The Cost Effectiveness of Educating Limited Resource Youths on Food and Nutrition

Ranju Baral, George C. Davis, Elena Serrano, Mary McFerren, and Wen You

Selected Paper prepared for presentation at the Agricultural and Applied Economics Association's 2013 AAEA & CAES Joint Annual Meeting, Washington, DC, August 4-6, 2013.

Ranju Baral is a Ph.D. candidate, George C. Davis is a Professor, and Wen You is an Assistant Professor, all in the Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, VA 24061. Elena Serrano is an Assistant Professor and Mary McFerren is a FNP Project Director, both in the Department of Human Nutrition, Foods, and Exercise, Virginia Tech, Blacksburg, VA 24061.

Abstract

A framework for estimating cost effectiveness of the youth Expanded Food and Nutrition Education Program (EFNEP), one of the largest nutrition education programs in the US, is developed. Using costs and effects data from 15 program counties in Virginia for the school year 2011/2012, the cost effectiveness ratio (CER) of the Healthy Weights for Healthy Kids (HWHK) program was estimated. Improvements in nutrition related behavior, and improvements in nutrition related self-efficacy, from pretest to posttest, were considered as two indicators of program effects. Direct costs associated with the provision of the HWHK program were considered program costs.

The total cost of the program was estimated to be $134,333 for one year. Among 1,864 participants, 1,786 improved in behavior indicators and 1,782 improved in self-efficacy indicators. The estimated CER is about $75 per improvement for each outcome indicator. Cross county comparisons reveal a wide variation in CER estimates across counties. The results from this study provide the first piece of information on the CER for youth EFNEP which quantify the effects of investment on youth EFNEP program in Virginia in achieving the stated program objectives.

1. Introduction

1.1. Problem Statement and Significance of the Study

The Expanded Food and Nutrition Education Program (EFNEP) is one of the largest federally funded nutrition education programs in the United States and is administered by the United States Department of Agriculture (USDA) (General Accounting Office 2004). The aim of EFNEP is to serve limited resource youths and adults by helping them acquire "the knowledge, skills, attitudes, and changed behavior necessary for nutritionally sound diets, and to contribute to their personal development and improvement of the overall family diet and nutritional well-being" (USDA 2013). In operation for more than 40 years and now in all 50 states and 6 territories, the EFNEP has become a cornerstone in US nutrition education (USDA 2013).

EFNEP appropriates more than $66 million dollars of the federal money, every year. Given the tight budgetary conditions, there has been an increasing pressure for the evaluation of federally funded programs to improve their effectiveness and to ensure program accountability (General Accounting Office [GAO] 2004). While several research studies have examined and documented the effectiveness, and cost benefits for *adult* EFNEP programs (Rajgopal et al. 2002; Joy et al. 2006; Wessman and Jensen 2002; Schuster et al. 2003; Dollahite et al. 2008), there is a dearth of research on the youth program, which constitutes more than 75% of program beneficiaries (more than 500,000 youth participants in year 2011). Townsend and colleagues (2006) are the only scholars to study and document the effectiveness of youth EFNEP to date. No cost effectiveness analyses of youth EFNEP programs exist.

Part of the reason for the dearth of evaluation studies on youth EFNEP has been the lack of comprehensive, valid, and reliable assessment tools to measure program effectiveness (USDA

2008). Further, unlike the adult EFNEP, no federal guideline mandates the use of any specific survey questions/instruments, or even curricula for youth EFNEP. For evaluation, the state programs often create their own survey instruments. This results in the lack of consistency and standardization across youth EFNEP programs and evaluation.

To address the challenges in evaluating the youth EFNEP programs, this project envisioned developing a cost effectiveness analysis (CEA) model for the youth EFNEP program in Virginia. A valid and reliable instrument to measure the program outcomes was developed and used in conducting the CEA. Results are presented in terms of the cost effectiveness ratios (CER) which expresses the amount of resources spent (in $) by the youth EFNEP program in producing the outcomes as dictated by the program goals. The project was guided, and closely monitored, by an advisory board consisting of nationally recognized faculty with expertise in nutrition, EFNEP, economics, item response theory, and program evaluation. Inputs from subject matter experts were instrumental to develop the perspective for the study and to enrich the validity and usability of this study.

