegetable consumption separate rather than merge them into a single variable.

[20] An interpretation of $w^*_{Y,t}$ that does not involve restricting its support is that $w^*_{Y,t}$ represents an increasing affine transformation of an underlying non-negative amount of consumed fruit.

[21] We believe that parental poverty status, which is included in $\mathbf{x}_t$, is less problematic, since it is computed using income of all family members.

Table 4: Explanatory variables

Variable in $\mathbf{x}_t$	$\mathbf{x}_{Y,t}$	$\mathbf{x}_{F,t}$	$\mathbf{x}_{P,t}$	Description
constant	$\surd$	$\surd$	$\surd$	Constant term
Y_age	$\surd$			Age of Y
Y_age2	$\surd$			Age squared of Y
Y_male[a]	$\surd$	$\surd$		Indicator of male sex of Y and F
F_age		$\surd$		Age of F
F_age2		$\surd$		Age squared of F
F_black		$\surd$		Indicator of African American race of F
P_age			$\surd$	Age of P
P_age2			$\surd$	Age squared of P
P_higher_educ[b]	$\surd$		$\surd$	Indicator of college education of P (with or without degree)
P_married	$\surd$		$\surd$	Indicator of married P
P_poverty	$\surd$		$\surd$	Indicator of P in poverty
Y_rel_price	$\surd$			Relative price of fruit faced by Y
F_rel_price		$\surd$		Relative price of fruit faced by F
P_rel_price			$\surd$	Relative price of fruit faced by P
	$k_Y = 8$	$k_F = 6$	$k_P = 7$	

Notes:
[a] Y and F are always of the same sex by the FACHS Wave 4 design.
[b] The omitted education category comprises P's with a high school degree or less.

In the system (1), parameter γ_{FY} measures an endogenous effect of fruit consumption by F from triplet t, $w^*_{F,t}$, on the consumption of fruit by Y from the same triplet, $w^*_{Y,t}$. Parameters γ_{PY}, γ_{YF}, and γ_{YP} have similar meaning. To derive a reduced form for the system (see Subsection 3.4), we assume that these parameters satisfy an inequality $\gamma_{PY} \cdot \gamma_{YP} + \gamma_{FY} \cdot \gamma_{YF} \neq 1$.

Next, a $k_Y \times 1$ vector $\boldsymbol{\beta}_Y$ represents parameters measuring a "conditional" effect of $\mathbf{x}_{Y,t}$ on $w^*_{Y,t}$ given fixed $w^*_{F,t}$ and $w^*_{P,t}$. It is important to recognize that in the simultaneous equation system (1), the conditional effect of $\mathbf{x}_{Y,t}$ on $w^*_{Y,t}$ does not coincide with its "full" effect, which does not involve fixing $w^*_{F,t}$ and $w^*_{P,t}$ and can only be ascertained from the reduced form. Likewise, a $k_F \times 1$ vector $\boldsymbol{\beta}_F$ and $k_P \times 1$ vector $\boldsymbol{\beta}_P$ measure conditional effects.

Lastly, an error term $\epsilon_{Y,t}$ represents the effect of unobservable variables on $w^*_{Y,t}$. Error terms $\epsilon_{F,t}$ and $\epsilon_{P,t}$ have similar meaning. We assume that the vector $(\epsilon_{Y,t}, \epsilon_{F,t}, \epsilon_{P,t})'$ is i.i.d. across t conditional on $\mathbf{x}_t$ as a mean zero normal random vector:

$$(\epsilon_{Y,t}, \epsilon_{F,t}, \epsilon_{P,t})' \,|\mathbf{x}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}), \tag{2}$$

where $\mathbf{\Sigma}$ is the covariance matrix. Notice that for a given t, we allow $\epsilon_{Y,t}$, $\epsilon_{F,t}$, and $\epsilon_{P,t}$ to be correlated with each other.

The system (1) may be interpreted as an approximation to a demand system that incorporates (possibly, asymmetric) social interactions between Y and F and between Y and P. Several behavioral mechanisms may underlie these interactions. For example, since Y considers F to be his or her best friend, they may share many experiences and perceptions (good or bad) with each other. The shared perceptions about foods would include preferences for fruits and vegetables. Also, Y and F may occasionally eat together, in which case the interdependence of their food consumption behaviors may result from one of them "mimicking" the other. Similar mechanisms may underlie the endogenous effects between Y and P, because Y and P are likely to communicate on a regular basis and sometimes eat together.[22] In contrast, since the extent of exposure of F and P to each other's food choices is limited, we rule out endogenous effects between them. In fact, less than 30% of parents in our sample report that they know the youth's friends very well, in the first place, let alone what specific foods the friends eat.

In addition, the model incorporates conditional effects on fruit consumption by an individual of his or her own age in the cases of Y, F, and P,[23] of sex in the cases of Y and F, and of race in the case of F. Allowing for one's own age, sex, and race effects is in line with prior research on correlates of food intake (e.g., Videon and Manning, 2003; Stewart and Menning, 2009), but we are unable to include a full range of them. Since Y and F are always of the same sex by the FACHS Wave 4 design, the effects of their sexes are not separately identifiable ($\mathbf{x}_{Y,t}$ and $\mathbf{x}_{F,t}$ contain the same variable *Y_male*). Also, we do not include indicators for the race of Y and race or sex of P, because few youths in the sample are not African American and few parents are not African American or are male and hence the corresponding effects would be difficult to identify.

The specification does not allow for conditional effects on fruit intake by an individual of demographic characteristics of other members of his or her triplet, which amounts to ruling out the corresponding contextual effects. We acknowledge that the restriction is a limitation of our analysis. However, inability to fully account for contextual effects is a common feature of the current empirical literature on social interactions.[24] Also, the specification allows for a conditional effect of the relative fruit price faced by an individual himself or herself, but not for conditional effects of prices faced by others, as the price variables are specific to place of residence and interview date.

However, completely ruling out contextual effects makes little sense in the case of parental education, marital status, and poverty. More educated parents may have better knowledge

[22] It may be that the diets of many youths in our sample are, to a large degree, determined by their parents' food purchasing decisions. It is also possible that a youth's food preferences affect his or her parent's preferences and vice versa.

[23] We include second-order polynomials in age to account for possibly nonlinear effects. The robustness of the results to the exclusion of quadratic age terms is discussed in Appendix C.

[24] It is difficult to imagine why, for example, the friend's age would (additionally) affect the youth's own intake of fruit when the endogenous effect of the friend's intake on the youth has already been taken into account. In fact, it is more reasonable to hypothesize that the *strength* of the endogenous effect itself would vary depending on how "close" demographically the youth and friend are. However, allowing for such heterogeneity of the endogenous effects is beyond the scope of our paper. We defer a comprehensive analysis of this interesting possibility to future research.

of the benefits of fruit consumption and may communicate this knowledge to their children. Thus, education of P may affect food preferences of Y. In addition, since more educated parents tend to have higher incomes, P's education may affect Y's budget constraint if there are intra-family transfers. Marital status and poverty may have similar effects. Thus, we include the corresponding indicators in both $\mathbf{x}_{Y,t}$ and $\mathbf{x}_{P,t}$, while still ruling out their contextual effect on F.

