# The Effect of the National School Lunch Program on Childhood Obesity

Yongwon Cho*

June 1, 2013

**Abstract**

Childhood obesity has been recognized as not only a personal disease, but also as a social problem inducing government development of policies to address the childhood obesity problem. Since the National School Lunch Program (NSLP) has been one of the most important nutrition supplement programs for children, the government has tried to solve the childhood obesity problem through the NSLP. We set a goal of the paper as investigating theoretical relationship between the NSLP and the children's BMI and establishing support for a new school lunch program policy emphasizing quality of foods to reduce the childhood obesity epidemic. To develop an efficient nutrition supplement policy on the childhood obesity problem, this paper studies the significance of participation in the NSLP on the improvement in childhood obesity and clarifies conditions to encourage parents' continued participation in the NSLP using both theoretical and empirical models. The theoretical results show that participating in the NSLP affects the children's BMI change and contributes to improving the childhood obesity problem. Moreover, the results exhibit if the marginal rate of substitution between high and low calories in the NSLP is greater than eating at home, the parent

*Doctoral Candidate, School of Economic Science, Washington State University. mail to: yongwon_cho@wsu.edu.

of an overweight child is more likely to continue participating in the NSLP to solve their child's obesity problem. The theoretical results are tested empirically by using the ECLS-K data and we simulate the child's BMI percentile change based on the empirical results.

# 1 Introduction

Childhood obesity has been recognized as not only a personal disease, but also as a social problem inducing government development of policies to address the childhood obesity problem. The National School Lunch Program (NSLP) has been one of the most important nutrition supplement programs for children; the government has tried to solve the childhood obesity problem through the NSLP policy. For this reason, previous studies have investigated the effect of the NSLP on children's Body Mass Index (BMI) and they have found that the NSLP is useful in improving the childhood obesity problem. However, these studies have focused on empirical analysis so that the relationship between the NSLP and the childhood obesity still remains as black box. Thus, we set a goal of this paper as investigating theoretical relationship between the NSLP and children's BMI change and then establishing theoretical, empirical support for a new school lunch program policy emphasizing quality of foods to reduce the childhood obesity epidemic. To develop an efficient nutrition supplement policy on the childhood obesity problem, this paper studies the significance of participation in the NSLP on the improvement in childhood obesity and clarifies conditions to encourage parents' continued participation in the NSLP using both theoretical and empirical models.

We can check several current issues related to children's foods and beverages purchasing habits outside of the NSLP through the School Health Policies and Program Studies 2006 (SHPPS 2006; see tables in the appendix)[1]. According to the SHPPS report (2006), over 33% of elementary schools, 71% of middle schools, and 89% of high schools have a vending

[1]SHPPS is a national survey periodically conducted to assess school health policies and programs at the state, district, school, and classroom levels.

machine or a school store, canteen, or snack bar where students could purchase foods or beverages. Only 4% of states required that schools makes fruits or vegetables available to students whenever food was offered or sold whereas 18.4% of states required schools make healthful beverages such as bottled water or low-fat milk available to students whenever beverages were offered or sold. The study also shows that 11.9% of elementary schools, 25.4% of middle schools, and 48% of high schools allowed students to purchase foods and beverages high in fat, sodium, or added sugars from a vending machine or in a school store, canteen, or snack bar during lunch periods. Moreover, 12.9% of elementary schools, 28.7% of middle schools, and 58.2% of high schools allowed students to buy soda pop, fruit drinks that are not 100% juice, or sports drink from a vending machine or in a school store, canteen, or snack bar during lunch times. And only 2% of states required that schools prohibits advertising for candy, fast food restaurants, or soft drinks on school property. Furthermore, children's foods and drinks consumption patterns changes from 2000 to 2006. For example, the children's foods purchasing percentage from vending machines or in school stores decreased in non-low-fat baked goods, ice cream or frozen yogurt, salty snacks, and whole milk whereas purchasing percentage of bottled water increased. Finally, the government has tried to reduce children's accessibility to junk food in school settings. From 2000 to 2006, the percentage of states that required prohibiting junk food decreased in each school setting. For instance, percentage of prohibiting junk food in a la carte breakfast or lunch periods increases twice (from 20% to 42%) whereas that of vending machines increases 4 times (from 8% to 32%) and that of school stores, canteens, or snack bars increases over 5 times (from 6% to 32%) (Kann, Telljohann, and Wooley, 2007).

In the theoretical development, we set up two-periods optimization problem between a child and parent. A two-stage sequential game between a child and her/his parent is used to solve the child's food choice in school subject to the parent's decision to participate in the NSLP in each period. When the period moves from the first to the second term, the parent decides whether to keep participating in the NSLP after observing the child's BMI

percentile change during the previous period. The theoretical results show that participating in the NSLP affects the children's BMI percentile change and contributes to improving the childhood obesity problem. Moreover, the results exhibit if the marginal rate of substitution between high and low calories in the NSLP is greater than eating at home, the parent of an overweight child is more likely to continue participating in the NSLP to solve their child's obesity problem.

These results are tested empirically using the Early Childhood Longitudinal Study-Kindergarten Cohort (ECLS-K) data collected by the U.S. Department of Education following children in the kindergarten class of 1998-1999 to the eighth grade. The ECLS-K data has wide range of information about children, their families, and schools including children's weight and activities, family demographic structures, home environments, food security, and accessibility of various types of food. Since the ECLS-K data includes several categorical data, we introduced interval regression to change the categorical variables to continuous variables. We employed a new BMI calculation method from the Center for Disease Control and Prevention (CDC) to consider children's Biological Implausible Values, which are different from the method for adult's BMI calculation.

Our empirical estimation procedure has three steps. First, the parent's participation decision in the NSLP is estimated by using probit estimation method and then we obtain a probability of participating in the NSLP for each family. Second, we estimate the child's food choice in the school based on the parent's participation decision. We let the probability of participating in the NSLP as the parent's decision observed by the child. Since we have frequencies of sweet, salty snack, and drink consumption in school from the ECLS-K data, the Seemingly Unrelated Regression (SUR) procedure is used to obtain more efficient estimates than the Ordinary Least Squares (OLS) method. We use the probability of participating in the NSLP as an instrument variable for the parent's decision to solve the selectivity problem in this step. Third, the child's BMI percentile change is estimated by using the first and the second estimation results. We use the children's biological information and their family

backgrounds, the child's food choice in school, and the parent's participation decision in the NSLP as the independent variable and estimate by using a simple OLS method. Finally we forecast the children's BMI change by using Monte Carlo simulation so that we are able to test whether the children's probability to be obese decreases if they participate in the NSLP.

From the theoretical and empirical models, we obtain several policy implications to improve the childhood obesity epidemic. One of the important policy implications of the paper is that a well-managed NSLP, which provides low-calorie and high quality food are able to keep children participating in the NSLP. Thus, the new NSLP will play a substantial role in improving childhood obesity since the NSLP has a significant effect on the children's BMI. The well-managed program can be used to solve the low-income childhood obesity problem by retaining the low-income children in the NSLP, which the government can control.

# 2 Literature Review

The sequential game has been proposed by Kreps and Wilson (1982) which proposes an equilibria of extensive game and develops the properties of sequential equilibria. They expand a new criterion of the sequential rationality from single person decision problem and justify a standard roll-bak procedure which is similar to backward induction to construct a sequentially optimal strategy in a problem described by a decision tree. In our study, we employ this model to two-players which the parent decides first and the children move after observing their parent's decision. Rubinstein (1982) characterizes the perfect equilibrium partitions under certain assumptions and shows several properties which the players' preference possess are assumed. We introduce the paper's game rule which one player has made an offer, then the other must decide either to accept or to reject it and continue the bargaining. The sequential game structure has been applied in various fields. For example, Varian (1994) examines applied sequential game involving private contribution to a public good. The author shows that less of the public good will be supplied if agents move se-

quentially than if they move simultaneously. Ferrall and Smith (1999) estimate parameters of sequential game model of best-of-$n$ championship series in sports fields. The authors control for measured and unmeasured differences in team strength and boostrapping the maximum likelihood estimates to improve small sample properties. Moreover, Hudson and Lusk (2004) have applied the sequential game structure in the agricultural field. The authors show strategic interaction between food companies and activists using a game theoretic model of sequential bargaining in incomplete information. They find that it is always in the best interest of the food company to comply with the activist's demand. In other field, Gay et al. (1989) attempt to determine the impact of a defendant's strategic choice of trial mode on the judicial process by using a sequential signalling game. They find out more defendants will choose a jury trial rather than bench trial, nevertheless the equilibrium conviction rate is higher and they test it empirically. Though we do not use a signalling game structure, we introduce the paper's solving process which decomposes the problem into each case and find solution for the each case.

Since the 1996 Welfare Reform declined cash welfare benefits, food assistance programs have become an increasingly important part of the nation's social safety net for low-income families. And food programs have been instrumental in protecting low-income individuals from the adverse effects of the Great Recession (USDA, 2011). Recently, the Health, Hunger-Free Kids Act of 2010 mandates that school meals contain more fruits, vegetables, whole grains, fat-free and low-fat milk, and less saturated fat, sodium, calories and trans-fats. Initial provisions of the Act are scheduled to take effect during the 2011-2012 school year (USDA, 2011). Moreover, Gleason and Suitor (2003) use fixed effect model to control for time-invariant unobservable and demonstrate that the NSLP participants. Millimet, Tchernis, and Husain (2010) consider the impact of simultaneous participation in both the School Breakfast Program (SBP) and NSLP on obesity. They demonstrate there is positive selection in the SBP by children with higher body weight, but little evidence of selection into the NSLP. They show a positive correlation SBP participation and obesity, and it inflates the

impact of the NSLP on obesity. These effects are smaller for children who enter kindergarten with normal weight and larger for children who start school overweight or obese.