The economic cost of a program in producing an impact is a basic question from a policy perspective. At times of economic pressure, education programs are easy targets for reducing the government spending. In such times, estimates of the CER could assist policy makers to prioritize the allocation of scarce resources. This study provides a concrete application and example of conducting a CEA for youth EFNEP (or other nutrition education programs) that other states/programs can consider applying to their own programs. EFNEP stakeholders could examine their programs utilizing this model in order to ensure that EFNEP youth nutrition programs are cost effective.

1.2. Goal and Objectives

The ultimate goal of this study is to create a cost effectiveness modeling procedure so that state and national level Extension faculty, administrators, and researchers can calculate and communicate the cost effectiveness of youth EFNEP programs (as well as other youth nutrition education programs).

This study demonstrates the use of the cost effectiveness model with a case study of the Virginia youth EFNEP. In this context, the specific objectives of this study are

- To measure costs of implementing youth EFNEP program in Virginia

- To measure effects of youth EFNEP program in Virginia

- To estimate the cost effectiveness ratio of the youth EFNEP in Virginia

1.3. Description of Virginia Youth EFNEP: Healthy Weights for Healthy Kids Program

The core of the youth EFNEP program in Virginia is the Healthy Weights for Healthy Kids (HWHK) curriculum, which is the foundation of this study.

Based on the Dietary Guideline for Americans, the HWHK curriculum was developed by Virginia Cooperative Extension [VCE] as a response to the childhood obesity epidemic and is targeted to the youths aged between 7 and 14 years (Serrano et al. 2007). The HWHK curriculum is based on the experiential learning model and contains the following six lessons with different activities possible for each lesson:

1. Smart Foods – Children use MyPyramid to learn about nutrition and to make smart food choices;

2. Smart Choices – Children explore different ways to enjoy eating, focusing on "quality" over "quantity," by measuring portion sizes of different meals and food products, planning a healthy plate or meal, eating slowly, and using different senses when eating;

3. Smart Drinks – Children investigate the amount of sugar and fat in popular beverages and can even learn how to make a nutritious homemade soda;

4. Smart Snacks – Children have the option of using the food label to compare snacks, making healthy "snackwiches," exploring what puts "whole" into whole grains, or creating advertising campaigns for snacks;

5. Smart Activities – Children get an opportunity to have fun and move in this lesson. They also can find out about calories and what inventions and discoveries have taken place that limit our country's need to be active; and

6. Smart Image – Children view, reflect, and discuss different media images and societal attitudes toward body size to improve attitudes and respect toward different sized and shaped individuals and to focus on "what is inside."

The curriculum is designed to be delivered in six one hour lessons, and cover activities from each of the six lessons. The curriculum was pilot-tested by EFNEP and the Supplemental Nutrition Assistance Program – Education Program (SNAP-Ed) (formerly Food Stamp Nutrition Education Program) educators and peer-reviewed before adoption.

The youth EFNEP program is delivered by para-professionals called *program assistants* (PAs). The PAs manage and deliver the HWHK curriculum to the participants in their designated

counties. First, PAs select the schools that have at least 50% of students eligible for free or reduced school lunch. In collaboration with local schoolteachers, PAs then identify the group of students' to participate in the program, and deliver the program. The curriculum is provided during school (considered "enrichment), in after-school care programs and through 4-H EFNEP clubs, day camps, residential camps, community centers, neighborhood groups, and home gardening workshops. Given the significant heterogeneity among youth participants (for example, differences in geographic location (rural and urban), and age range), the delivery of programs takes on various forms. The PAs, local to the community in which they serve, decide the content from within the lessons, order the content, and determine the frequency of meeting for class for each group of participants.

2. Methods

2.1. Cost Effectiveness Analysis

Cost effectiveness analysis is a common method used to evaluate and to compare the effects and costs of programs designed to improve health (Gold et al. 1996). Cost effectiveness analysis (CEA) adopted in this study is an alternative to the cost benefit analysis (CBA), commonly adopted for evaluating the adult EFNEP program. In CBA, benefits of health intervention are expressed in dollar terms rather than in terms of a nonmonetary effectiveness measure in CEA. Although the scope of application of CBA is broader than that of CEA, the CEA is also informative in stating the program success in achieving its goal. In fact, CEA are favored over CBA for evaluating health programs as it avoids the ethical concerns over converting the program benefits into monetary units (Gold et al. 1996).