3.3 Identification

Since the dependent variables $w^*_{Y,t}$, $w^*_{F,t}$, and $w^*_{P,t}$ are observed only ordinally, variances of the errors $\epsilon_{Y,t}$, $\epsilon_{F,t}$, and $\epsilon_{P,t}$ are impossible to identify (Maddala, 1983, p. 47). Therefore, we specify the covariance matrix of the errors as

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \rho_{YF} & \rho_{YP} \\ \rho_{YF} & 1 & \rho_{FP} \\ \rho_{YP} & \rho_{FP} & 1 \end{pmatrix}, \tag{3}$$

where each diagonal entry is normalized to one and parameters ρ_{YF}, ρ_{YP}, and ρ_{FP} are correlation coefficients to estimate.

We must also impose a normalization on the thresholds:

$$\alpha_Y(2) = \alpha_F(2) = \alpha_P(2) = 0, \tag{4}$$

which leaves a total of nine thresholds, $\{\alpha_Y(j), \alpha_F(j), \alpha_P(j)\}_{j=3}^{5}$, to estimate.

Given the normalizations, the identification approach is analogous to a textbook approach for a system of linear equations (for an explanation of why the analogy holds, see Maddala and Lee, 1976, pp. 531-533). Thus, we can employ the conventional order and rank conditions (Greene, 2008, pp. 368-369). First, we show that the number of variables comprising $\mathbf{x}_t$ that are excluded from each equation of the system (1) is at least as large as the number of dependent variables included on the equation's right-hand side. In the first equation, the number of the excluded explanatory variables is $k - k_Y = 15 - 8 = 7$, while the number of the included dependent variables on the right-hand side is 2 (namely, $w^*_{F,t}$ and $w^*_{P,t}$). Thus, the order condition holds. Similarly, it holds for the second and third equations, because $k - k_F = 9 > 1$ and $k - k_P = 8 > 1$, respectively.

Second, to explain why the rank condition also holds, we rewrite the system (1) in a matrix form:

$$\left(w^*_{Y,t}, w^*_{F,t}, w^*_{P,t}\right) \cdot \mathbf{\Gamma} + \mathbf{x}'_t \cdot \mathbf{B} = \left(\epsilon_{Y,t}, \epsilon_{F,t}, \epsilon_{P,t}\right), \tag{5}$$

where a $k \times 3$ matrix $\mathbf{B}$ consists of zeros and elements of the vectors $-\boldsymbol{\beta}_Y$, $-\boldsymbol{\beta}_F$, and $-\boldsymbol{\beta}_P$ that are arranged according to the layout in Table 4 and a 3×3 matrix $\mathbf{\Gamma}$ is

$$\mathbf{\Gamma} = \begin{pmatrix} 1 & -\gamma_{YF} & -\gamma_{YP} \\ -\gamma_{FY} & 1 & 0 \\ -\gamma_{PY} & 0 & 1 \end{pmatrix}.$$

Now, consider a $(3+k) \times 3$ matrix $\mathbf{A}$ in which $\mathbf{B}$ is stacked below $\boldsymbol{\Gamma}$: $\mathbf{A} = \begin{bmatrix} \boldsymbol{\Gamma} \\ \mathbf{B} \end{bmatrix}$. Let $\mathbf{a}_1$ be the first column of $\mathbf{A}$ and $\mathbf{A}_1$ be a submatrix consisting of the other columns so that $\mathbf{A} = [\mathbf{a}_1, \mathbf{A}_1]$. Define $\tilde{\mathbf{A}}_1$ as a submatrix of $\mathbf{A}_1$ consisting only of rows of $\mathbf{A}_1$ corresponding to rows of $\mathbf{a}_1$ with zero entries. Formally, the rank condition holds for the first equation of the system (1) if rank $\left(\tilde{\mathbf{A}}_1\right) = 2$. Given the exclusion restrictions embedded in $\mathbf{x}_{Y,t}$, it is straightforward to verify that this equality holds. Showing that the rank condition also holds for the second and third equations is analogous.

3.4 Estimation Strategy

Given the assumption that $\gamma_{PY} \cdot \gamma_{YP} + \gamma_{FY} \cdot \gamma_{YF} \neq 1$, matrix $\boldsymbol{\Gamma}$ is nonsingular, and we can solve for the reduced form of the system (5) as

$$\left(w^*_{Y,t}, w^*_{F,t}, w^*_{P,t}\right) = \mathbf{x}'_t \cdot \boldsymbol{\Pi} + \left(v_{Y,t}, v_{F,t}, v_{P,t}\right), \tag{6}$$

where a $k \times 3$ matrix $\boldsymbol{\Pi} = -\mathbf{B} \cdot \boldsymbol{\Gamma}^{-1}$ and a 1×3 vector of the reduced form errors $\left(v_{Y,t}, v_{F,t}, v_{P,t}\right) = \left(\epsilon_{Y,t}, \epsilon_{F,t}, \epsilon_{P,t}\right) \cdot \boldsymbol{\Gamma}^{-1}$ is i.i.d. across t conditional on $\mathbf{x}_t$ as a mean zero normal random vector:

$$\left(v_{Y,t}, v_{F,t}, v_{P,t}\right)' | \mathbf{x}_t \sim N\left(\mathbf{0}, \boldsymbol{\Omega}\right), \tag{7}$$

where $\boldsymbol{\Omega} = \left(\boldsymbol{\Gamma}^{-1}\right)' \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{\Gamma}^{-1}$ is the covariance matrix.

The likelihood contribution of a triplet is the probability of observing actual answers about fruit consumption by the three triplet members. We derive this probability using the reduced form (6) and joint distribution (7) of the reduced form errors. Let $\boldsymbol{\theta}$ be a vector of all identifiable parameters of the model:

$$\boldsymbol{\theta} = \left(\left\{\alpha_Y(j), \alpha_F(j), \alpha_P(j)\right\}_{j=3}^{5}, \rho_{YF}, \rho_{YP}, \rho_{FP}, \gamma_{FY}, \gamma_{PY}, \gamma_{YF}, \gamma_{YP}, \boldsymbol{\beta}'_Y, \boldsymbol{\beta}'_F, \boldsymbol{\beta}'_P\right)'.$$

The parameters of the reduced form, matrices $\boldsymbol{\Pi}$ and $\boldsymbol{\Omega}$, are known functions of $\boldsymbol{\theta}$. For convenience, we partition $\boldsymbol{\Pi}$ as $\boldsymbol{\Pi} = [\boldsymbol{\pi}_Y, \boldsymbol{\pi}_F, \boldsymbol{\pi}_P]$, where the three $k \times 1$ vectors $\boldsymbol{\pi}_Y$, $\boldsymbol{\pi}_F$, and $\boldsymbol{\pi}_P$ are also known functions of $\boldsymbol{\theta}$. Then, the likelihood contribution of triplet t is