# 3 Theoretical Model

In the previous chapter, we have known that the parent and their child have own decision making mechanism to decide participating in the NSLP and food choice in the school, respectively. Since the child's food and exercise choice sets are limited by the parent's decision, the theoretical model of the previous chapter sets up a dynamic game model wich the parent moves first and the child optimizes his/her actions after observing the parent's action. Therefore, both the parent and the child's decision are included in one-time period and then they choose own actions after observing results of the previous period (see figure 1). In the first period, the parent decides whether participate in the NSLP through solving their utility maximization problem constrained by family budget, and then the child chooses food in school through solving hie/her own utility maximization problem constrained by the parent's decision. From this procedure, the child's BMI is determined and the parent's use this information to decide to keep participating in the NSLP at the second period. We know this procedure is one of dynamic games of complete and perfect information (Gibbons, 1992). Because both the parent's and the child's decision are made in sequencially so that the child is able to observe his/her parent's all previous actions before she/he chooses food in school. In addition, we assume that each party's payoff from each feasible combination of choices are common knowledge. Thus, we solve these optimization problems through backward induction and obtain a sequence of optimal actions in each period. Also, this decision making game between the parent and the child is repeated until the child graduates so that this game is finitely repeated game. We know the finitely repeated game is similar to the non-repeated game so that we analyze two period game between the parent and the child.

## 3.1 Child's Food Choice in School

Since this paper studies the effect of NSLP on childhood obesity, we focus on overweight case only, and then the parent's decision of the second period ($\lambda_{t+1}$) depends on

$$\lambda_{t+1} = 1 \text{ if } BMI_{ct+1} \leq BMI_{ct} \tag{1}$$
$$\lambda_{t+1} = 0 \text{ otherwise}$$

where $BMI_{ct}$ is the child's BMI in the first period. The condition (1) implies the parent's decision structure that they will keep participating in the NSLP when their child's BMI reduced or leave from the NSLP if their child's BMI increases. If we focus on underweight or malnutrition case, the parent's decision condition (1) has opposite sign. Now we are able to extend the parent's decision condition into two period as (2):

$$\text{When } \lambda_t = 1, \quad \begin{array}{l} \lambda_{t+1} = 1 \text{ if } \triangle BMI_{ct} \leq 0 \\ \lambda_{t+1} = 0 \text{ otherwise} \end{array} \tag{2}$$

$$\text{When } \lambda_t = 0, \quad \begin{array}{l} \lambda_{t+1} = 1 \text{ if } \triangle BMI_{ct} \geq 0 \\ \lambda_{t+1} = 0 \text{ otherwise} \end{array}$$

where $\lambda_t$ is the parent's participation decision in the first period and $\triangle BMI_{ct}$ is the child's BMI change. The condition (2) means if an overweight child participated in the NSLP and his/her BMI is reduced during the first period, then the parent decide to keep participating in the NSLP because they think the NSLP is effective to reduce their child's obese problem. On the other hands, if the overweight child did not participate in the NSLP and his/her BMI increases during the first period, then the parent decides to participate in the NSLP to seek alternative ways for reducing their child's obese problem. Finally the parent's and the child's optimization problem is similar to the previous chapter and solve the child's problem first and then solve the parent's problem contrained by the child's actions through backward induction.

Suppose the parent decides to participate in the NSLP at the second period ($\lambda = 1$), we solve the child's food choice problem such that

$$\underset{\{f_h, f_l, a_p\}}{Max}\ u = u\left(\lambda c_s, c_o, a_p, BMI_c, X\right) \tag{3}$$

$$s.t.\ \triangle BMI = \delta CI\left(f_h, f_l\right) - CO\left(a_p, G\right)$$

where $c_s$ is the child's food consumption in school whereas $c_o$ is his/her food consumption in other places except school, both are function of food choice $(f_h, f_l)$. $a_p$ is a time share of physical activities, and $X$ is consumption of all other goods except foods. $\delta$ is the child's exogenous metabolism rate, $CI$ is the child's total amount of calorie-in which is function of food choices whereas $CO$ is that of calorie-out which is function of the phycal activities time share and the child's genetic characteristics $(G)$. The child's optimization problem has a biological constraint which means the child's BMI is determined by difference between total calorie-in and total calorie-out. Whereas if the parent decides to leave from the NSLP ar the second period, then we solve the child's food choice problem different from the previous case as

$$\underset{\{f_h, f_l, a_p\}}{Max}\ u = u\left(\left(1-\lambda\right) c_s^h, c_o, a_p, BMI_c, X\right) \tag{4}$$

$$s.t.\ \triangle BMI = \delta CI\left(f_h, f_l\right) - CO\left(a_p, G\right)$$

where $c_s^h$ is food consumption prepared by the child's family. And the child has the same biological constraints as before. Since we use backward induction, we solve the child's problem first and then solve the parent's problem later.

Suppose the parent decides to participate in the NSLP, the first order conditions of the

child's problem are such that

$$\begin{aligned}
\lambda\frac{\partial u}{\partial c_s}\frac{\partial c_s}{\partial f_h}+\frac{\partial u}{\partial c_o}\frac{\partial c_o}{\partial f_h}+\frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial f_h}&=\mu\delta\frac{\partial CI}{\partial f_h}\\
\lambda\frac{\partial u}{\partial c_s}\frac{\partial c_s}{\partial f_l}+\frac{\partial u}{\partial c_o}\frac{\partial c_o}{\partial f_l}+\frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial f_l}&=\mu\delta\frac{\partial CI}{\partial f_l}\\
\frac{\partial u}{\partial a_p}+\frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial a_p}&=-\mu\frac{\partial CO}{\partial a_p}\\
\delta CI\left(f_{ht},f_{lt}\right)-CO\left(a_{pt},G\right)&=\triangle BMI_t
\end{aligned}\tag{5}$$

On the other hands, if the parent decides not to participate in the NSLP, the child's first order conditions changes as

$$\begin{aligned}
(1-\lambda)\frac{\partial u}{\partial c_s^h}\frac{\partial c_s^h}{\partial f_h}+\frac{\partial u}{\partial c_o}\frac{\partial c_o}{\partial f_h}+\frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial f_h}&=\mu\delta\frac{\partial CI}{\partial f_h}\\
(1-\lambda)\frac{\partial u}{\partial c_s^h}\frac{\partial c_s^h}{\partial f_l}+\frac{\partial u}{\partial c_o}\frac{\partial c_o}{\partial f_l}+\frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial f_l}&=\mu\delta\frac{\partial CI}{\partial f_l}\\
\frac{\partial u}{\partial a_p}+\frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial a_p}&=-\mu\frac{\partial CO}{\partial a_p}\\
\delta CI\left(f_{ht},f_{lt}\right)-CO\left(a_{pt},G\right)&=\triangle BMI_t
\end{aligned}\tag{6}$$

From the first order conditions of (5) and (6), we are able to let $\frac{\partial u}{\partial c_s}\frac{\partial c_s}{\partial f_k}+\frac{\partial u}{\partial c_o}\frac{\partial c_o}{\partial f_k}+\frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial f_k}=\frac{du}{df_k}$, $k=h,l$, and then we have

$$\frac{\frac{du}{df_k}}{\frac{\partial u}{\partial a_p}}=-\frac{\frac{\partial CI}{\partial f_k}}{\frac{\partial CO}{\partial a_p}},\quad k=h,l\tag{7}$$

It implies that ratio of marginal utility of high/low calorie food consumption and marginal utility of physical activity share equals to ratio of marginal calorie intake of high/low calorie food consumption and marginal calorie expenditure of physical activity share. And we note that the parent's participation decision does not affect to the child's optimal action since the parent's decision is exogenous. Moreover, we derive another result by modifying the

equation (7) as

$$\frac{\frac{\partial u}{\partial f_h}}{\frac{\partial u}{\partial f_l}} = \frac{\frac{\partial CI}{\partial f_h}}{\frac{\partial CI}{\partial f_l}} \tag{8}$$

which means marginal utility ratio of high and low calorie equals to marginal calorie intake ratio of high and low calorie. And we know that the parent's participation decision does not affect to the child's optimal food choice because the parent's decision is exogenous to the child.

## 3.2 Parent's Decision to Participate in the NSLP

After solving the child's optimization problem, we move to the parent's participation decision making problem. Since the parent's participation decision in the NSLP is a dichotomous problem, we analyze the parent's decision making procedure by using discrete choice model (Train and McFadden, 1978). When the parent decides whether participate in the NSLP, they consider family backgrounds such as family income, employment status, expenditure for food, and prices of NSLP or outside restaurants. We assume the family background the parent considers is the family's budget constraint. Moreover, the parent observes the child's optimal food choices in school and his/her BMI from the first stage. Based on these conditions, the parent decides whether to participate in the NSLP or not through solving utility maximization problem such as

$$\begin{aligned} &\underset{\{\lambda_t, Et\}}{Max}\, u = u\left(\lambda_t p_s, p_o, Z_t - \lambda_t N_t, E_t, BMI_{ct}\right) \\ &s.t.\ \ I_t \leq \lambda_t p_s N_t + p_o\left(Z_t - \lambda_t N_t\right) + E_t \\ &\quad N_t^* = N\left(f_{ht}^*, f_{lt}^*\right); \quad BMI_{ct} = \overline{BMI} \end{aligned} \tag{9}$$

where $\lambda$ is a dichotomous variable which is one if parent decides to participate in the NSLP, otherwise $\lambda$ is zero. We assume the parent's utility function is

We assume that the parent's utility is determined by food prices of both types, total

food expenditure, overall expenditures except foods. And $p_s$, $p_o$ represents exogenous price of NSLP and other places including home, other restaurants, respectively. When the parent decides to participate in NSLP, then $N$ represents child's total amount of food expenditure in school which is determined from the first stage by her utility maximization problem. $Z$ is total amount of overall food expenditures of the child. Finally, $E$ is overall expenditures except food consumption and $I$ is family's income. If we assume that family's budget constraint is binding for simplicity, we are able to plug the budget constraint into the utility function directly as

$$\max_{\lambda} U\left(\lambda p_s, p_o, p_o^{-1}\left(I - E - \lambda p_s N^*\right), E, \triangle BMI_c^*\right) \tag{10}$$

Since one of our decision making variables ($\lambda$) is binary variable, we cannot find the optimal $\lambda$ from deriving first order conditions. Rather, we obtain the parent's decision through introducing Random Utility Model (RUM). This model is derived from the parent's utility maximization problem and it can be used to represent their decision making on participating NSLP that does not entail the utility maximization problem (Train, 2003). We assume that parent's utility is decomposed by representative utility $V\left(p, \lambda, N^*, E, \triangle BMI_c^*\right)$ and a vector $\varepsilon$ such that

$$U_\lambda\left(p, \lambda, N^*, E, \triangle BMI_c^*\right) = V_\lambda\left(p, \lambda, N^*, E, \triangle BMI_c^*\right) + \varepsilon_\lambda, \quad \lambda = 0, 1 \tag{11}$$

where $\lambda$ is one when they decide to participate in NSLP, otherwise zero. And $\varepsilon$ captures factors that affect to the parent's utlity but are not included in the representative utilty $V(\cdot)$. Furthermore, the vector $\varepsilon$ is treated as random since the researchers do not know about that and its distribution depends critically on the researcher's specification of the representatitive utility.