7

The CEA describes a program in terms of the (incremental) cost effectiveness ratio (CER), which is expressed as

$$\text{CER} = \frac{\text{Total Cost}_A - \text{Total Cost}_B}{\text{Total Effects}_A - \text{Total Effects}_B}$$

where, A and B could be two different programs being compared. When the program is compared with no program (status quo), as is the case in this study, then the formula becomes

$$\text{CER} = \frac{\text{Total Cost}}{\text{Total Effects}} \qquad (1)$$

Equation (1) is the key formula used in this analysis. The numerator, total cost, refers to the dollar value of inputs used to design, deliver, and to maintain the program. The denominator, total effects, may be the targeted final outcomes of the program or could be intermediate outcomes (Gold et al. 1996) specific to the program's goal.

When estimating the CER using data from multiple units (such as different counties or programs), the estimates can be different depending upon the level of aggregation and averaging done in costs and effects data that comes from various units. In this study, CER derived by using equation 1 are expressed as total CER and average CER. The distinction between the two measures of CER is important at the outset to avoid confusion. Let TCER be the total CER, where the numerator is sum of all cost and the denominator is sum of all effects, over all counties. Let ACER be the average CER across counties, that is, the average of CER's estimated separately for each county. Let, n be the number of programs/counties included in the study, and $C_1, C_2, \ldots, C_n$ be the total cost incurred for programs/counties $1, \ldots, n$. Similarly, $E_1, E_2, \ldots, E_n$ be the effects of each county. By definition,

$$TCER = \frac{\sum_i C_i}{\sum_i E_i}, \text{ and } ACER = \frac{1}{N}\sum_i \frac{C_i}{E_i}.$$

Empirically, the estimates of TCER and ACER derived using the same data can either be greater than, equal to, or less than each other. In an attempt to disentangle the relationship between the two estimators, we proceed as follows:

$$TCER \lesseqgtr ACER$$

Or, $\frac{\sum_i C_i}{\sum_i E_i} \lesseqgtr \frac{1}{N}\sum_i \frac{C_i}{E_i}$

Some algebraic manipulations establish the following relationship

$$0 \lesseqgtr \sum_i \frac{C_i}{E_i}(\bar{E} - E_i) \quad [\because \bar{E} = \frac{\sum_i E_i}{N}]$$

The relationship between TCER and ACER is determined by the above expression. Because $C_i \geq 0$ and $E_i \geq 0 \; \forall \; i$, the ratio $\frac{C_i}{E_i}$ is always positive. The relationship between TCER and ACER thus depends on terms inside the parenthesis only, that is the difference between the average effects across all counties to effects of county i. The estimates of TCER and ACER will be equal only when $E_i = \bar{E} \; \forall \; i$. Whether TCER < ACER or TCER > ACER cannot be predetermined because it depends on whether $E_i > \bar{E}$, or $E_i < \bar{E}$. From basic mathematics, we know that E_i cannot be always greater than or less than $\bar{E} \; \forall \; i$. Even in cases where E_i is greater than or less than $\bar{E}$ for some i, the total sum of the product of the ratio $\frac{C_i}{E_i}$ and the mean deviation of effects can be greater or less than zero rendering the estimates of either TCER or ACER to be greater or less than each other.

In addition to the differences in TCER and ACER, the estimates of the CER[1] and its interpretations vary widely with the measures of costs and effects. In order to ensure the comparability of the CER across similar programs standard procedures to measure costs and effects were developed and used in this study.

2.2. Conceptualization of Costs and Effectiveness Measures

This study is a part of the project which envisioned developing a cost effectiveness model for the youth EFNEP program that could be used by youth EFNEP programs across the US and also by other similar programs. The whole project cycle was closely monitored and guided by the "advisory board" that consists of nationally recognized faculty with expertise in nutrition, EFNEP, economics, item response theory, and program evaluation. This is a popular mechanism for assuring content validity in the design of the instruments. A brief background on the formation of the advisory committee, and the process in which the project evolved provides a context for the discussion of the development of the instruments developed.

2.2.1. Formation of Advisory Committee and Organization of Workshop

A 2-day expert panel workshop was conducted at the beginning of the project, 11-12 May 2009, at Virginia Tech, Blacksburg, VA. The panel members consisted of 12 State and National experts who were representative of economics, nutrition, behavioral science, EFNEP, extension, and evaluation within the Extension context. Their inputs were instrumental in conceptualizing the project, developing instruments, interpreting findings, and guiding the different cost effectiveness approaches.