$$\begin{aligned} L_t(\boldsymbol{\theta}) \equiv L\left(w_{Y,t}, w_{F,t}, w_{P,t} | \mathbf{x}_t; \boldsymbol{\theta}\right) = \Pr\big[\alpha_Y\left(w_{Y,t}\right) < w^*_{Y,t} \leq \alpha_Y\left(w_{Y,t}+1\right), \\ \alpha_F\left(w_{F,t}\right) < w^*_{F,t} \leq \alpha_F\left(w_{F,t}+1\right), \alpha_P\left(w_{P,t}\right) < w^*_{P,t} \leq \alpha_P\left(w_{P,t}+1\right) | \mathbf{x}_t; \boldsymbol{\theta}\big] = \\ = \Pr\left[\alpha_Y\left(w_{Y,t}\right) - \mathbf{x}'_t \cdot \boldsymbol{\pi}_Y < v_{Y,t} \leq \alpha_Y\left(w_{Y,t}+1\right) - \mathbf{x}'_t \cdot \boldsymbol{\pi}_Y,\right. \\ \alpha_F\left(w_{F,t}\right) - \mathbf{x}'_t \cdot \boldsymbol{\pi}_F < v_{F,t} \leq \alpha_F\left(w_{F,t}+1\right) - \mathbf{x}'_t \cdot \boldsymbol{\pi}_F, \\ \left.\alpha_P\left(w_{P,t}\right) - \mathbf{x}'_t \cdot \boldsymbol{\pi}_P < v_{P,t} \leq \alpha_P\left(w_{P,t}+1\right) - \mathbf{x}'_t \cdot \boldsymbol{\pi}_P | \mathbf{x}_t; \boldsymbol{\theta}\right] = \\ = \int\limits_{\alpha_Y\left(w_{Y,t}\right)-\mathbf{x}'_t\boldsymbol{\pi}_Y}^{\alpha_Y\left(w_{Y,t}+1\right)-\mathbf{x}'_t\boldsymbol{\pi}_Y} \int\limits_{\alpha_F\left(w_{F,t}\right)-\mathbf{x}'_t\boldsymbol{\pi}_F}^{\alpha_F\left(w_{F,t}+1\right)-\mathbf{x}'_t\boldsymbol{\pi}_F} \int\limits_{\alpha_P\left(w_{P,t}\right)-\mathbf{x}'_t\boldsymbol{\pi}_P}^{\alpha_P\left(w_{P,t}+1\right)-\mathbf{x}'_t\boldsymbol{\pi}_P} f\left(v_{Y,t}, v_{F,t}, v_{P,t} | \mathbf{x}_t; \boldsymbol{\theta}\right) dv_{P,t} dv_{F,t} dv_{Y,t}, \end{aligned} \tag{8}$$

where $f\left(v_{Y,t}, v_{F,t}, v_{P,t} | \mathbf{x}_t; \boldsymbol{\theta}\right)$ is a trivariate normal density function, as implied by (7).

Computation of $L_t(\boldsymbol{\theta})$ in (8) requires evaluation of a trivariate normal rectangle probability. This evaluation problem was extensively studied in the literature, and numerical algorithms are available (see Genz, 2004).

Given the assumption that $(w_{Y,t}, w_{F,t}, w_{P,t}, \mathbf{x}_t')'$ is i.i.d. across t, we can obtain an estimate of $\boldsymbol{\theta}$ by the maximum likelihood method as

$$\widehat{\boldsymbol{\theta}}_{MLE} = \arg\max_{\boldsymbol{\theta}} \sum_{t=1}^{T} \ln L_t(\boldsymbol{\theta})$$

and conduct statistical inference using standard techniques (e.g., Greene, 2008, Ch. 16).

4 Results

4.1 Fruit Consumption

The estimated model of fruit consumption is presented in Table 5. In Panel A, we list estimates of the thresholds (there are three identifiable thresholds for each triplet member). Panel B provides estimates of the endogenous effects. Panel C contains estimates of conditional effects of the explanatory variables. Estimates of the correlations among the error terms are given in the notes to the table.[25]

For convenience, we arrange the estimates in columns corresponding to the three equations of the system (1). Missing entries in Panel C indicate exclusion restrictions (in line with the layout in Table 4). Note that while the thresholds are an essential component of the econometric model and Panel A shows that they are precisely estimated, we are primarily interested in coefficients in Panels B and C.

Panel B reveals the presence of endogenous effects in the consumption of fruit. More specifically, we estimate a statistically significant positive impact of the parent's consumption on the youth's consumption ($\widehat{\gamma}_{PY} = 0.620$) and of the youth's consumption on the parent's consumption ($\widehat{\gamma}_{YP} = 0.382$). We do not detect the presence of endogenous effects between the youth and his or her best friend at a conventional significance level. The latter finding stands in contrast to a well-established result that peer effects play a key role in facilitating young people's substance use (e.g., Gaviria and Raphael, 2001; Powell et al., 2005; Krauth, 2006; Lundborg, 2006; Clark and Lohéac, 2007; Fletcher, 2010). Perhaps eating habits are formed mainly through observing and "mimicking" dietary choices of parents, while risky health behaviors are primarily learned from peers.

Panel C shows estimated conditional effects of the demographic and socioeconomic characteristics and of relative fruit prices. We infer that (given fixed consumption of fruit by the friend and parent) the youth's consumption of fruit declines with the youth's age, as the coefficient on the quadratic age term is $-0.184 \cdot 10^{-2}$ (the coefficient on the linear term

[25] In order to ensure that the positive definiteness of the normalized matrix of the errors and the constraints imposed on the thresholds were true, we reparameterized the model prior to estimation. All estimates were obtained by numerically maximizing the sample log-likelihood function, and standard errors were computed using outer products of numerical gradients of the log-likelihood contributions (Berndt et al., 1974). We then recovered estimates of the original parameters and computed corresponding standard errors by the delta method (Greene, 2008, pp. 1055-1056).

Table 5: Estimated model of fruit consumption

	Youth: $w^*_{Y,t}$ Coeff.	Std. Err.		Friend: $w^*_{F,t}$ Coeff.	Std. Err.		Parent: $w^*_{P,t}$ Coeff.	Std. Err.
Panel A: Thresholds								
$\widehat{\alpha}_Y\,(3)$	0.753**	(0.101)	$\widehat{\alpha}_F\,(3)$	0.764**	(0.074)	$\widehat{\alpha}_P\,(3)$	0.861**	(0.078)
$\widehat{\alpha}_Y\,(4)$	1.442**	(0.172)	$\widehat{\alpha}_F\,(4)$	1.546**	(0.111)	$\widehat{\alpha}_P\,(4)$	1.946**	(0.109)
$\widehat{\alpha}_Y\,(5)$	1.766**	(0.207)	$\widehat{\alpha}_F\,(5)$	1.801**	(0.124)	$\widehat{\alpha}_P\,(5)$	2.181**	(0.118)
Panel B: Endogenous Effects								
$\widehat{\gamma}_{FY}$	0.285	(0.182)	$\widehat{\gamma}_{YF}$	−0.251	(0.243)	$\widehat{\gamma}_{YP}$	0.382**	(0.192)
$\widehat{\gamma}_{PY}$	0.620**	(0.142)						
Panel C: Effects of Explanatory Variables								
constant	1.658**	(0.592)		1.584**	(0.442)		0.495	(0.345)
Y_age $\times 10^{-1}$	−0.400	(0.358)						
Y_age2 $\times 10^{-2}$	−0.184**	(0.051)						
Y_male[a]	0.038	(0.071)		0.084	(0.100)			
F_age $\times 10^{-1}$				−0.028	(0.190)			
F_age2 $\times 10^{-2}$				−0.029	(0.027)			
F_black				0.186	(0.135)			
P_age $\times 10^{-1}$							0.229**	(0.049)
P_age2 $\times 10^{-2}$							−0.015**	(0.001)
P_higher_educ[b]	0.134	(0.107)					0.026	(0.103)
P_married	−0.195*	(0.105)					0.216*	(0.107)
P_poverty	−0.050	(0.114)					0.152	(0.112)
Y_rel_price	−0.594	(0.486)						
F_rel_price				−0.717*	(0.416)			
P_rel_price							−1.012*	(0.548)

Notes:
[a] Y and F are always of the same sex by the study design.
[b] The omitted education category comprises P's with a high school degree or less.
* and ** denote significance at 10% and 5% levels, respectively.
The estimated correlations (std. errors) are $\widehat{\rho}_{YF} = -0.117$ (0.155), $\widehat{\rho}_{YP} = -0.725^{**}$ (0.084), and $\widehat{\rho}_{FP} = 0.009$ (0.074).
The sample log-likelihood is −2244.13.

is not statistically significant). Age also affects the parent's fruit intake. Specifically, the parent's consumption appears to increase with age at a decreasing rate, as the coefficients on the linear and quadratic age terms are $0.229 \cdot 10^{-1}$ and $-0.015 \cdot 10^{-2}$, respectively. It is worth noting that the discovered effects may not be the effects of age per se but may rather be cohort effects reflecting different attitudes of younger and older generations toward fruit consumption. In a cross-sectional setting such as the one in this paper, cohort and age effects are not separately identifiable.