Parent decides to select NSLP only if the utility derived from participating is greater than that of other type of lunches such as buying food from outside restaurant. Then probability

that the parent chooses to participate in NSLP is such that

$$\begin{aligned} \Pr(\lambda = 1) &= \Pr(U_1 > U_0) = \Pr(V_1 + \varepsilon_1 > V_0 + \varepsilon_0) \\ &= \Pr(\varepsilon_0 < \varepsilon_1 + V_1 - V_0) = P_1 \end{aligned} \tag{12}$$

Now we specify our model by defining density of unobserved factor. In this paper, the density of unobserved factor follows independently and identically distributed (*i.i.d.*) type I extreme value for all alternatives. Therefore, we choose Logit model for our analysis[2]. Moreover, we add assumptions that the unobserved factors are uncorrelated over alternatives as well as having the same variance for all alternatives[3]. And each choice is independent of the other. From those assumptions, we note that the unobserved factors related to one alternative might be similar to those related to another alternative. Then we are able to derive the Logit choice probability as equation (12)[4]. If we assume that the random vector follows type I extream value, then the cumulative distribution of the unobserved random vectors is such as

$$F(\varepsilon_0) = \exp[-\exp(-\varepsilon_0)] = \exp[-\exp(-(\varepsilon_1 + V_1 - V_0))] \tag{13}$$

Since we know the unobserved factor follows independently and identically distributed, the

[2]While applying the Logit model, we face two identification problems; only difference in utility matter and scale of utility is arbitrary. i) Only difference in utility matter implies that the absolute level of utility is irrelevant to both the parent's behavior model so that the only parameters that can be identified capture differences across alternatives. ii) The overall scale of utility is irrelevant means that adding a constant to the utility of all alternative does not change the parent's choice so that normalizing the scale of utility is equivalent to normalizing the variance of the error term.

[3]This assumption makes our analysis more convenient form for the choice variable

[4]Train (2003) represents the power of logit model; i) taste variation, ii) substitution patterns, and iii) repaeated choices over time. In our case, we employ the logit model to capture the parent's systematic taste variation. In other words, the parent's taste variation is not random, but is related to observed characteristics of the parents.

conditional probability and logit probability are represented as

$$P_1|\varepsilon_1 = \prod_{\lambda=0}^{1} \exp\left[-\exp\left(-\varepsilon_\lambda\right)\right] \tag{14}$$
$$P_1 = \frac{\exp\left(V_1\right)}{\exp\left(V_0\right) + \exp\left(V_1\right)}$$

The representative utility is able to have various functional forms. Regardless of the representative functional form, McFadden (1974) demonstrated that the log-likelihood function with these choice probabilities are globally concave in parameters. It helps in the numerical maximization procedure. After estimating logit model, we can obtain the parameters of determinants which affect to parent's participation decision and derive the ratio of coefficients which have economic meaning.

We note that the result (8) is satisfied regardless of the parent's participation decision. To optain optimal conditions for continuing participation, we derive the child's optimal food choices by specifying consumption functions. For simplicity, we define food consumption functional forms as Cobb-Douglas production functions such that

$$u\left(\cdot\right) = c_s^2 + c_o^2 + BMI_c^2 \tag{15}$$
$$c_s = T_1 f_h^\alpha f_l^\beta \ ; \quad c_s^h = T_2 f_h^a f_l^b$$
$$\triangle BMI_c = \delta\left(f_h + f_l\right) - E$$

and we let $\theta_1^*$ and $\theta_0^*$ are vector of the child's optimal high and low calorie food consumption when they participate in the NSLP or not, respectively. Suppose the child expects the parent will participate in the NSLP. By using the first and second conditions of $(5)$, we have

$$\frac{\frac{\partial u}{\partial c_s}\frac{\partial c_s}{\partial f_h} + \frac{\partial u}{\partial c_o}\frac{\partial c_o}{\partial f_h} + \frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial f_h}}{\frac{\partial u}{\partial c_s}\frac{\partial c_s}{\partial f_l} + \frac{\partial u}{\partial c_o}\frac{\partial c_o}{\partial f_l} + \frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial f_l}} = \frac{\frac{\partial CI}{\partial f_h}}{\frac{\partial CI}{\partial f_l}} \tag{16}$$

To make a calculation easier, we ignore two terms: $\frac{\partial u}{\partial c_o}\frac{\partial c_o}{\partial f_h} + \frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial f_h}$ and $\frac{\partial u}{\partial c_o}\frac{\partial c_o}{\partial f_l} + \frac{\partial u}{\partial BMI}\frac{\partial BMI}{\partial f_l}$

and let $\frac{\frac{\partial CI}{\partial f_h}}{\frac{\partial CI}{\partial f_l}} = D$. Then the equation (16) is modified as

$$\frac{\frac{\partial u}{\partial c_s}\frac{\partial c_s}{\partial f_h}}{\frac{\partial u}{\partial c_s}\frac{\partial c_s}{\partial f_l}} = \frac{\frac{\alpha c_s}{f_h}}{\frac{\beta c_s}{f_l}} = D \tag{17}$$

and we get $f_l = \frac{\beta}{\alpha} f_h D$. Plug the low calorie food consumption into the child's BMI change function, then we have the optimal food consumption of high calorie and low calorie:

$$\begin{aligned} &\theta_1^* = \left(f_h^1, f_l^1\right), \quad \text{where} \\ &f_h^1 = \frac{\triangle BMI + E}{\delta\left(\frac{\beta D}{\alpha} + 1\right)}; \quad f_l^1 = \frac{\beta D}{\alpha}\frac{\triangle BMI + E}{\delta\left(\frac{\beta D}{\alpha} + 1\right)} \end{aligned} \tag{18}$$

On the other hands, we have the optimal food consumption of high and low calorie when the parent decides not participate in the NSLP are such that

$$\begin{aligned} &\theta_0^* = \left(f_h^0, f_l^0\right), \quad \text{where} \\ &f_h^0 = \frac{\triangle BMI + E}{\delta\left(\frac{bD}{a} + 1\right)}; \quad f_l^0 = \frac{bD}{a}\frac{\triangle BMI + E}{\delta\left(\frac{bD}{a} + 1\right)} \end{aligned} \tag{19}$$

If the parent keeps participating in the NSLP, the high calorie food consumption under participation is less than high calorie food consumption under not participation whereas the low calorie food consumption under participation is more than low calorie food consumption under not participation: $f_h^1 \leq f_h^0$ and $f_l^0 \leq f_l^1$. In other words, we can represent as

$$\begin{aligned} \frac{\triangle BMI + E}{\delta\left(\frac{\beta D}{\alpha} + 1\right)} &\leq \frac{\triangle BMI + E}{\delta\left(\frac{bD}{a} + 1\right)} \\ \frac{bD}{a}\frac{\triangle BMI + E}{\delta\left(\frac{bD}{a} + 1\right)} &\leq \frac{\beta D}{\alpha}\frac{\triangle BMI + E}{\delta\left(\frac{\beta D}{\alpha} + 1\right)} \end{aligned} \tag{20}$$

and (20) is equivalent to

$$\frac{\frac{\partial c_s^h}{\partial f_l}}{\frac{\partial c_s^h}{\partial f_h}} \leq \frac{\frac{\partial c_s}{\partial f_l}}{\frac{\partial c_s}{\partial f_h}} \tag{21}$$

the equation (21) is represented to elasticity form: $\frac{\varepsilon_{c_s^h f_l}}{\varepsilon_{c_s^h f_h}} \leq \frac{\varepsilon_{c_s f_l}}{\varepsilon_{c_s f_h}}$. It implies that an elasticity ratio between low and high calorie food of participating in NSLP is greater than that of non participation. Furthermore, (21) means that marginal rate of consumption between high and low calorie in school is greater than the marginal rate of consumption in school. What if well-managed NSLP which share of low calorie food is greater than that of high calorie food, then children will choose low calorie foods rather than high calorie foods for consuming same calorie intake in school and it makes the children's BMI decrease.

An especially appealing aspect of the two-stage sequential game model is that the structure helps clarify and avoid the hybrid model problem discussed by Rosenzweig and Schultz (1983) of including exogenous determinants, and endogeneous determinants as independent variables in the same model. Since the goal of the theoretical model is to determine the parent's participation decision in the NSLP, the child's food choice in school and effect of the parent's decision on the child's health outcome, as measured by the child's BMI, we substitute out the child's foods choice with his/her best response to the parental decision. Therefore, the child's indirect BMI production function is the natural choice for our empirical model because of the parental inputs are then the argument (Capogrossi and You, 2012).