---

[1] The term CER is used to refer both TCER and ACER, unless specifically mentioned.

The main goal of the workshop was to gather and document opinions from these experts in order to create a cost effectiveness model. Specifically, the goals were to identify input variables (costs and effects) for estimating the CER and to identify/design valid, reliable, and sensitive instruments to measure these inputs. What constitutes costs and effects of youth EFNEP and how they should be measured were discussed in the workshop. Arguments for measurement of specific constructs were justified and passed by the panel members.

2.2.2. Decisions on Measuring Costs and Effects: The following decisions were made to precede the measurement of costs and effects.

- In order to ensure the comparability of the CER across common studies, the Panel on Cost effectiveness in Health and Medicine (PCEHM) recommends a methodological guide to conduct CEA (Gold et al. 1996). This guide was followed in this study in measuring both costs and effects as far as applicable.
- CEA are done with a perspective in mind, which is essentially determined by the purpose of the analysis. Because the key target audiences for the cost effectiveness model are EFNEP administrators, this study utilizes the program manager's perspective to CEA. CEA done from this perspective, although narrow, is appropriate in the present study where the program is being evaluated for its effectiveness in achieving its stated goals.
- Because lack of a universal instrument to measure the effects of youth EFNEP is one of the primary reasons for lack of evaluation studies on youth EFNEP, this project developed its own instrument to measure the effects of the youth EFNEP for CEA.

Details on conceptualization of cost effectiveness model for Virginia, and the decisions made in the workshop can be found in Serrano et al. 2011.

Time Frame for the Cost Effectiveness Analysis for Virginia

The timeframe for the cost effectiveness analysis was set to be one school calendar year, which ranged from September 2011 to May 2012. Costs data for the same period were collected from 15 EFNEP unit offices under which the program was implemented, and also from the State Office. Effects data were collected from the pre-post survey administered to all youth EFNEP participants in 15 counties of Virginia, during the same time period. The instruments used to collect both costs and effects data are available upon request.

2.3. Measuring Effects – the Denominator

The measures of program effects, the denominator in equation (1), used in CEA are usually guided by the purpose of the program. Given the *purpose of the youth EFNEP program is to improve participants' practice to healthy nutritional behavior,* change in behavior was identified as one of the primary measure of program effects. Other potential outcome measures such as self-efficacy and behavior intentions were also explored.

Different models of individual health behavior suggest that behavior change is mediated by variables such as self-efficacy and behavior intention (Glanz et al. 2008). Self-efficacy (SE) is the person's confidence in performing a particular behavior (Glanz et al. 2008; Pajeres 2010). It is a belief one has that leads to a particular course of action. Self-efficacy affects the amount of effort put by an individual in taking that particular action and thus the behavior. Behavior intention, on the other hand, measures a person's strength of intention to perform a behavior and is determined by one's attitude towards performing the behavior and his subjective norm associated with the behavior (Glanz et al. 2008; Ajzen 1991). The review of literature related to youth nutrition suggests that self-efficacy and behavior intention towards healthy nutrition are

the most commonly used predictors of nutrition behavior (Vries et al. 1988; Vries et al. 1995; Parcel et al. 1995; Reynolds et al. 1993; Vereecken et al. 2005; Lein et al. 2002; Baranowski et al. 2000; Contento 1995). In the context of CEA and the goals of the program, these are clearly intermediate outcomes.

Based on the empirical evidence and suggestions from the advisory board, self-efficacy towards healthy nutrition, behavior intention towards healthy nutrition, and nutrition behavior were selected as three indicators of youth EFNEP outcomes/effects. These outcome indicators are latent constructs, which cannot be measured directly. In such a situation, several items to measure the different dimensions of the construct are formulated. The responses to each item serve as indicators of latent constructs (Emberston 2010; Wilson 2010).

Given the above background, the first step of the analysis was to develop a new standardized instrument for the purpose of measuring these outcomes of youth EFNEP. Details on the process of instrument development to measure program effects are discussed next.