In addition, we find a negative conditional effect of the parent's being married on the youth's consumption (-0.195) and a positive effect of being married on the parent's own consumption (0.216). The negative effect of the parent's being married on the youth's consumption of fruit is not easy to explain. All else equal, married couples are likely to confer more resources on their children than are single parents, which includes providing more access to all foods. Perhaps the magnitude of the relationship between non-fruit food consumption and family wealth is much larger than that between fruit intake and wealth among the youths in our sample. As a result, the resource-induced increase in non-fruit consumption may "crowd out" fruit intake, which would explain the negative coefficient. Other demographic and socioeconomic characteristics do not have a statistically significant effect.

In line with intuition, the coefficients on the relative prices are negative, but the corresponding effects are fairly weak, since the coefficients are only marginally significant at a 10% level in the cases of the friend (-0.717) and parent (-1.012) and not statistically significant in the case of the youth.[26]

Lastly, we find a statistically significant correlation between the errors $\epsilon_{Y,t}$ and $\epsilon_{P,t}$ (-0.725) indicating the presence of the correlated effect between the youth and parent. An analysis of the robustness of the results is presented in Appendix C. It shows that our main results remain qualitatively the same across different specifications of the model.

4.2 Vegetable Consumption

The estimated model of vegetable consumption is presented in Table 6, which follows the layout of Table 5. Again, while Panel A shows that the identifiable thresholds are precisely estimated, our primary interest lies in coefficients in Panels B and C.

Panel B reveals the existence of a positive endogenous effect of the parent's consumption of vegetables on youth's consumption ($\widehat{\gamma}_{PY} = 0.586$). We do not detect an impact in the reverse direction, as the estimate of γ_{YP} is not statistically significant, indicating the asymmetry of social interactions. Similarly to the case of fruit consumption, we do not find statistically significant endogenous effects between the youth and friend, which again suggests that eating behaviors are primarily learned from one's parents rather than peers.

Panel C shows that the conditional effects of age are nonlinear. We infer that the youth's intake of vegetables declines until the youth is approximately 20 years old (given fixed intakes of vegetables by the parent and friend) but increases thereafter, as the coefficients on the linear and quadratic age terms are $-1.256 \cdot 10^{-1}$ and $0.319 \cdot 10^{-2}$, respectively. The finding that the youth's consumption decreases in the late teens is broadly consistent with Stewart

[26] Because the consumption levels are not observed, we cannot compute price elasticities.

Table 6: Estimated model of vegetable consumption

	Youth: $w^*_{Y,t}$			Friend: $w^*_{F,t}$			Parent: $w^*_{P,t}$	
	Coeff.	Std. Err.		Coeff.	Std. Err.		Coeff.	Std. Err.
Panel A: Thresholds								
$\widehat{\alpha}_Y\,(3)$	0.672**	(0.134)	$\widehat{\alpha}_F\,(3)$	0.811**	(0.087)	$\widehat{\alpha}_P\,(3)$	1.140**	(0.134)
$\widehat{\alpha}_Y\,(4)$	1.489**	(0.286)	$\widehat{\alpha}_F\,(4)$	1.760**	(0.149)	$\widehat{\alpha}_P\,(4)$	2.286**	(0.214)
$\widehat{\alpha}_Y\,(5)$	1.737**	(0.331)	$\widehat{\alpha}_F\,(5)$	2.044**	(0.166)	$\widehat{\alpha}_P\,(5)$	2.547**	(0.231)
Panel B: Endogenous Effects								
$\widehat{\gamma}_{FY}$	−0.351	(0.273)	$\widehat{\gamma}_{YF}$	−0.168	(0.384)	$\widehat{\gamma}_{YP}$	−0.234	(0.386)
$\widehat{\gamma}_{PY}$	0.586**	(0.250)						
Panel C: Effects of Explanatory Variables								
constant	2.147**	(1.090)		1.204**	(0.562)		0.271	(0.413)
Y_age $\times 10^{-1}$	−1.256**	(0.057)						
Y_age2 $\times 10^{-2}$	0.319**	(0.075)						
Y_male[a]	−0.006	(0.099)		0.071	(0.095)			
F_age $\times 10^{-1}$				0.328*	(0.177)			
F_age2 $\times 10^{-2}$				−0.037**	(0.018)			
F_black				0.123	(0.119)			
P_age $\times 10^{-1}$							0.524**	(0.059)
P_age2 $\times 10^{-2}$							−0.045**	(0.002)
P_higher_educ[b]	0.052	(0.093)					0.012	(0.105)
P_married	−0.039	(0.128)					0.320**	(0.110)
P_poverty	0.006	(0.094)					−0.010	(0.122)
Y_rel_price	−1.559*	(0.902)						
F_rel_price				−1.352[c]	(0.839)			
P_rel_price							0.485	(0.672)

Notes:
[a] Y and F are always of the same sex by the study design.
[b] The omitted education category comprises P's with a high school degree or less.
[c] Significant at 11% level. * and ** denote significance at 10% and 5% levels, respectively.
The estimated correlations (std. errors) are $\widehat{\rho}_{YF} = 0.467^{**}$ (0.235), $\widehat{\rho}_{YP} = -0.306$ (0.207), and $\widehat{\rho}_{FP} = 0.107$ (0.132).
The sample log-likelihood is −2152.12.

and Menning's (2009) result that adolescents' propensity to eat vegetables declines with age in Wave 2 of the National Longitudinal Study of Adolescent Health (Add Health).[27] Also, we infer that the consumption of the friend increases with age at a decreasing rate, as the coefficients on the linear and quadratic age terms are $0.328 \cdot 10^{-1}$ and $-0.037 \cdot 10^{-2}$, respectively. Similarly, the parent's consumption increases with age at a decreasing rate (the coefficients on the linear and quadratic age terms are $0.524 \cdot 10^{-1}$ and $-0.045 \cdot 10^{-2}$, respectively). Analogously to the case of fruit consumption, these effects may reflect cohort effects rather than the effects of age per se.

In addition, we find that the conditional effect of the parent's being married on his or her own intake of vegetables is positive and statistically significant (0.320), but the effect on the youth's consumption is negative and not significant. It is worth noting that Stewart and Menning (2009) estimate that adolescents from two-parent households have a higher propensity to eat vegetables, which is consistent with our finding that the reduced-form impact of the parent's being married on the youth's vegetable intake (not reported in Table 6) is positive.[28] Other demographic and socioeconomic characteristics do not tend to exert a significant impact.

In line with intuition, we estimate negative conditional effects of the relative vegetable prices faced by the youth (−1.559) and friend (−1.352). Both effects are fairly weak, since they are only marginally statistically significant (at 10% in the case of the youth and 11% in the case of the friend). The estimate on the relative price faced by the parent is not statistically significant.