# 4 Empirical Model

## 4.1 Data Description

In the empirical model, we use the Early Childhood Longitudinal Study- Kindergarten Cohort (ECLS-K) database to test our hypothesis. This is a panel of data collected by the U.S. Department of Education following children in the kindergarten class of 1998-1990 to eighth grade. The data includes approximately 15,000 students in 100 different schools across the United States. The longitudinal study collected a wide range of information about children, students, their families, and schools including children's weight and exercise, activities, diet, family demographic structures, home environments, relationship between parent and chil-

dren, family income, food security, school characteristics, school facilities, and availability of different type of foods (IES, 2011). Among various waves, we selected the sixth and seventh waves because these have information about individual children's total quantity of food consumption in school, types of food consumed, place at which the food is purchased, and frequency of consuming food in school. Furthermore, these waves provide information about children's weight, height, activities in school, and their family background including parent's biological information, time allocation pattern, education level, employment status, and total family income.

Some previous published studies on childhood obesity have employed data from the National Health and Nutrition Examination Survey (NHANES). This survey is a program designed to assess the health and nutrition status of adults and children in the United States through interviews and direct physical examination. These data include various detailed information on dietary and health-related outcomes collected from self-reports, nutritional well-being, and physiological measurements. Though the NHANES has been widely used in previous researches on health and nutrition related to children's outcome, it has two problems: reporting bias derived from self-reporting and oversampling in vulnerable groups. Because of weight and wellbeing status are influenced by self-reporting, those variables can be under-reported especially by girls and overweight groups. In addition, this survey oversamples vulnerable groups including Hispanics and African Americans because these groups have been treated as important target populations in health surveys. On the other hands, the ECLS-K data reduce bias associated with non-response by adjusting for differential non-response (IES, 2011). The ECLS-K data also allow children's BMI to be calculated for differential selection probabilities. Moreover, we will be able to solve an identification problem associated with metabolic rate. Despite children's BMI being highly correlated with metabolic rate, the NHANES does not include any information on metabolic rate. However, the ECLS-K is able to solve this identification problem by using children's parents' weight and the children's initial weight. The logic of this method is based on previous studies in

biology convincingly linking heavier maternal weights during pregnancy with increased baby birth weight and the child's later risk of obesity (Harmon, 2010). For all of these reasons, we will employ the ECLS-K data to test our research question.

Though our data contain plenty of information, it is hard to apply to our empirical model directly because the ECLS-K data includes several categorical questions that we need to change to continuous variables. In particular, because family income is one of the essential variables in determining children's BMI, we modify this categorical variable to a continuous variable by using interval regression. Interval regression is used to change ordered categories into the exact value for each observation. We assume that the family income is determined by parent's age, education level, number of earners, race, and region. Based on these assumptions, we set up a family income model using those determinants as independent variables from the interval regression model. We find that marginal effect of the parent's age is decreasing whereas those of education level, number of earners are constantly increasing. After estimating the interval regression model, we calculate the exact family income of each observation and use this value in our empirical model. In addition, though the ECLS-K data contain children's BMIs, we employ a new BMI calculation method from the Center for Disease Control and Prevention (CDC). Because the new method contains indices of the anthropometric status of children based on the 2000 CDC growth chart that are considered to be biologically implausible values (BIVs), it is different compared with the adult BMI calculation method computed from their height and weight. After calculating students' BMI percentile using the new method, we use those values as dependent variables in our empirical model and redefine obese children who are over 95% of BMI percentile.

## 4.2 Model Description

For each family $i$, we model a two-stage decision process: the parent's dichotomous participation decision in the NSLP and the child's food choice in school. After estimating the two-stage decision process, we test effect of participating in the NSLP to the child's BMI

percentile change. To estimate the parent's participation decision in the NSLP, we consider the following empirical model specification as

$$\begin{aligned} D_{it} &= \sum_n x_{nit}\alpha_n + \triangle BMI\%_{it-1}\alpha + u_{it} \\ &= X_{it}\alpha + u_{it} \end{aligned} \tag{22}$$

where $D_{it} = 1$ if the parent decide to participate in the NSLP, $D_{it} = 0$ otherwise. The parents consider their child's BMI percentile change in last period when they decide to keep participating in the NSLP at current period. After observing the parent's decision, the child chooses foods in school as follows

$$C_{it} = Z_{it}\beta + \varepsilon_{it} \tag{23}$$

where $C_{it}$ denotes the child's food choice in school defined as a frequency of buying the food. From the two-stage decision process results, the child's BMI percentile change is determined by his/her food choice in school ($C_{it}$), the parent's participation decision ($D_{it}$), and other independent variables ($X_{it}, Z_{it}$) as

$$\begin{aligned} \triangle BMI\%_{it} &= f\left(C_{it}, D_{it}, X_{it}, Z_{it} : \theta\right) + \eta_{it} \\ &= Y_{it}\theta + \eta_{it} \end{aligned} \tag{24}$$

We assume functional form of $f$ is a linear defined as product of independent variable matrix $Y_{it}$ and estimates $\theta$, and $\eta_{it}$ is unobservable to the researcher. The row vector of independent variables, $X_{it}$ and $Z_{it}$ represent causal factors influencing the parent's participating decision in the NSLP and the child's food choice in school, respectively. Common elements include both time-variant and time-invariant variables such that family income, the parent's employment status, the child's food choice in school, participating in the NSLP, and number of earners are time-variant independent variables whereas family's biological characteristics including gender, race, the family's living area and the parent's education level are

time-invariant independent variables. From the theoretical model, we note that the parents consider their child's BMI percentile change of the last period when they decide to participate in the NSLP at current period. So we employ the child's BMI percentile change of the last period which is calculated from the result of the previous empirical chapter. When the parents decide to keep participating in the NSLP, they concern the effect of the NSLP during the last period. However, the child's BMI percentile change are affected by several factors including the NSLP: for example, the child's food consumption in school except the NSLP, the child's food consumption in home, and food environment of family. For this reason, we use the child's BMI percentile change of the last period calculated by using the last period information to capture the effect of the NSLP only so that we can be able to reflect the effect of the NSLP to the parent's decision making procedure. Finally, the column vectors $\alpha$, $\beta$, and $\theta$ represent the coefficients of the independent variables of the parent's participation decision, the child's food choice in school, and the child's BMI percentile change which we seek to estimate.

The independent variables $X_{it}$ in the parent's decision equation include the following variables: the family income, state's average and standard deviation spending for health care per capita, and state's average and standard deviation spending for family-children care per capita, the parent's education level, the parent's time spending with the child, the parent's age, number of workers in the family, and the parent's working hours. Moreover, the independent variables $Z_{it}$ in the child's food choice equation include the probability of participating in the NSLP, the family income, the parent's education level, the family's living areas, the parent's average age, number of the family, and the parent's spending hours with children. Finally, the independent variables in the child's BMI percentile change equation contain the independent variables both in the parent's decision equation and the child's food choice equation: the probability of participating in the NSLP, the child's gender, family income, the parent's average age, the family's living area, the parent's spending hours with children, number of the family, and the child's food consumption in school.

From previous researches, we know that parents have different incentives to participate in the NSLP with respect to family income. Low income families tend to join to the NSLP because the program provides reduced or even free price to the low-income family children so that the parents are able to spend more budget for their children's nutrition intake. On the other hands, non-low income families want to participate in the NSLP because the program is convenient then the parents do not need to prepare their children's lunch, therefore, the child's food choice in school is affected by the parent's decision and it leads to sample selectivity problem. In particular, the child's food choice and the child's BMI percentile change equation estimates are subject to bias derived from the selectivity problem when unobservable determinants that affect to the parent's decision or affect to the child's food choice in school either. Thus, we reflect those sample selectivity problems through modeling the correlation between the error terms ($u_{it}$ and $\varepsilon_{it}$) of the parent's decision and the child's food choice equations.

According to previous studies related to the selectivity problem, Inverse Probability Weighting (IPW) and Matching methods used to correct selection on observables have problem in their assumption: it is too strong when conducting food assistance program evaluations. The difference in a self-selected treatment status of two different individuals in the treatment group and in the control group who share the same observable characteristics may be due to the unobservable that determine treatment selection (Meyerhoefer and Yang, 2011). If those unobservable variables in the selection process are dependent on the unobservable variables that determine the two potential outcomes, then those indicate the presence of selection on unobservable. Moreover, the previous literatures have introduced instrumental variables as treatment effects highlight the importance of allowing for heterogeneous effects of NSLP participation. Breunig and Dasgupta (2005) show effect of the Supplemental Nutrition Assistance Program (SNAP) participation on adult obesity is different in single adult and multiple adult households, although these two groups are pooled in empirical studies.

We seek to the estimate the individual-level response of the parent's and the child's choice

while controlling for endogenous use of independent variables due to sample selectivity. The panel selection model specified in (22) and (23) allows us to account for consideration, but estimating our model presents problem: estimation of the child's food choice equation is complicated by the presence of selectivity. Although a panel estimator (using fixed effects, random effects, or differencing) can be used to remove the individual-specific effects from (23), the selectivity bias does not disappear with differencing since the selection process changes over time. For this reason, the time-varying selection process have to be modeled explicitly in order to get consistent estimates of coefficients in our empirical models. As a solution, we employ regression method in a model of selection which Heckman (1976) introduced. Our sample rule is that $C_{it}$ is observed only when $D_{it}$ is greater than zero. It means we can observe food choice of the child participating in the NSLP only if the parent's decision is observed. Suppose error terms in both equations $\varepsilon_{it}$ and $u_{it}$ have a bivariate normal distribution with zero means and correlation $\rho$, and plug (22) and (23) into the theorem[5], then we obtain the model which applies to the observations in our sample as

$$\begin{aligned} E\left[C_{it}|D_{it}>0\right] &= E\left[C_{it}|u_{it}>-X_{it}\alpha\right] \\ &= Z_{it}\beta + E\left[\varepsilon_{it}|u_{it}>-X_{it}\alpha\right] \\ &= Z_{it}\beta + \beta_{\lambda}\lambda_i\left(\alpha_u\right) \end{aligned} \tag{25}$$

where $\beta_{\lambda} = \rho\sigma_{\varepsilon}$, $\alpha_u = \frac{-X_{it}\alpha}{\sigma_u}$, and $\lambda_i\left(\alpha_u\right) = \frac{\phi\left(\frac{X_{it}\alpha}{\sigma_u}\right)}{\Phi\left(\frac{X_{it}\alpha}{\sigma_u}\right)}$, called the inverse Mills ratio. Since $u_{it}$ is normally distributed and $\alpha_u$ is normalized, we can obtain the density of $u_{it}$ by using the