2.3.1 (Step 1): Development of Instrument to Measure Youth EFNEP Effects

2.3.1. a. Literature Review of Evaluation Instruments

The process of instrument development began with an extensive review of literature to identify valid, reliable and sensitive instruments on diet and physical activity for youth, focusing on instruments developed specifically for limited resource youth and Extension-delivered programs. A total of 15 instruments were reviewed, of which 10 were specific to the youth EFNEP.[2] Most of these instruments designed specifically for youth EFNEP evaluation were not tested (rigorously) for reliability and validity. The instruments focused only on some aspects

---

[2] A separate review article summarizing these instruments have been published (Hernandez-Garbanzo et al. 2013).

related to nutrition. More general type of instrument which measures multiple aspects of nutrition which would be particularly beneficial and practical for a broader set of programs such as EFNEP was lacking (Hermandez-Garbanzo et al. 2013).

As a starting point, a pool of potential question was developed based on the 15 selected existing instruments and was shared with the advisory board. Because, none of these instruments were specifically designed to capture the goals of youth EFNEP, the advisory board requested that the team develop a list of 'indicators' based on the Dietary Guidelines for Americans, to ensure that the instrument was comprehensive and 'could' possibly be used by other states.

2.3.1. b. Identification of Topics and Measurable Outcomes from the Dietary Guidelines

The Dietary Guidelines for Americans 2005 were reviewed and the entire lists of topics most appropriate for low-income youth ages 7 to 14 were selected. Out of over 30 guidelines addressing topics such as alcoholic beverages and sodium intake, 15 were selected. In addition, topic areas that the dietary guideline specifically encompassed were identified. The topics were presented to the advisory board who then voted the selected topics in terms of relevance priority, and its appropriateness for the population under study. In addition, advisory board members were asked to identify topics that they felt were implicit in the guideline and thus should be measured. The following topics emerged from the discussion:

- different food groups from MyPyramid,
- sugar-sweetened beverages,
- discretionary calories,
- physical activity,
- portion size,

- sodium intake.

The advisory board members were then asked to align the selected topics with the outcome measures such as knowledge, skills, attitudes, behavior, to understand how the topics would be best measured. The advisory board members later agreed that instead of knowledge, skills, attitudes, and behaviors, the evaluation instrument should measure behavior, self-efficacy, and behavior intentions. The choice of measure followed because the overall goal of nutrition education programs is to impact behavior, and prior research had established that self-efficacy and behavior intentions are two mediators that are strongly linked to behavior outcomes among youths (Cerin et al. 2009). Therefore, the decision was made that the selected dietary guidelines should be measured in terms of behavior, self-efficacy, or behavior intention outcomes.

2.3.1. c. Construction of the Evaluation Instrument

New questions were generated as well as the questions identified from the literature review were modified to address all of the selected topics from the Dietary Guidelines and framed in terms of self-efficacy, behavior intention, and behavior. Specifically, 111 questions were derived, covering the various topics. The advisory board was again asked to assess the questions. Given their feedback, questions with vague or inappropriate topics were removed from the list of questions. Definitions and clarifications were added when needed (i.e., definition for low-fat, etc.).

Lastly, the format of the response choice had to be determined. Two formats for asking questions and framing the response choices were common in the literature: i) "amount of food" and ii) "frequency of eating". For example, the "amount of food" type response asks students how many ounces of whole grain you ate yesterday. On the other hand, the "frequency of eating"

15

type question asks students, for example, how many times you ate cereals yesterday. Given the age of the target population (young children), tradeoffs between precision and accuracy of measurement by using both amount type and frequency type questions were discussed. It was later decided to get feedback from Virginia youth EFNEP Program Assistants (PAs), to confirm the appropriateness of the proposed indicators and provide guidance on the response categories. Four highly effective youth EFNEP PAs from Virginia were consulted about whether the questions should be framed in terms of asking about the "amount of food" or the "frequency of eating" the food. All four PAs agreed that the response should be measured by frequency of food (e.g. times per week) rather than the quantity of food (i.e. cup, ounces), given the level of comprehension of the target group (youths). The PAs shared their experience that very few children sit down and actually eat a given amount of food at one time. In addition asking for the quantity (amount of food) would require kids to do something abstract and add up amounts over the course of a day, and their estimates would likely be imprecise. Therefore, the PAs recommended using frequency of eating type response format over the amount of food type response format. The PAs also helped eliminate as well as add some questions as based on feedback.

2.3.2 (Step 2): Pilot-Testing of the Evaluation Instrument

A pool of 93 items was initially developed to measure self-efficacy, behavior intention, and behavior, each measured using 31 items. These items formed three separate sets of preliminary instruments. Similarly to other studies measuring self-efficacy, behavior intentions, and behavior, the items in all three instruments were the same except that the wording in the beginning of items was specific to the corresponding outcome.