Lastly, we find a statistically significant correlation between the errors $\epsilon_{Y,t}$ and $\epsilon_{F,t}$ (0.467), indicating the presence of a correlated effect between the youth and friend. An analysis of the robustness of the results (see Appendix C) shows that our main results remain qualitatively unchanged across different specifications of the model.

4.3 Implications for Policy Design

The results have several implications for designing policy interventions to facilitate healthy eating by African American youths. Importantly, since our sample is comparable to nationally representative samples in terms of food consumption frequencies (see Appendix A) and basic demographic characteristics (Appendix B), the conclusions may apply not only to African American youths in Georgia and Iowa, but also more broadly to the population of all African American youths in the U.S.[29]

First, we detected the presence of endogenous effects between the youth and parent in the consumption of fruits and vegetables. This result is in line with existing evidence that children's eating behaviors are affected by observing food-selection patterns of their parents (Cullen et al., 2001). Moreover, it suggests that the process of shaping eating

[27] Stewart and Menning only estimate a linear age effect and do not control for endogenous effects.

[28] The difference between the conditional effect of the parental marital status in our case and Stewart and Menning's estimate underscores the importance of distinguishing between conditional and reduced-form effects of explanatory variables in a model with social interactions.

[29] We do not believe that small differences between our sample and the CPS subsample in terms of the fraction of married parents and with respect to parental poverty status and educational attainment pose a problem, since these characteristics are controlled for in the empirical models.

behaviors persists beyond childhood years into the late teens. Most notably, the result indicates the existence of social multipliers within a family, suggesting that a health policy intervention focusing on increasing fruit and vegetable intake by parents of African American youths would simultaneously increase intake by the youths themselves, even when the youths are not a direct target of the intervention.[30] Hence, it may be cost-effective to design policy interventions that target only one member in an African American family such as, in particular, the mother, who is the most likely primary caregiver.

Second, we found little evidence for endogenous effects between the youth and his or her best friend. This result suggests that the family rather than peer groups should be the main focus of interventions aimed at increasing consumption of fruits and vegetables by African American youths.

Third, our estimates imply that relative prices may affect individual decisions about consumption of fruits and vegetables. However, the price effects tend to be statistically weak, since the estimates are only marginally significant. We find that the relative price of fruit is more important for the parent's consumption than for the youth's consumption. In contrast, the relative price of vegetables tends to affect the youth's consumption, but not the parent's consumption. Given our estimates of the endogenous effects, these results suggest that decreasing the relative price of fruit through subsidies or by taxing other foods may increase the intake of fruit by parents of African American youths and, because of the social multiplier effect, increase the intake among the youths themselves. In contrast, decreasing the relative price of vegetables may increase the intake by African American youths but is unlikely to have any spillover consumption effects.

5 Conclusion

In this paper, we analyze determinants of fruit and vegetable consumption by African American youths. We contribute to the literature by focusing on the role of social interactions and the relative prices of fruits and vegetables in the consumption of these foods by a youth, his or her parent, and best friend. The richness of the behavioral data in the FACHS allows us to distinguish between the impact of the parent's food intake and the effect of the friend's intake on the youth's own consumption. Comprehensive food price data from the ERS's QFAHPD, which includes detailed information on fruits and vegetables, enables estimation of the relative price effects. Knowledge of how various factors affect food consumption is crucial for designing policy interventions to facilitate healthy eating by young people in order to overcome the obesity epidemic in the U.S.

We construct and estimate a simultaneous equation ordered probit model, which, to the best of our knowledge, is novel in the literature on social effects. This methodological approach allows us to incorporate assumptions about the endogenous effects and helps to mitigate the limitations of the available food consumption data. We find statistically significant endogenous effects of the parent's consumption of fruits and vegetables on the youth's consumption, as well as of the youth's consumption of fruit on the parent's consumption. However, no statistically significant social interactions are detected between the youth and

[30] For an experiment implementing a parent-focused intervention aimed at reducing fat and sugar intake by children, see Epstein et al. (2001).

friend. We also find that the relative fruit and vegetable prices tend to negatively affect the intakes of these foods, but the estimated price effects are statistically weak.

The results imply that policy interventions aimed at increasing fruit and vegetable consumption by only one family member such as, in particular, the mother – the most likely primary caregiver – may be a cost-effective way to facilitate healthy eating by African Americans, since an increase in parental fruit and vegetable intake tends to raise the intake of these foods by adolescent children because of the social multiplier effect. The results also indicate that lowering the relative prices of fruits and vegetables by subsidizing these foods or by taxing other foods may be used to promote consumption of fruits and vegetables by African Americans.

Our analysis can be extended in several directions. We investigated fruit consumption and vegetable consumption separately from each other and in isolation from other health-related behaviors. This methodological choice was necessitated by the modest size of our sample and the difficulty of locating restrictions to identify a more inclusive model. As more and better data are collected, it may be interesting to estimate a richer model of health behaviors to offer more comprehensive policy recommendations than the ones provided in this paper. Also, we focused on African Americans, but they are not the only population group in the U.S. at risk for developing poor nutritional habits and gaining excess weight. Future research should address the role of social interactions and prices in healthy eating by other minority populations, as well as by the nation as a whole.

References

Auld, M. C., Powell, L. M., 2009. Economics of food energy density and adolescent body weight. Economica 76, 719–740.

Berndt, E. K., Hall, B. H., Hall, R. E., Hausman, J. A., 1974. Estimation and inference in nonlinear structural models. Annals of Economic and Social Measurement 3/4, 653–665.

Brock, W. A., Durlauf, S. N., 2001. Discrete choice with social interactions. Review of Economic Studies 68 (2), 235–260.

Cameron, A. C., Trivedi, P. K., 1986. Econometric models based on count data: Comparisons and applications of some estimators and tests. Journal of Applied Econometrics 1 (1), 29–53.

Cawley, J., 2004. An economic framework for understanding physical activity and eating behaviors. American Journal of Preventive Medicine 27 (3S), 117–125.

Centers for Disease Control and Prevention, 2010. State-specific trends in fruit and vegetable consumption among adults - United States, 2000-2009. Morbidity and Mortality Weekly Report 59 (35), 1125–1130.

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

Christakis, N. A., Fowler, J. H., 2007. The spread of obesity in a large social network over 32 years. New England Journal of Medicine 357 (4), 370–379.

Christeson, W., Taggart, A. D., Messner-Zidell, S., 2010. Too fat to fight: Retired military leaders want junk food out of America's schools. Report, Mission: Readiness, Military Leaders for Kids, Washington, DC, http://cdn.missionreadiness.org/MR_Too_Fat_to_Fight-1.pdf (accessed August 27, 2010).

Clark, A. E., Lohéac, Y., 2007. "It wasn't me, it was them!" Social influence in risky behavior by adolescents. Journal of Health Economics 26, 763–784.

Cullen, K. W., Baranowski, T., Rittenberry, L., Cosart, C., Hebert, D., de Moor, C., 2001. Child-reported family and peer influences on fruit, juice and vegetable consumption: Reliability and validity of measures. Health Education Research 16 (2), 187–200.

de Castro, J. M., Brewer, E. M., 1991. The amount eaten in meals by humans is a power function of the number of people present. Physiology and Behavior 51, 121–125.

DeNavas-Walt, C., Proctor, B. D., Smith, J., 2007. U.S. Census Bureau. Current Population Reports, P60-233. Income, Poverty, and Health Insurance Coverage in the United States: 2006. GPO, Washington, DC.