[5] The theorem is Moments of the Incidentally Truncated Bivariate Normal Distribution. If $y$ and $z$ have a bivariate normal distribution with mean $\mu_y$ and $\mu_z$, standard deviation $\sigma_y$ and $\sigma_z$, and correlation $\rho$, then

$$\begin{aligned} E\left[y|z>a\right] &= \mu_y + \rho\sigma_y\lambda\left(\alpha_z\right) \\ Var\left[y|z>a\right] &= \sigma_y^2\left[1-\rho^2\delta\left(\alpha_z\right)\right] \end{aligned}$$

where $\alpha_z = \frac{(a-\mu_z)}{\sigma_z}$, $\lambda\left(\alpha_z\right) = \frac{\phi(\alpha_z)}{1-\Phi(\alpha_z)}$, $\delta\left(\alpha_z\right) = \lambda\left(\alpha_z\right)\left[\lambda\left(\alpha_z\right)-\alpha_z\right]$

characteristics of the normal distribution as

$$E\left(u_{it}|u_{it}>\alpha_u\right)=\int_c^{\infty}\frac{u_{it}\phi\left(u_{it}\right)}{1-\Phi\left(c\right)}du_{it} \tag{26}$$
$$=\left(1-\Phi\left(c\right)\right)^{-1}\left(8\pi\right)^{-1/2}\int_c^{\infty}u_{it}\exp\left(-u_{it}^2\right)du_{it}$$

We know that $\frac{d\phi(u_{it})}{du_{it}}=\frac{-u_{it}}{(2\pi)^{1/2}}\exp\left(\frac{-u_{it}^2}{2}\right)$ so that

$$\int_c^{\infty}-\left(\frac{d\phi\left(u_{it}\right)}{du_{it}}\right)du_{it}=\int_c^{\infty}\frac{-u_{it}}{\left(2\pi\right)^{1/2}}\exp\left(\frac{-u_{it}^2}{2}\right)du_{it}=\phi\left(c\right) \tag{27}$$

When we use symmetric property of the normal distribution, we have

$$E\left(u_{it}|u_{it}>\alpha_u\right)=\int_c^{\infty}\frac{u_{it}\phi\left(u_{it}\right)}{1-\Phi\left(c\right)}du_{it}=\frac{\phi\left(-u_{it}\right)}{\Phi\left(-u_{it}\right)} \tag{28}$$

From the equation $(25)$, we know that

$$C_{it}|D_{it}>0=E\left(C_{it}|D_{it}>0\right)+v_{it} \tag{29}$$
$$=Z_{it}\beta+\beta_\lambda\lambda_i\left(\alpha_u\right)+v_{it}$$

Thus, we can reformulate the empirical model as follows:

$$\text{Selection mechanism}:\ D_{it}=X_{it}\alpha+u_i \tag{30}$$
$$\text{where}\ \Pr\left(D_{it}=1|X_{it}\right)=\Phi\left(X_{it}\alpha\right)\ \text{and}\ \Pr\left(D_{it}=0|X_{it}\right)=1-\Phi\left(X_{it}\alpha\right)$$
$$\text{Regression model}:\ C_{it}=Z_{it}\beta+\varepsilon_{it}$$
$$\text{where}\ \left(u_{it},\varepsilon_{it}\right)\ \tilde{}\ \text{bivariate normal}\ \left[0,0,1,\sigma_\varepsilon,\rho\right]$$

To estimate parameters in the model $(\mathbf{??})$, we employ Heckman's (1979) estimation procedure: i) Estimating the probit equation by maximum likelihood to obtain estimates of $\alpha$. For each observation in the selected sample, compute $\widehat{\lambda}_i=\frac{\phi(X_{it}\alpha)}{\Phi(X_{it}\alpha)}$ and $\widehat{\delta}_i=\widehat{\lambda}_i\left(\widehat{\lambda}_i-X_{it}\widehat{\alpha}\right)$.

ii) Estimate $\beta$ and $\beta_\lambda = \rho\sigma_e$ by OLS regression of $C_{it}$ on $Z_{it}$ and $\widehat{\lambda}_i$. After estimating parameters in the model (30), we are able to obtain consistent estimators of the individual parameters $\rho$ and $\sigma_\varepsilon$. The true condistional variance of the error term at each observation is $\sigma_i^2 = \sigma_\varepsilon^2 (1 - \rho^2\delta_i)$ and the average conditional variance for the sample converges to $\sigma_\varepsilon^2 \left(1 - \rho^2\overline{\delta}\right)$ which we can estimate by the least squares residual variance: $\frac{e'e}{n}$. For the squares of the coefficient on $\lambda$, we have plim$\widehat{\beta_\lambda^2} = \rho^2\sigma_\varepsilon^2$, whereas based on the probit results we have plim$n^{-1}\sum \widehat{\delta}_i = \overline{\delta}$. Now we are able to get a consistent estimator of $\sigma_\varepsilon^2$ by using $\widehat{\sigma_\varepsilon^2} = n^{-1}e'e + \widehat{\overline{\delta}}\widehat{\beta_\lambda^2}$, and an estimator of $\rho^2$ is obtained from $\widehat{\rho^2} = \frac{\widehat{\beta_\lambda^2}}{\widehat{\sigma_\varepsilon^2}}$ which provides a complete set of esimators of the model's parameters. Comparing to OLS estimation procedure, the marginal effect of the independent variable ($Z_{it}$) on the dependent variable ($C_{it}$) in the observed sample consists of two components: a direct effect on the mean of the dependent variable ($\beta_k$) and a particular effect appears only in a probability of the sample selection ($D_{it}$) is positive so that it influences to $C_{it}$ through its presence in $\lambda_i$. Thus, the full effect of changes in the dependent variable that appears in both $X_{it}$ and $Z_{it}$ on $C_{it}$ is such that

$$\frac{\partial E\left(C_{it}|D_{it} > 0\right)}{\partial z_{ik}} = \beta_k - \gamma_k \left(\frac{\rho\sigma_\varepsilon}{\sigma_u}\right) \delta_i\left(\alpha_u\right) \tag{31}$$

where $\delta_i\left(\alpha_u\right) = \lambda_i^2\left(\alpha_u\right) - \alpha_i\lambda_i\left(\alpha_u\right)$. We know the marginal effect is affected by size of $\delta_i\left(\alpha_u\right)$ depeding on the model setting and Greene (2008) points out that it frequently to be overlooked in empirical studies. To improve an inefficiency due to the property of the Heckman's two-step estimation procedure, Vella and Verbeek (1999) proposes a Full Information Maximum Likelihood (FIML) approach which removes the inefficiency in the Heckman's second stage estimation. Though the FIML estimation procedure is more efficient, the procedure is difficult and often infeasible to implement because the likelihood function is coomplex and it sometimes does not converge (Vella and Verbeek, 1999). For this reason, we left this estimation approach for future studies. Asymptotically, these two methods are equivalent, however in small samples the results can differ. Previous simulations have shown that FIML

is more efficient than the two-step approach but also more sensitive to mis-specification due to non-normal disturbance terms. When we estimate the child's BMI equation, we use the Seemingly Unrelated Regression (SUR) model to obtain more efficient estimates rather than estimating equation-by-equation by using standard OLS model. Even the OLS estimates are consistent, however those estimates are not as efficient as the SUR model generally. Because the error terms of SUR are assumed to be correlated across three equations (Davison and MacKinnon, 1993).

Moreover, after estimating the empirical model, we forecast the child's simulated probability of BMI change by employing Monte Carlo Simulation method because it has distinguished properties than other sampling methods. First, the Monte Carlo Simulation method can handle problem of greater complexity and size than most other methods. One of strengths of this approach is robustness and simplicity. Second, the Monte Carlo Simulation method is intuitively based on laws of large numbers and central limit theorem so that the marginal cost of learning is low (Jude, 1998). For these reasons, we employ the Monte Carlo method to simulate our empirical results. Though we are not able to know the child's exact BMI value, we can draw the estimate of the child's BMI based on certain conditions such as the child's demographic information and the family backgrounds. From the estimated range of the child's BMI, we will be able to know the distribution of possible BMI values through the mean and standard deviation of the child's BMI. By using a range of possible values, instead of a single guess, we can generate a more realistic forecasting of the child's BMI since the forecasted output of the empirical model is a range either. Therefore, the Monte Carlo Simulation tells us how likely the child's BMI are based on how we generate the range of estimates. This probabilistic nature of the Monte Carlo methods has important implications: the result of any Monte Carlo procedure is a random variable. Though any numerical method has error, however the probabilistic nature of the Monte Carlo error puts structure on the error so that the accuracy of the Monte Carlo methods can be controlled by adjusting the sample size. Thus, we can find answers in low cost but useful accuracy with

the Monte Carlo methods (Jude, 1998). We conduct the Monte Carlo Simulation procedure as follow: First, we create a parametric model from the empirical part. Second, we generate a set of random independent variables from a probability distribution over domain. Third, we evaluate the model and store the dependent variable and repeat the second and the third steps $n$-times, then we obtain the range of dependent variable. Finally, we analyze and forecast the range of dependent variable using various statistics.

In order to analyze our empirical results, because this study principally concerns the effect of the NSLP on childhood obesity, we categorized our sample into two groups: obese and normal-weight children. In defining those groups, we employed a definition based upon BMI percentile provided by the U.S. Center for Disease Control and Prevention (CDC). According to CDC Reports (2013), although BMI for children are calculated in the same way as for adults, the criteria used to interpret BMI in children and teens are different. BMI percentile, rather than BMI itself, is used in children and teens for two reasons: first, the amount of body fat changes with age; second, the amount of body fat differs between girls and boys[6]. According to CDC guidelines, children are obese when their BMI exceeds the 95th percentile; the same guidelines define normal-weight children as falling between the 5th and 95th percentile.