16

In consultation with advisory board and program assistants in Virginia, the team recognized the response burden of 93 questions was too great for the targeted age group. Consequently the decision was made to pilot test all questions and to determine which were redundant and could be culled to a more feasible number.

Six youth EFNEP PAs pilot-tested the preliminary instrument with youth EFNEP participants in Virginia, during December 2010 to February 2011. Two sets of instruments were pilot -tested on each group of the participants over the course of two days. In the first day, all the participants were given the behavior component. On the second day, about half of the participants were given the self-efficacy survey and the other half of the participants received the behavior intention survey. A total of 272 surveys on behavior, 115 surveys on behavior intention, and 139 surveys on self-efficacy were completed.

2.3.3 (Step 3): Statistical Procedures for Culling Questions, and Reliability and Validity of Instrument

Data gathered from the pilot test was analyzed to select the best items from the preliminary instruments that would be used to develop the final instrument to measure the effects of the program. A conventional method based on classical test theory; item analysis, was conducted to select best items from the pool of items. Selection of items was later guided by another statistical method called exploratory factor analysis.

Item Analysis: Item analysis is a commonly used method in psychometrics for developing instruments. It plays an important role in construct validation. Item analysis extracts items that measure a uni-dimensional construct by exploring how the test takers respond to each item, and how the items relate to their overall performance (Nunnally and Bernestein 1994). The decision

about good and bad items is based on the item-total correlation statistics between the scores of an individual item and the total score on the test excluding that item. This measure checks whether a given item is consistent with the other items in the instrument. If a given item is consistent with the overall test, the item-total correlation should be high. A small item-total correlation means that the item is not measuring the same underlying construct as the rest of the items, thus should be revised or discarded. Item analysis can be used to determine the most valid questions for the development of a shorter, efficient questionnaire.

Using the software "jMetrik", separate analyses were done for the three sets of preliminary instruments: self-efficacy, behavior intention, and behavior. We started with all 31 items in the first iteration. As a rule of thumb, any items that do not correlate strongly ($r < .30$) with the total test score were regarded as bad items (Field 2005). Items with a negative or low correlation in each iteration were discarded until all the remaining items had an item total correlation > 0.3.

Among 31 items, 11 met the criteria of good items in each instrument. Good items that measured self-efficacy and behavior intention were all negative items, and the good items that measured behavior were all positive items.[3] The findings suggested that the preliminary instrument was probably measuring multiple constructs. Violation of uni-dimensionality would have been the cause for culling entire positive items in the self-efficacy and behavior intention instrument, and all negative items in the behavior instrument. Rather than do this, exploratory factor analysis (EFA) was used to explore the number of underlying constructs (factors) in the preliminary instrument.

---

[3] Negative items refer to questions about reducing the intake of unhealthy food, and positive items refer to questions about increasing the intake of healthy food.

Exploratory Factor Analysis: In EFA, all the items are modeled as linear combinations of potential factors that have common characteristics. The coefficient of each of the items, known as factor loadings, gives the relative contribution of an item to make a factor. Items with higher factor loadings are important variables for explaining a particular factor. If all the items have high correlations with only one factor, this provides evidence that the instrument is measuring a uni-dimensional construct. EFA not only identifies the number of latent factors but also identifies good items that measure the underlying construct. The purpose of EFA was thus twofold: first, to check dimensionality (that is whether the instrument is measuring one or more underlying latent constructs), and second, to select the best items to measure the underlying construct. Statistical software SPSS was used to perform EFA.

EFA for checking dimensionality: There are different, often debated, criteria in deciding the number of factors being measured by a set of instrument. According to Kaiser's criterion, any component with eigenvalue >1 is considered as a separate factor. The eigenvalues when graphed against the factors with which it is associated, a scree plot, shows the relative importance of each factor. The scree plot provides useful information regarding the number of important factors being measured. The cutoff point for selecting the number of important factors is at the point of inflection of the scree plot (Field 2005). The scree plots for all three preliminary set of items derived using pilot test data show that there are two main underlying factors for each of the preliminary instruments, which provides evidence that each instrument measured more than one constructs. To further explore the multi-dimensionality of the instrument, and for item selection, exploratory factor analysis was conducted.