Dietary Guidelines Advisory Committee, 2010. Report of the Dietary Guidelines Advisory Committee on the Dietary Guidelines for Americans, 2010. Tech. rep., U.S. Department of Agriculture, U.S. Department of Health and Human Services, http://www.cnpp.usda.gov/DGAs2010-DGACReport.htm (accessed August 27, 2010).

Dong, D., Lin, B.-H., 2009. Fruit and vegetable consumption by low-income Americans: Would a price reduction make a difference? Economic Research Report 70, U.S. Dept. of Agriculture, Economic Research Service, http://www.ers.usda.gov/publications/err70/err70.pdf (accessed August 27, 2010).

Epstein, L. H., Gordy, C. C., Raynor, H. A., Beddome, M., Kilanowski, C. K., Paluch, R., 2001. Increasing fruit and vegetable intake and decreasing fat and sugar intake in families at risk for childhood obesity. Obesity Research 9 (3), 171–178.

Finkelstein, E. A., Trogdon, J. G., Cohen, J. W., Dietz, W., 2009. Annual medical spending attributable to obesity: Payer- and service-specific estimates. Health Affairs 28 (5), w822–w831.

Fletcher, J. M., 2010. Social interactions and smoking: Evidence using multiple student cohorts, instrumental variables, and school fixed effects. Health Economics 19, 466–484.

Freedman, D. S., Wang, J., Thornton, J. C., Mei, Z., Pierson, Jr., R. N., Dietz, W. H., Horlick, M., 2008. Racial/ethnic differences in body fatness among children and adolescents. Obesity 16 (5), 1105–1111.

Gaviria, A., Raphael, S., 2001. School-based peer effects and juvenile behavior. Review of Economics and Statistics 83 (2), 257–268.

Genz, A., 2004. Numerical computation of rectangular bivariate and trivariate normal and t probabilities. Statistics and Computing 14 (3), 251–260.

Gibbons, F. X., Gerrard, M., Cleveland, M. J., Wills, T. A., Brody, G., 2004. Perceived discrimination and substance use in African American parents and their children: A panel study. Journal of Personality and Social Psychology 86 (4), 517–529.

Granberg, E. M., Simons, L. G., Simons, R. L., 2009. Body size and social self-image among adolescent African American girls: The moderating influence of family racial socialization. Youth and Society 41 (2), 256–277.

Greene, W. H., 2008. Econometric Analysis, 6th Edition. Pearson Prentice Hall, Upper Saddle River, NJ.

Harris, J. E., López-Valcárcel, B. G., 2008. Asymmetric peer effects in the analysis of cigarette smoking among young people in the United States, 1992-1999. Journal of Health Economics 27, 249–264.

He, F. J., Nowson, C. A., Lucas, M., MacGregor, G. A., 2007. Increased consumption of fruit and vegetables is related to a reduced risk of coronary heart disease: Meta-analysis of cohort studies. Journal of Human Hypertension 21 (9), 717–728.

He, F. J., Nowson, C. A., MacGregor, G. A., 2006. Fruit and vegetable consumption and stroke: Meta-analysis of cohort studies. Lancet 367 (9507), 320–326.

Krauth, B. V., 2006. Simulation-based estimation of peer effects. Journal of Econometrics 133 (1), 243–271.

Lundborg, P., 2006. Having the wrong friends? Peer effects in adolescent substance use. Journal of Health Economics 25, 214–233.

Maddala, G. S., 1983. Limited-Dependent and Qualitative Variables in Econometrics. Cambridge University Press, Cambridge, UK.

Maddala, G. S., Lee, L., 1976. Recursive models with qualitative endogenous variables. Annals of Economic and Social Measurement 5 (4), 525–545.

Manski, C. F., 1993. Identification of endogenous social effects: The reflection problem. Review of Economic Studies 60 (3), 531–542.

Moffitt, R. A., 2001. Policy interventions, low-level equilibria, and social interactions. In: Durlauf, S. N., Young, H. P. (Eds.), Social Dynamics. MIT Press, Cambridge, MA, pp. 45–82.

Neumark-Sztainer, D., Story, M., Perry, C., Casey, M. A., 1999. Factors influencing food choices of adolescents: Findings from focus-group discussions with adolescents. Journal of the American Dietetic Association 99, 929–934, 937.

Ogden, C. L., Carroll, M. D., Curtin, L. R., Lamb, M. M., Flegal, K. M., 2010. Prevalence of high body mass index in US children and adolescents, 2007-2008. Journal of the American Medical Association 303 (3), 242–249.

Olshansky, S. J., Passaro, D. J., Hershow, R. C., Layden, J., Carnes, B. A., Brody, J., Hayflick, L., Butler, R. N., Allison, D. B., Ludwig, D. S., 2005. A potential decline in life expectancy in the United States in the 21st century. New England Journal of Medicine 352 (11), 1138–1145.

Pomery, E. A., Gibbons, F. X., Gerrard, M., Cleveland, M. J., Brody, G. H., Wills, T. A., 2005. Families and risk: Prospective analyses of familial and social influences on adolescent substance use. Journal of Family Psychology 19 (4), 560–570.

Powell, L. M., 2009. Fast food costs and adolescent body mass index: Evidence from panel data. Journal of Health Economics 28 (5), 963–970.

Powell, L. M., Tauras, J. A., Ross, H., 2005. The importance of peer effects, cigarette prices and tobacco control policies for youth smoking behavior. Journal of Health Economics 24, 950–968.

Renna, F., Grafova, I. B., Thakur, N., 2008. The effect of friends on adolescent body weight. Economics and Human Biology 6, 377–387.

Soetevent, A. R., 2006. Empirics of the identification of social interactions; an evaluation of the approaches and their results. Journal of Economic Surveys 20 (2), 193–228.

Stewart, S. D., Menning, C. L., 2009. Family structure, nonresident father involvement, and adolescent eating patterns. Journal of Adolescent Health 45, 193–201.

Subar, A. F., Dodd, K. W., Guenther, P. M., Kipnis, V., Midthune, D., McDowell, M., Tooze, J. A., Freedman, L. S., Krebs-Smith, S. M., 2006. The food propensity questionnaire: Concept, development, and validation for use as a covariate in a model to estimate usual food intake. Journal of the American Dietetic Association 106, 1556–1563.

Task Force on Childhood Obesity, 2010. Solving the problem of childhood obesity within a generation. Report to the President, White House Task Force on Childhood Obesity, http://www.letsmove.gov/pdf/TaskForce_on_Childhood_Obesity_May2010_FullReport.pdf (accessed August 27, 2010).

Todd, J. E., Mancino, L., Leibtag, E., Tripodo, C., 2010. Methodology behind the Quarterly Food-at-Home Price Database. Technical Bulletin 1926, U.S. Dept. of Agriculture, Economic Research Service, http://www.ers.usda.gov/Publications/TB1926/TB1926.pdf (accessed August 27, 2010).

Trogdon, J. G., Nonnemaker, J., Pais, J., 2008. Peer effects in adolescent overweight. Journal of Health Economics 27, 1388–1399.

Unnevehr, L., Eales, J., Jensen, H., Lusk, J., McCluskey, J., Kinsey, J., 2010. Food and consumer economics. American Journal of Agricultural Economics 92 (2), 506–521.

U.S. Department of Health and Human Services and U.S. Department of Agriculture, 2005. Dietary Guidelines for Americans, 2005, 6th Edition. GPO, Washington, DC.