Our empirical estimation procedure had three steps. First, we estimated the probability of a parent participating in the NSLP using a probit model. Second, we estimated the child's in-school food choices based on the parent's participation decision. We let the probability of participating in the NSLP as the parent's decision observed by the child. Since we have frequencies of in-school sweets, snacks, and drinks consumption from the ECLS-K data, we used a Seemingly Unrelated Regression (SUR) procedure to obtain more efficient estimates instead of the OLS method. We used the probability of participating in the NSLP as an instrument variable for the parent's decision to solve the selectivity problem in this step.

[6]The CDC BMI for-age growth charts take into account these differences and allow translatioin of a BMI number into a percentile for a child's gender and age. For adults, on the other hand, BMI is interpreted through categories that do not take into account gender or age.

Third, the child's BMI percentile change was estimated by using the first and the second estimation results. We used the child's biological information, family background, in-school food choices, and the parent's participation decision in the NSLP as independent variables and made our estimates using a simple OLS method.

We employed independent variables in our empirical models as follows. Since family income is a categorical variable in the ECLS-K data, we converted it to a continuous variable by using an interval regression employing family background as an independent variable and family income as a dependent variable. Education refers to the number of years of education of the parent with the longer education history, be that the mother or the father. The ECLS-K data provides regional information as four categories: North East, Midwest, South, and West[7]. We selected the North East and West as our baseline and the Midwest and South as dummy variables. Spending hours refers to the number of hours a parent spends with the child. We calculated average parental age and used it as a stand-in for individual parent age. Snack, sweet, and drinks refer to the frequency with which a child consumed these items at school over the past week. Finally, we included the number of family members as an independent variable.

Among explanatory variables determining the parent's participating decision and the child's food choice in school, we needed to identify two groups explaining two different stages. Had we not done so, we would have faced a linear dependency problem in estimating a child's in-school food choices. We used the probability of participating in the NSLP calculated from the first stage in the second stage and that probability is determined by other independent variables used in the second stage. The instrumental variables we employ in the first stage are highly correlated with parental budget and time-spending constraints. We therefore employ state averages and standard deviations for health care spending, for family and child care,

[7]The ECLS-K data employs the Bureau of Labor Statistics classification: The North East areas include PA, NY, ME, VT, NH, MA, RI, CT, and NJ. The South areas contain MD, DE, DC, VA, WV, NC, SC, KY, TN, GA, FL, GA, AL, MS, AR, LA, OK, and TX. The Midwest areas have MI, OH, IN, IL, WI, MN, IA, MO, ND, SD, NE, and KS. The West areas have MT, WY, ID, WA, OR, CO, NM, UT, AZ, NV, CA, AK, and HI.

and for parental working hours as the instrumental variables to be used in the first-stage estimation.

## 4.3 Empirical Results

In the first step, we analyzed the parent's participation decision from the result of the probit model. We find that the parents who have obese children, are younger, live in the Midwest or South areas, or spend more time with their children are more likely to participate in the NSLP than parents for whom at least one of these factors is not true. Our results also show that parents of obese children whose BMI percentile increases in the previous period are more likely to participate in the NSLP. These results imply that parents expect the NSLP to have a standard for foods provided to children participating in the program that will result in improvements in their child's health and a decrease in their child's BMI. Moreover, we know that the parents with low family incomes are more likely to participate in the NSLP than other parents. Since NSLP pricing varies according to family income, low-income families can receive lunch at reduced prices and even for free, providing an extra incentive for these parents to join the program. These results support the results of previous studies showing that family income is negatively correlated with both childhood and adulthood obesity so that parents with obese children are more likely to join to the NSLP than parents with normal-weight children. In addition, it is worth noting that a greater percentage of the obese population lives in the South and Midwest regions than in the West and North East (CDC report, 2011); our empirical results are in-line with the CDC report[8]. Parents of normal-weight children have similar propensities compared with parents of obese children. Our results show that family income, parental education level, living areas, and number of hours spent with the child are significant determinants of the likelihood of NSLP participation of parents with normal-weight children. Thus, parents with low family incomes, who have fewer years of education, and who live in the South and Midwest are more likely to participate in

[8] Alabama (32.0) , Louisiana (33.4) , West Virginia (32.4) , Michigan (31.3) , Iowa (29.0) , and Ohio (29.6) .

the NSLP. Parents who are younger, spend more time with their children, and have more numerous families are also more likely to participate in the NSLP.

Based on the results of the probit model, we can calculate a probability describing how likely each family is to participate in the NSLP and a consequent estimation of obese and normal-weight children's in-school food consumption following the parental decision to participate. The SUR model is used to obtain more efficient estimates rather than estimating equation-by-equation via the standard OLS model. Even though the OLS estimates are consistent, they are generally not as efficient as estimates from the SUR model because the error terms of the SUR are estimated to be correlated across three equations: the consumption frequencies of sweets, snacks, and drinks (Davison and MacKinnon, 1993). The SUR estimation results of the obese and the normal children are summarized in table 7 and table 8. Our results show that obese children's in-school food choices are primarily affected by their family income and their number of family members. Obese children are more likely to consume more salty snacks in school if their family income is low and are more likely to purchase more drinks if their families are more numerous. If obese children participate in the NSLP, they are less likely to consume sweets and drinks at school. Also, if the obese children observe the probability of participating in the NSLP is higher then they are less likely to spend sweet foods and drinks in school, even the coefficient is insignificant so that we know that participating in the NSLP helps the child obese children to reduce spending foods in school. The NSLP will substitute for part of obese children's in-school food consumption. If those obese children participate in a NSLP that provides low-calorie, high-quality foods, therefore, then their in-school consumption of snacks and drinks will decrease. Thus, obese children's BMI percentile will decrease because their total calorie intake decreases. The propensity of the normal children's food consumption in school is similar to that of obese children. Region and family income are important factors determining the in-school food consumption of normal-weight children; normal-weight children with higher family incomes and who live in the South are more likely to consume more foods in school. The probability

of participating in the NSLP is higher for children who spend less money on food in school.

After estimating parental likelihood of participating in the NSLP (first stage) and the child's food choices in school (second stage), we finally estimated the child's BMI percentile change by using the probability of participating in the NSLP, the child's gender, family income, parental level of education, living region, hours the parent spends with the child, parental age, number of family members, and the child's in-school consumption of sweets, snacks, and drinks. The results of these estimations are summarized in table 9. Obese children's BMI percentile are likely to decreases when their probability of participating in the NSLP increases, which implies that the decision of parents to participate in the NSLP results in their children consuming fewer food calories at school. In addition, our results indicate that family income and living region are significant factors in determining the BMI percentile change of obese children; children whose family income is low are more likely to undergo an increase in BMI percentile, as are children who live in Midwest and South regions compared with children living in the West and North East. These results are similar to the results of previous studies dealing with childhood or adulthood obesity. From these results, we know that participating in the NSLP will play an important role in improving the childhood obesity problem because the marginal effect of the probability of participating in the NSLP is bigger than that of other independent variables. For this reason, we can therefore conclude that our theoretical hypothesis is reasonable and derive policy implications from our results. Finally, we note that our BMI estimation model has low $R^2$ because the children's BMI is mainly affected by calorie intake which is highly correlated to not only children's in-school food choice but also parental food selection or buying patterns. The ECLS-K data provides the children's buying frequencies of snacks, sweets, and drinks in school; however, much about children's tastes, personal preferences, food tastes, quality of foods, and parental food consumption tendencies remain unobservable and therefore missing from the data. We therefore anticipate obtaining more explanatory power if we gain information about these heretofore unobserved variables.

From our empirical results, we can simulate the effect of participating in the NSLP on an obese child's BMI percentile change. We use the average values of the independent variables, except for the dummy variables, plus the probability of participating in the NSLP. Since the main model has regional and gender dummy variables, we fix those variables as obese male children living in the South regions. We then calculate the obese child's BMI percentage change by substituting the probability of participating in the NSLP in three cases: 0%, 50%, and 100%. The simulation results show that when the probability of participating in the NSLP is higher, the obese child's BMI percentile decreases. For example, if the probability of participating in the NSLP is 100%, then the mean of the obese child's BMI percentile change is 8.74. When the probability decreases to 50%, then the mean of the obese child's BMI percentile increases to 39.05 and it goes up to 87.65 if the probability of NSLP participation decreases to 0%. From these simulated results, we can say that the marginal effect of the probability of participating in the NSLP is positive on reducing childhood obesity.

# 5 Policy Implications

The results of our theoretical model and empirical estimations allow us to derive two policy implications related to reducing childhood obesity. First, policy makers should promote the positive effect of the NSLP on reducing childhood obesity. Moreover, policy makers should advertise the importance of their children not being obese to parents so that parents consider their children's weight when making decisions about family food expenditures. Second, policy makes should improve the food quality standards of the NSLP. We know that participating in the NSLP reduces a child's likelihood of remaining obese. We can derive necessary conditions for retaining the likelihood of obese children's participation in the NSLP from our theoretical model. An elasticity ratio between low- and high-calorie food of participating in the NSLP is greater than that of non-participation, and the marginal rate of consumption between low and high calorie in the NSLP is greater than that of other food consumption in school

including cafeteria, outside restaurants, and vending machines.

# 6 Conclusion

Childhood obesity has been recognized as not only a personal disease, but also as a social problem which should induce the development of government policies aimed at reducing its incidence. Since the NSLP is one of the most important nutrition supplement programs for children in the United States, it has also been an important arm through which the government seeks to reduce childhood obesity. In this paper, we sought to investigate the theoretical relationship between the NSLP and children's BMI percentile change and establish grounds for a policy supporting an improved NSLP which emphasizes the quality of foods. To develop an efficient nutrition supplement policy on the childhood obesity problem, this paper studies the significance of participation in the NSLP on the improvement in childhood obesity and clarifies conditions to encourage parents' continued participation in the NSLP using both theoretical and empirical models.

In the theoretical model, we establish a two-period optimization problem between a child and parent. The theoretical results show that participating in the NSLP affects the children's BMI change and reduces childhood obesity. Moreover, the results demonstrate that if the marginal rate of substitution between high and low calories in the NSLP is greater than eating at home, the parent of an obese child is more likely to continue participating in the NSLP to solve their child's obesity problem. We tested these theoretical results empirically using ECLS-K data and we then simulated children's BMI percentile change based on our empirical results.