EFA for item selection: The relationship between the items and the factors, are measured as factor loadings. The factor loadings can be thought of as correlation coefficients or sometimes as regression coefficients between a factor and the items. The interpretability of factors can be enhanced by the technique called rotation (Field 2005). An oblique rotation was used which allows the underlying factors to be correlated. The higher the factor loading, the more variance in the factor is explained by the items. In selecting the good items, items with factor loadings > 0.3 were sorted as it is typical to consider factor loadings with this threshold value (> 0.3) to be important (Field 2004).

Exploratory factor analysis using the pilot data extracted two main factors for each outcome: self-efficacy, behavior intention, and behavior. Most of the positive items appeared to be constituents of one factor, while most of the negative items of another factor. A few items did not correlate (had r < 0.3) with any of the two factors. This result was consistent with all three sets of preliminary instrument.

Based on these results, selection of good items based on the item analysis may not be a sound method for measuring multi-dimensional latent constructs. Therefore, results from the factor analysis were used to select items for the final instrument based on their factor loadings. All items with factor loadings >0.3 were regarded as 'good' and used in the final instrument. Unfortunately, the resulting instrument however would still contain more than 50 items. It was recognized that the high number of items would substantially increase the response burden to the target group. After another review of the literature and a prolong discussion with "advisory board" members, the following decisions were made that finalized the instrument.

- Because a review of instruments that measure the effects of youth nutrition education program in the literature revealed that self-efficacy is more widely used than behavior intention, the behavior intention indicators were discarded from the final instrument.

- Most studies published in nutrition education focus on positive dimensions of self-efficacy. Therefore, negative items were discarded.

The final instrument included only two indicators: self-efficacy, and behavior. The numbers of total items were reduced to 35 items: 19 items to measure behavior, and 16 items to measure self-efficacy. In addition, questions regarding participants' demographic information, and food availability at home were collected. The final instrument is available upon request to authors.

Measuring Total Effects of the Program

A common and practical method to measure the effects of an educational program is to use a *pre* and *post*-test approach. The effects instruments were administered *pre* and *post* intervention to measure changes in responses for each participant. Improvements in individual participant's responses have to be amassed to generate the total effects for the program which is required to calculate the CER (denominator of equation 1). Two alternative approaches to generate total effects are possible.

- **Score based measures**: This measure uses the raw scores of individuals on all items and measure the change in raw scores from pre to post test. Standardized tests, such as GRE and SAT, are common examples of score based measures, which have been used as a measure of educational achievement (for example, Congressial Budget Office (CBO), 1987).

21

- **Count based measures**: This measure counts the number of individuals satisfying some criteria. These are very common in the education field. The "No Child Left Behind" uses the percentage of students achieving proficiency in a given grade and subject; National Assessment of Education Progress uses the percentage of students meeting or exceeding the national standards, both of which are basically the count based measures (Virginia Department of Education, 2012; US Department of Education National Assessment of Title I, final report, 2007).

Simple algebra reveals that improvements in total scores does not map one-to-one with changes in numbers of individuals improving in outcomes. In fact, it is possible for the total (or average scores) to improve while the percentages of individuals improving actually decreases. Because the objective of the program is to improve numbers of individuals (counts) not scores, the count based measure was favored over the score based measure and this decision approved by the advisory board. The count based measure is used in other areas, such as education and poverty (for example, Heck 2006) and is standard in the poverty literature (for example, Jolliffe et al. 2005). It is also consistent with the current practice of the USDA in measuring EFNEP outcomes which counts the number of individuals who improve on *at least one item* among a list of items. Effectively, in this study, the denominator of the CER is measured by two measures: 1) the total numbers of participants whose response in at least one of the 19 items measuring behavior improved, and 2) the total numbers of participants whose response in at least one of the 16 items measuring self-efficacy improved.

2.4. Measuring Costs – the Numerator

The underlying principle in measuring costs of the youth EFNEP program is that the estimates of costs should reflect the value of resources that are used in the provision of the

program (HWHK). Only the direct costs associated with implementing the program are considered. Indirect costs borne by program participants are ignored because the cost effectiveness analysis is done from program managers' perspective. The direct costs of implementing the HWHK program are divided into four main categories: 1) Labor Cost; 2) Capital Cost; 3) Material Cost; 4) Energy Cost.