Van Duyn, M. A. S., Pivonka, E., 2000. Overview of the health benefits of fruit and vegetable consumption for the dietetics professional: Selected literature. Journal of the American Dietetic Association 100, 1511–1521.

Videon, T. M., Manning, C. K., 2003. Influences on adolescent eating patterns: The importance of family meals. Journal of Adolescent Health 32, 365–373.

World Cancer Research Fund, 2007. Food, nutrition, physical activity, and the prevention of cancer: A global perspective. Tech. rep., American Institute for Cancer Research, Washington, DC, http://www.dietandcancerreport.org/ (accessed August 27, 2010).

Appendix A Comparison to NHANES

The NHANES is a continuous program of cross-sectional studies conducted by the National Center for Health Statistics of the CDC to assess the health and nutritional status of the U.S. civilian noninstitutional population. Each year, the survey covers a nationally representative probability sample of about 5,000 adults and children. Public use data are released biannually. Starting with the NHANES 2003-2004, respondents aged two years and older who have completed a 24-hour dietary recall interview are requested to additionally fill in a food frequency questionnaire (FFQ).[31] The FFQ is administered to ascertain information on food consumption in the past year. Details on the development of the FFQ are provided by Subar et al. (2006).

We employ the FFQ in the NHANES 2005-2006 to assess whether the fruit and vegetable consumption patterns of the FACHS Wave 4 participants are in line with the NHANES respondents' consumption habits, which are representative of the habits of the U.S. population. Since the underlying food frequency questions in the FACHS and NHANES are phrased differently, we can only provide a qualitative comparison of the patterns rather than perform a formal test of whether they are identical.

To obtain the fruit consumption patterns in the NHANES, we use raw answers on drinking various fruit juices and eating various fruits (a total of 15 distinct answers). Records with missing answers are dropped. The fruit juice drinking responses are recorded in ten separate categories (from "never" to "6 or more times per day" in the past year), while fruit eating answers comprise eleven categories (from "never" to "2 or more times per day" in the past year). We convert each answer into a weekly frequency using the midpoint of a corresponding response range. For example, if a respondent ate apples "1-6 times per year," we convert this frequency to $\frac{1+6}{2} \cdot \frac{7}{365} \cong 0.067$ times per week. The conversion to the weekly frequency is done for comparability with the FACHS, in which the reference period is the last week before the interview. Next, we sum the imputed weekly frequencies across the questions. If this sum is less than seven, the NHANES respondent is deemed to consume fruit less frequently than once a day in a "typical" week in the past year. Otherwise, his or her fruit consumption frequency is once a day or more. Analogously, we obtain the vegetable consumption patterns from 21 distinct vegetable eating questions in the FFQ. Using only two broad frequency categories – "less than once a day" and "once a day or more" – rather than narrower categories may help reduce sensitivity of the comparison to the imputation error.

Table A1 presents the fruit and vegetable consumption patterns in the FACHS along with the imputed patterns in the NHANES 2005-2006.[32] Given the distribution of age and race in our FACHS sample (Table 2), we focus on two separate subsamples in the NHANES: (non-Hispanic) African Americans ages 17-21 and African Americans ages 30-69. All subsample frequencies in the NHANES are computed using the FFQ sample weights.

As can be seen in Table A1, among African Americans ages 17-21 in our FACHS sample,

[31] Prior to 2003, the NHANES regularly included food frequency questions, but they varied in terms of the food group specificity, reference period, and so forth.

[32] The FACHS frequency responses "none" and "less than once a day (1-6 times)" are grouped together as "less than once a day," while the responses "once a day," "8-12 times," and "twice a day (or more)" are grouped as "once a day or more."

Table A1: Fruit and vegetable consumption in FACHS and NHANES

	17-21 y.o. African Americans		30-69 y.o. African Americans	
	FACHS (Last week)	NHANES (Typical week)	FACHS (Last week)	NHANES (Typical week)
Panel A: Fruit consumption				
Less than once a day, %	39.35	33.01	34.85	39.01
Once a day or more, %	60.65	66.99	65.15	60.99
(Subsample size)	(826)	(173)	(462)	(430)
Panel B: Vegetable consumption				
Less than once a day, %	40.56	37.00	24.03	25.01
Once a day or more, %	59.44	63.00	75.97	74.99
(Subsample size)	(826)	(173)	(462)	(437)

Note:
Statistics for the NHANES subsamples are computed using the FFQ sample weights.

approximately 39% ate fruit or drank fruit juice less frequently than once a day in the last week before the interview and the remaining 61% consumed fruit once a day or more often. In comparison, among African Americans ages 17-21 in the NHANES, 33% consumed fruit less than once a day in a typical week in the past year, while 67% did so once a day or more often. It is easy to compute that the difference between the corresponding fractions in the FACHS and NHANES in relative terms is within 9.5% ($\cong \frac{6.34}{66.99} \cdot 100\%$) to 19% ($\cong \frac{6.34}{33.01} \cdot 100\%$). We see even smaller differences between the vegetable consumption patterns of African Americans ages 17-21 (6% to 10% in relative terms) and fruit consumption patterns of African Americans ages 30-69 (6% to 12% in relative terms). The difference between the vegetable consumption patterns of African Americans ages 30-69 is particularly small (1% to 4% in relative terms).

The comparison indicates that the difference between the fruit and vegetable consumption patterns in the FACHS Wave 4 and NHANES 2005-2006 is qualitatively small. Thus, the food consumption habits of the FACHS participants appear to be in line with the habits of the corresponding U.S. population. However, it is important to note that our comparison is only suggestive rather than definitive because it is not possible to perform a formal test of whether the consumption patterns in the FACHS and NHANES are identical.

Appendix B Comparison to CPS

The CPS is a monthly survey of about 50,000 households conducted by the Bureau of the Census. By design, it is representative of the U.S. civilian noninstitutional population. Although the main purpose of the CPS is to collect employment data, its notable secondary goal is to obtain demographic information. The Annual Social and Economic (ASEC) Supplement to the CPS, also known as the March Supplement, is administered every year to collect socioeconomic information beyond basic employment and demographic data.

To explore whether characteristics of the FACHS participants are similar to characteristics of the corresponding population in the U.S., we extracted a subsample from the 2006 ASEC Supplement data file by selecting all households containing an African American youth between the ages of 17 and 21 and at least one parent (in what follows, we refer to this subsample as "the CPS subsample").[33] The CPS subsample includes 1,053 households.

Table B1 presents summary statistics for selected demographic and socioeconomic characteristics of parents in the CPS subsample. To be consistent with the FACHS design, whenever a household in the CPS subsample contains the mother (either single or married) of the youth, we use her characteristics (rather than characteristics of the father) as characteristics of the parent. Otherwise, when a household does not contain the mother, we use information on the father. All statistics are computed using the ASEC Supplement weights. We also provide z-statistics and P-values for tests of equality between respective means in the FACHS and CPS.

As can be seen in Table B1, parents in the CPS subsample are between the ages of 30 and 80 and, on average, are 45 years old. The parent is male in only 5% of the cases.[34] Ninety three percent of the parents are African American. Thirteen percent of them have no high school degree, 39% have a high school degree, 31% have some college education (including an associate degree), and 17% have a bachelor's or higher degree. Forty three percent are married, and 22% live in poverty.