From these theoretical and empirical models, we arrive at several policy implications. One important policy implication of these results is that a well-managed NSLP which provides low-calorie, high-quality food is able to keep children participating in the NSLP. Thus, a new and improved NSLP can play a substantial role in reducing childhood obesity since the

NSLP has a significant effect on the children's BMI. A well-managed NSLP can be used to reduce low-income childhood obesity by retaining low-income children in the program, under government control.

Though this paper investigates the theoretical relationship between the NSLP and children's BMI percentile change and supports that theoretical relationship empirically, we do not separate out the effects of school-provided breakfast and lunch. To test the effect of the NSLP only (excluding school-provided breakfast effects), we would need to generate a subsample of individuals who participate in the NSLP only. Moreover, we need to find instrumental variables to capture quality of parental care – how much parent's care about their children – and instrumental variables reflecting family food environments including, for example, food accessibility and food security. Finally, we need to improve the efficiency of our estimates by employing a Full Information Maximum Likelihood (FIML) approach, which would remove inefficiency in the second stage of the two-step estimation procedure. Having addressed those weaknesses, we will be able to derive more efficient estimates and address the selectivity issues present in the current study.

**Figure 1: Time Horizon**

| $t = 0$ | $t = 1$ | $\cdots$ | $t = T$ |
|---|---|---|---|
| Parent decides to participate | $\lambda_{t+1} = 1$if $\triangle BMI_{ct+1} \leq 0$ | $\cdots$ | $\lambda_T = 1$if $\triangle BMI_{cT} \leq 0$ |
| Child's food choice | Child's food choice | $\cdots$ | Child's food choice |

**Table 1: Percentage of Schools Always Used Healthy Food Prepatation Practices**

| | 2000 | 2006 |
|---|---|---|
| Used part-skim or low-fat cheese instead of regular cheese | 31.0 | 45.9 |
| Trimmed fat from mean or used lean meat | 56.3 | 66.4 |
| Removed skin from pultry or used skinless poultry | 40.2 | 54.6 |

Note: During the 30 days preceding the study

Resource: Center for Disease Control and Prevention, 2006

**Table 2: Percentage of Schools Offered Low-fat a la Carte Foods**

| | 2000 | 2006 |
|---|---|---|
| Bread sticks, rolles bagels, pita bread, or other bread products | 50.8 | 67.1 |
| Lettuce, vegetable, or bean salads | 52.6 | 72.8 |
| Low-fat salty snacks | 38.2 | 53.2 |
| Low-fat or nonfat yogurt | 35.5 | 50.3 |
| Vegetables other than potatoes | 51.0 | 70.8 |

Resource: Center for Disease Control and Prevention, 2006

**Table 3: Percentage of Schools to Prohibit Offering Junk Foods**

| | 2000 | 2006 |
|---|---|---|
| A la carte during breakfast or lunch periods | 20.0 | 42.0 |
| Concession stands | 2.0 | 6.1 |
| School stores, canteens, or snack bars | 6.0 | 32.0 |
| Student parties | 2.0 | 8.0 |
| Vending machine | 8.0 | 32.0 |

Note: Defined as foods or beverages that have low nutrient density

Resource: Center for Disease Control and Prevention, 2006

**Table 4. Descriptive Statistics**

| | Obese | | Normal | |
|---|---|---|---|---|
| | Mean | Std. dev. | Mean | Std. dev. |
| NSLP (Participate in NSLP=1, otherwise=0) | 0.809 | (0.394) | 0.769 | (0.421) |
| Log(Income) (Family income) | 4.835 | (0.239) | 4.914 | (0.240) |
| Education (Parent's education level) | 4.868 | (1.773) | 5.507 | (1.813) |
| Midwest | 0.325 | (0.469) | 0.362 | (0.481) |
| South | 0.328 | (0.470) | 0.273 | (0.446) |
| West | 0.144 | (0.351) | 0.170 | (0.376) |
| Northeast | 0.203 | (0.403) | 0.194 | (0.395) |
| Spending hours (Spending hour with children) | 5.722 | (0.454) | 5.643 | (0.505) |
| Age (Parent's average age) | 42.60 | (5.269) | 43.93 | (5.454) |
| Log(HC_m) (Mean of states' health care spending) | 7.038 | (0.108) | 7.030 | (0.109) |
| Log(HC_sd) (Std. dev. of states' health care spending) | 5.621 | (0.339) | 5.602 | (0.348) |
| Log(FC_m) (Mean of states' spending for family and children's) | 3.860 | (0.204) | 3.873 | (0.199) |
| Working hours (Parent's average working hours) | 42.03 | (7.344) | 41.14 | (7.912) |
| N_Fam (Number of Family) | 4.342 | (1.087) | 4.406 | (0.967) |
| Sweet (Buying frequency in school) | 0.993 | (0.926) | 1.143 | (1.122) |
| Snack (Buying frequency in school) | 0.960 | (0.897) | 1.067 | (1.003) |
| Drink (Buying frequency in school) | 0.871 | (0.943) | 0.953 | (1.035) |
| $\triangle$BMI$\%_{t-1}$(BMI% change from 5th to 6th grade) | 3.535 | (1.480) | 2.033 | (0.533) |
| Number of observation | 403 | | 2544 | |

**Table 5. Correlation among Variables (All students)**

| Variables | BMI percentile | Obese | Log(income) | NSLP |
|---|---|---|---|---|
| BMI percentile | 1.000 | | | |
| Obese | 0.502 | 1.000 | | |
| Log(income) | -0.141 | -0.112 | 1.000 | |
| NLSP | 0.044 | 0.033 | -0.179 | 1.000 |

Note: Obese is defined as over 95% percentile of BMI percentile.

BMI percentile calculated from Center for Disease Control.

This matrix is symmetric.

**Table 6. Probit Estimation Result**

| Dependent Var. | Obese | | Normal | |
|---|---|---|---|---|
| Participate in NSLP | Coef. | Std. Err. | Coef. | Std. Err. |
| Log(income) | -2.182 | (9.246) | -12.378*** | (4.006) |
| Log(income)$^2$ | 0.205 | (0.958) | 1.177*** | (0.403) |
| Log(HC_m) | 1.736 | (1.318) | -0.645 | (0.521) |
| Log(HC_sd) | -0.711 | (0.454) | -0.374** | (0.179) |
| Log(FC_m) | -1.134*** | (0.437) | -0.476*** | (0.179) |
| Education | -0.053 | (0.050) | -0.053** | (0.021) |
| Parent's age | -0.011 | (0.016) | -0.014** | (0.006) |
| N_Fam | 0.001 | (0.072) | -0.018 | (0.032) |
| Spending hours | 0.138 | (0.191) | 0.146* | (0.078) |
| Working hours | -0.004 | (0.013) | 0.009* | (0.005) |
| $\triangle BMI\%_{t-1}$ | 0.023 | (0.035) | 0.178*** | (0.018) |
| Constant | 2.817 | (23.23) | 41.11*** | (10.25) |
| Observations | 403 | | 2544 | |
| Prob. $>\chi^2$ | 0.015 | | 0.000 | |

*** indicates significance at 1% level, ** indicates 5% level,

and * indicates significance at 10% level.

**Table 7. SUR Estimation Results (Obese)**

| Dependent var. | Sweet | | Snack | | Drink | |
|---|---|---|---|---|---|---|
| Independent var. | Coef. | Std. Err. | Coef. | Std. Err. | Coef. | Std. Err. |
| Pr(NSLP) | 1.084 | (1.200) | 2.060 | (1.427) | -0.904 | (1.201) |
| Gender | 0.173 | (0.122) | 0.140 | (0.145) | -0.150 | (0.122) |
| Log(Income) | -0.399 | (0.590) | -0.728 | (0.702) | 1.140 | (0.591) |
| Log(Income)$^2$ | 0.043 | (0.078) | 0.084 | (0.093) | -0.153* | (0.078) |
| Education | 0.033 | (0.038) | 0.046 | (0.045) | 0.012 | (0.038) |
| Midwest | -0.147 | (0.228) | -0.568** | (0.272) | 0.331 | (0.229) |
| South | -0.005 | (0.235) | -0.336 | (0.279) | 0.278 | (0.235) |
| Spending hours | -0.031 | (0.118) | 0.049 | (0.141) | -0.083 | (0.119) |
| Parent's age | 0.013 | (0.011) | 0.011 | (0.013) | -0.009 | (0.011) |
| N_Fam | 0.073 | (0.046) | -0.003 | (0.054) | 0.115** | (0.046) |
| $R^2$ | 0.485 | | 0.223 | | 0.423 | |
| Prob.>$\chi^2$ | 0.000 | | 0.000 | | 0.000 | |

*** indicates significance at 1% level, ** indicates significance at 5% level, and * indicates significance at 10% level.