The tests of the equality between mean characteristics in the FACHS and CPS indicate no statistically significant difference (at a conventional level) between the two samples of parents with respect to the mean age and the sex and race compositions. We also see no statistically significant difference between the samples with respect to the fractions of parents with some college education and with a bachelor's or higher degree. However, the tests and comparison of Tables 2 and B1 reveal that the proportion of individuals without a high school degree among the FACHS parents is higher than among the CPS parents (18% vs. 13%, respectively), while the proportion of high school graduates among the FACHS parents is lower (34% vs. 39%, respectively). These differences are significant at the 5% level. Thus, the FACHS parents seem to have a somewhat lower educational attainment overall. Also, we see that fewer FACHS than CPS parents are married (36% vs. 43%, respectively) and more FACHS parents live in poverty (28% vs. 22%, respectively).[35] The latter differences

[33] More specifically, given the observed races of the FACHS youths, we selected the CPS households with youths who report their race as "African American only" or as any bi- or tri-racial combination involving "African American."

[34] In other words, there are only 5% of households with a single father in the CPS subsample.

[35] The incidence of poverty in the FACHS sample at 28% is also higher than the incidence of poverty among all African Americans in the U.S. in 2006 at 24% (DeNavas-Walt et al., 2007, p. 47).

Table B1: Characteristics of parents in CPS subsample

Characteristic of Parent	Mean	Std. Dev.	Min	Max	z-stat[a]	P-value[a]
Age in years	44.65	(6.68)	30	80	1.00	(0.32)
Indicator of male sex	0.05	(0.23)	0	1	−0.17	(0.87)
Indicator of African American race[b]	0.93	(0.25)	0	1	−0.68	(0.50)
Indicator of no high school degree[c]	0.13	(0.34)	0	1	2.05	(0.04)
Indicator of high school degree[c]	0.39	(0.49)	0	1	−1.96	(0.05)
Indicator of some college education[c]	0.31	(0.46)	0	1	1.49	(0.14)
Indicator of BA or higher degree[c]	0.17	(0.37)	0	1	−1.46	(0.14)
Indicator of married parent	0.43	(0.50)	0	1	−2.58	(0.01)
Indicator of poverty	0.22	(0.41)	0	1	2.52	(0.01)

Notes:
The number of parents in the CPS subsample is $1,053$. Statistics are computed using the ASEC Supplement weights.
[a] z-statistic and P-value refer to a two-sided test of the null hypothesis that the mean of the characteristic in the FACHS sample is equal to the mean in the CPS subsample.
[b] Four percent of parents are "White only," 2% are bi- or triracial with "African American" as one of the races, and 1% report some other race.
[c] Educational categories represent the highest level of educational attainment.

are significant at the 1% level.

We conclude that basic demographic characteristics of the FACHS sample of parents are practically the same as the characteristics of the CPS subsample, except that the proportion of married parents is lower in the FACHS. However, parents in the FACHS tend to have less income and a somewhat lower educational attainment than parents in the CPS.

Appendix C Robustness Analysis

We estimated several different specifications of the empirical model to check the robustness of the findings reported in Subsections 4.1 and 4.2.[36]

First, we replaced the relative price measures with "raw" price indices, that is, the expenditure weighted averages of prices of the fruit and vegetable groups in the QFAHPD (see Subsection 2.4 for details). In the case of fruit, we find little difference from the results with the relative price measures. The new estimates of the endogenous effects between the youth and parent are still positive and statistically significant, while the ones between the youth and friend remain insignificant. There are a few minor changes in the magnitude and significance of the constant terms and age terms. However, we still find that the youth's consumption decreases with age, while the consumption of the parent increases with age. The coefficients on the parental marital status are practically unchanged. As before, all other demographic and socioeconomic characteristics do not have a statistically significant impact, and estimates of the correlations among the errors indicate the presence of a correlated effect between the youth and parent. Most notably, we find a statistically significant negative effect of the "raw" fruit price index on the parent's consumption and marginally significant (at 11%) negative impact on the friend. The coefficient on the price faced by the youth is negative but not significant.

The results for vegetable consumption tend to be slightly more sensitive to the change in the price variables, but we still find many similarities. As before, we estimate a positive and significant endogenous effect of the parent's consumption on the youth's consumption and no effect in the reverse direction. The endogenous effects between the youth and his or her friend remain insignificant. There are a few small differences in the magnitude of coefficients on the constant terms and age terms. We still find that the consumption of the youth declines in the late teens, but the estimates no longer imply that it increases after age 20. As before, the friend's consumption tends to increase with age at a decreasing rate. Likewise, the parent's consumption tends to increase with age at a decreasing rate. The impact of the parent's being married on his or her own intake of vegetables remains positive and significant. The coefficients on all other demographic and socioeconomic characteristics remain insignificant. There are a few changes in the significance of the estimated correlations among the errors (namely, the estimate of ρ_{YF} loses significance, while the estimate of ρ_{YP} becomes significant at the 10% level), but the direction of every correlation is unchanged and the magnitude of the difference from the previous results is small. Most notably, we do not find the "raw" vegetable prices to exert a significant impact on consumption. Perhaps the corresponding effects are statistically weak and our sample size is insufficiently large to estimate them precisely. It is more likely, however, that the specification with the "raw" prices is too crude to correctly capture price effects, since it ignores prices of substitute goods. Therefore, we believe that the results with the relative prices in Subsection 4.2 are more informative.

Second, we re-estimated the models while excluding quadratic age terms. Apart from minor changes in the magnitude and significance of the estimates, the results are similar to the ones with quadratic age terms included. In the case of fruit, we find that the youth's

[36] Numerical estimation results discussed in this appendix are available from the authors on request.

consumption declines with the youth's age, while the parent's consumption increases, which is broadly in line with the earlier findings. We also obtain slightly larger estimates of the endogenous effect of the friend on the youth and of the friend's being African American. In addition, the estimate of the effect of the relative price faced by the friend becomes marginally insignificant. All other results are practically unchanged.

In the case of vegetables, we find that the friend's consumption and parent's consumption increase with age, which is broadly consistent with the earlier estimates. As before, we find that the youth's consumption decreases with age, but the corresponding estimate becomes insignificant.[37] Also, there are some changes in the significance of the correlations among the errors (the estimate of ρ_{YF} loses significance, while the estimate of ρ_{YP} becomes significant at the 10% level), but the change in the magnitude of the correlations is negligible. All other results are practically unchanged.

Third, as noted in Subsection 2.4, a small number of the FACHS respondents were interviewed in 2007 for which no data are available in the QFAHPD. Previously, instead of dropping these observations, we merged all FACHS records from 2007 with the price data from the fourth quarter of 2006. Such imputation may introduce errors in the price variables. Thus, we re-estimated the models while excluding all triplets with imputed prices from the sample (there are eight such triplets, which leaves a total of 494 triplets to re-estimate the models). We find that the results remain practically unchanged after the exclusion.

Lastly, we were surprised to find no impact of parental education on the consumption of fruits and vegetables. To check whether this result may be due to our specification of the two education categories (a high school degree or less vs. some college with or without a degree), we estimated versions of the empirical model with two different education categories (namely, some college without a degree or less vs. BA or a higher degree), as well as with four categories (no high school degree, high school degree, some college but no BA degree, and BA or a higher degree). The coefficients on education indicators remain insignificant. Perhaps the sample size is insufficiently large to estimate the impact of education precisely.

To conclude, the additional analyses indicate that the main results reported in Subsections 4.1 and 4.2 are, overall, robust to changes in the specification of the empirical model and are not substantially affected by limitations of the available data.

[37] It should be noted that a likelihood ratio test rejects the null hypothesis of no quadratic age effects at the 1% level. Thus, the insignificance of the estimate may have resulted from incorrect specification.