**Table 8. SUR Estimation Results (Normal)**

| Dependent var. | Sweet | | Snack | | Drink | |
|---|---|---|---|---|---|---|
| Independent var. | Coef. | Std. Err. | Coef. | Std. Err. | Coef. | Std. Err. |
| Pr(NSLP) | -0.291 | (0.188) | -0.522*** | (0.170) | 0.092 | (0.172) |
| Gender | -0.075 | (0.047) | -0.082* | (0.043) | -0.064 | (0.043) |
| Log(Income) | 0.604*** | (0.199) | 0.325* | (0.180) | 0.352* | (0.182) |
| Log(Income)$^2$ | -0.048 | (0.029) | -0.009 | (0.026) | -0.032 | (0.027) |
| Education | -0.031 | (0.016) | -0.056*** | (0.015) | -0.018 | (0.015) |
| Midwest | 0.018 | (0.059) | 0.028 | (0.054) | 0.161** | (0.054) |
| South | 0.357*** | (0.060) | 0.252*** | (0.054) | 0.244*** | (0.055) |
| Spending hours | -0.030 | (0.047) | 0.086** | (0.043) | 0.002 | (0.043) |
| Parent's age | 0.0001 | (0.005) | 0.007 | (0.004) | 0.002 | (0.004) |
| N_Fam | -0.021 | (0.023) | -0.016 | (0.021) | -0.028 | (0.022) |
| $R^2$ | 0.509 | | 0.517 | | 0.468 | |
| Prob.>$\chi^2$ | 0.000 | | 0.000 | | 0.000 | |

*** indicates significance at 1% level, ** indicates significance at 5% level,

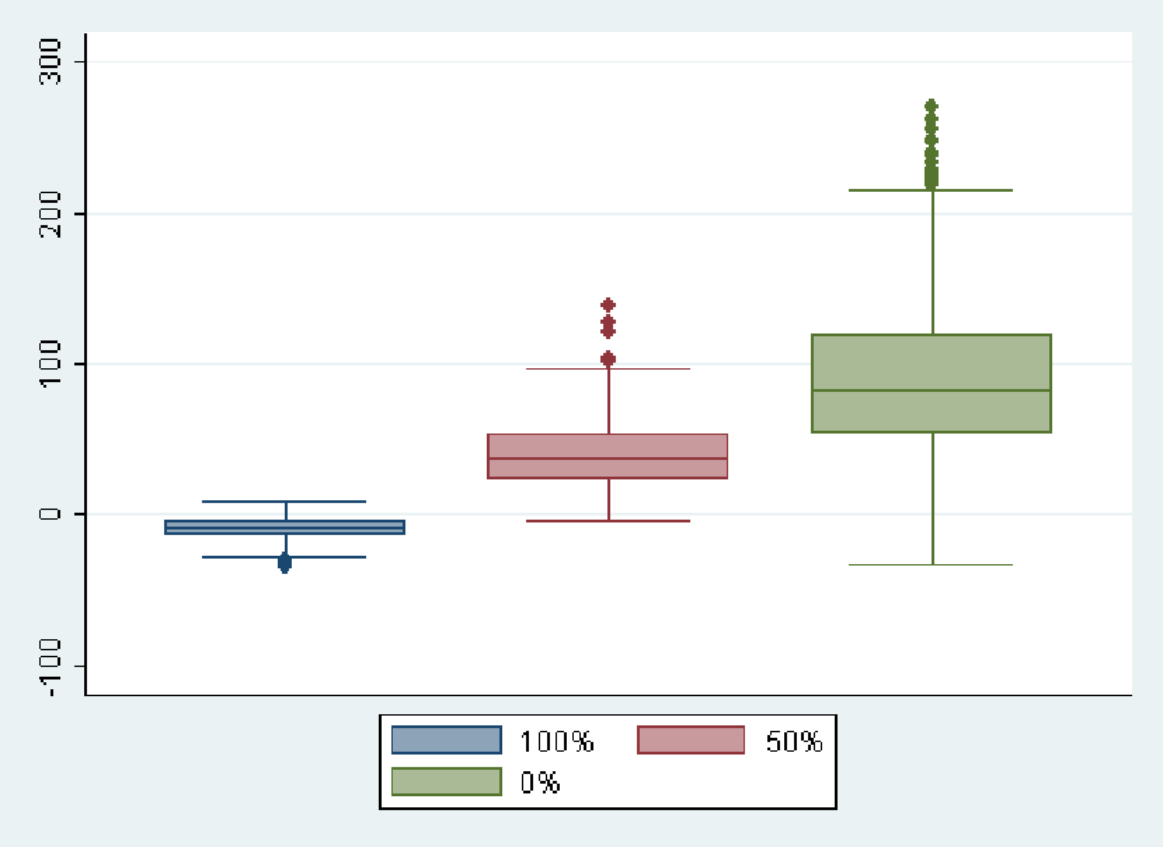
and * indicates significance at 10% level.

**Table 9. BMI Estimation Results**

| Dependent Var. | Obese | | Normal | |
|---|---|---|---|---|
| △BMI% | Coef. | Std. Err. | Coef. | Std. Err. |
| Pr(NSLP) | -44.94* | (25.77) | -3.918 | (3.089) |
| Gender | -3.141 | (2.580) | 3.643*** | (0.775) |
| Log(Income) | 19.30 | (12.73) | 5.031 | (3.249) |
| Log(Income)$^2$ | -2.452 | (1.666) | -0.583 | (0.477) |
| Education | -0.057 | (0.749) | -0.312 | (0.267) |
| Midwest | 5.594 | (4.996) | 3.107*** | (0.969) |
| South | 5.341 | (4.866) | 1.242 | (0.983) |
| Spending hours | 1.400 | (2.280) | 0.190 | (0.781) |
| Parent's age | -0.117 | (0.215) | -0.059 | (0.075) |
| N_Fam | 0.229 | (0.899) | 0.217 | (0.388) |
| Sweet | 2.417 | ()1.510 | 0.873** | (0.384) |
| Snack | 0.475 | (1.267) | -0.446 | (0.432) |
| Drink | -1.128 | (1.401) | 0.186 | (0.390) |
| Observation | 403 | | 2542 | |
| Prob.>$\chi^2$ | 0.020 | | 0.000 | |

*** indicates significance at 1% level, ** indicates at 5% level, and * indicates significance at 10% level.

**Figure 2. Simulation Result**



# References


[1] Breuning, R. and Dasgupta, I., 2005, Do Intra-household Effects Generate the Food Stamp Cash-out Puzzle? *American Journal of Agricultural Economics*, Vol. 87, No. 3, 552-568.

[2] Capogrossi, K. and You, W., 2012, Do School Lunch Programs Influence Child Weight? A Treatment Effect Analysis, 2012 Annual Meeting, Agust 12-14, 2012, Seattle, Washington, Agricultural and Applied Eonomic Association.

[3] Center for Disease Control and Prevention (CDC), 2006, School Health Policies and Program Study (SHPPS): A Summary Report, *Journal of School Health*, Special Issue, October 2006.

[4] Davison, R. and MacKinnon, J. G., Estimation and Inference in Econometrics, Oxford University Press, USA, January 14, 1993.

[5] Ferall, F. and Smith, A. A., 1999, A Sequential Game Model of Sports Championship Seires: Theory and Estimation, *The Review of Economics and Statistics*, MIT Press, Vol. 81, No. 4, 704-719.

[6] Gay, G. D., Grace, M. F., Kale, J. R., and Noe, T. H., 1989, Noisy Juries and the Choice of Trial Mode in a Sequential Signalling Game: Theory and Evidence, *RAND Journal of Economics*, Vol. 20, No. 2, 196-213.

[7] Gibbons, R., Game Theory for Applied Economists, Princeton University Press, Jul. 13, 1992.

[8] Gleason, P. M. and Suitor, S. W., 2003, Eating at School: How the NSLP Affects Children's Diet, *American Journal of Agricultural Economics*, Vol. 85, No. 4, 1047-1069.

[9] Greene, W., Econometric Analysis (6th Edition), Prentice Hall, 2008.

[10] Heitmueller, A., 2006, Public-Private Sector Pay Differentials in a Developed Scotland, *Journal of Applied Economics*, Vol. 9, No. 2, 295-353.

[11] Hudson, D. and Lusk, J. L., 2004, Activists and Corporate Behavior in Food Processing and Retailing: A Sequence Bargaining Game, *Journal of Agricultural and Resource Economics*, Western Agricultural Economics Association, Vol. 29, No. 1.

[12] Istitute of Education Sciences (IES), 2011, The Early Childhood Longitudinal Study, Kindergarten Class of 1998-99, U.S. Department of Education.

[13] Kann, L., Telljohan, S. K., and Wooley, S. F., 2007, Health Education: Result from the School Health Policies and Program Study 2006, *Journal of School Health*, Vol. 77, 408-434.

[14] Kreps, D. M. and Wilson, R., 1982, Sequential Equilibria, *Econometrica*, Vol. 50, No. 4, 863-894.

[15] Millimet, D. L., Tchernis, R., and Husain, M., 2010, School Nutrition Programs and the Incidence of Childhood Obesity, Journal of Human Resources, University of Wisconsin Press, Vol. 45, No. 3.

[16] McFadden, D., 1974, The Measurement of Urban Travel Demand, *Journal of Public Economics*, Elsevier, Vol. 3, No. 4, 303-328.

[17] Meyerhoefer, C. D. and Yang, M., 2011, The Relationship between Food Assistance and Health: A Review of the Literature and Empirical Strategies for Identifying Program Effects, *Applied Economic Perspectives and Policy*, Vol. 33, No. 3, 304-344.

[18] Mohanty, M. S., 2012, Effects of Positive Attitude and Optimism on Wage and Employment: A Double Selection Approach, *The Journal of Socio-Economics*, Vol. 41, Issue 3, 304-316.

[19] Rosenzweig, M. R. and Schultz, T. P., 1998, Consumer Demand and Household Production: The Relationship between Fertility and Child Materity, *American Economic Review*, American Economic Association, Vol. 73, No. 2, 38-42.

[20] Rubinstein, A., 1982, Perfect Equilibrium in a Bargaining Model, *Econometrica*, Vol. 50, No. 1, 97-109.

[21] Train, K. and McFadden, D., 1978, The Goods / Leisure Trade-off and Disaggregate Work Trip Mode Choice Models, *Transportation Research*, Vol. 12, 349-353.

[22] Train, K., Discrete Choice Methods with Simulation, Cambridge University Press, 1st Edition, 2003.

[23] Varian, H. R., 1994, Sequential Contributions to Public Goods, *Journal of Public Economics*, Vol 53, Issue 2, 165-186.

[24] Vella, F. and Verbeek, M., 1999, Two-step Estimation of Panel Data Models with Censored Endogenous Variables and Selection Bias, *Journal of Econometrics*, Elsevier, Vol. 90, No. 2, 239-263.

[25] Vella, F. and Verbeek, M., 1999, Estimating and Interpreting Model with Endogenous Treatment Effects, *Journal of Business and Economic Statistics*, American Statistical Association, Vol. 17, No. 4, 473-81.