Each cost category is a composite of multiple components. The cost attributed to the HWHK should be reflective of the share of the HWHK program within the EFNEP and the Virginia Cooperative Extension (VCE) under which the program are usually implemented. Thus, adjustments are made where appropriate. Further, the costs are prorated for months in which the program is actively implemented within the time frame for which the CEA is done.

2.4.1. Labor Cost

The cost related to the human resources directly involved in delivering and managing the youth EFNEP are considered as labor cost. Youth EFNEP program assistants (PAs) are the key personnel involved in the management and the implementation of the HWHK program. The EFNEP administrators and program support staffs are also involved in managing the HWHK program, but their contribution to the HWHK is relatively small. Therefore, labor costs associated only with the youth EFNEP PAs are considered. Labor costs include expenditures on salaries and benefits, travel, and training.

2.4.1. a. Salaries and Benefits

Salaries and benefits consist of the compensation provided to the PAs in the form of salaries and fringe benefits. Some youth PAs are responsible for the youth EFNEP only, while others are responsible for both the youth and the adult EFNEP program. Consequently, the share

of time allocated to adult programs is not included. Furthermore, within the HWHK, PAs spend their time in various sub-activities such as recruiting youths, studying and preparing for the classes, preparing food for demonstration, teaching/delivering the HWHK curriculum, traveling, and filling out forms for administrative purposes. Within the youth EFNEP, the youth PAs deliver other curricula, for example, Literacy, Education, Activity Program (LEAP), Organ Wise Guys, Teen Cuisine, and Professor Popcorn, in addition to the HWHK curriculum in Virginia.

The share of salaries and benefits attributed to the HWHK program are reflective of the share of time spent by each PA on the provision of the HWHK curriculum only. Further, salaries and benefits are only measured for months in which the actual HWHK program is implemented, that is, only for months active in the HWHK. Thus, the annual salaries and benefits of the youth PAs are adjusted for the share of their work time spent on the HWHK, and for months active in the HWHK program. The total labor cost for salaries and benefits attributed to the HWHK are derived as follows:

Cost of salaries and benefits[4] = *Annual salaries and benefits of the youth PAs * Share of PA work time spent on the HWHK * Proportion of months active in the HWHK program*

2.4.1. b. Training

Training costs include costs associated with the training of the youth EFNEP PAs. It consist of cost associated with training PAs after their initial recruitment into the youth program, and cost associated with attending all trainings/workshops, related to the HWHK program, attended by the youth PAs within the time frame in which the program evaluation is done.

---

[4] Source: State EFNEP office and PAs bimonthly time allocation survey

> **Cost of training**[5] = *Costs associated with youth PAs initial training after recruitment + Costs associated with attending trainings/workshop/meetings during the time frame of program evaluation*

2.4.1. c. Travel

Travel costs include costs associated with PAs travel in order to manage and implement the HWHK program. Some of the common causes for travel include school visit for student recruitment, purchase of supplies, curriculum delivery, attending meetings and workshop. The product of the numbers of miles traveled by PAs in their personal vehicles in order to deliver the program and the rate of reimbursement for travel yields the travel cost incurred for each PA.

> **Cost of travel**[6] = *Numbers of miles traveled by PAs in their personal vehicles in order to deliver the HWHK program * Rate of reimbursement for travel cost per mile*

2.4.2. Capital Cost

Capital costs are costs associated with the capital resources directly used in support of delivering and managing the HWHK program. Only the costs of capital resources used at the county level are considered. This is because the share of capital attributed to the HWHK is relatively small at the EFNEP offices higher in the hierarchy (such as district level and state level). Office space and equipment (for example, computers and copiers) are the main capital resources used for the HWHK.

2.4.2. a. Office Space

---

[5] Source: State EFNEP office
[6] Source: State EFNEP office

Office space includes the cost of office space used by the staffs directly involved in the HWHK program. The HWHK program staffs at the county level are housed at the EFNEP unit offices. Most of the EFNEP unit offices are usually housed within the Cooperative Extension offices. The Cooperative Extension office building is often shared by multiple agencies. Most of the Cooperative Extension office buildings are owned by the local government and some of them are rented. The measurements of cost of office space for the office housed in the owned and in the rented buildings are different, hence are dealt separately.

2.4.2. a. 1 Owned Office Space

The cost of owned office space is derived from the current market value of the building in which the program is housed. The Office of Management and Budgeting (OMB) Circular 87 on "Cost Principles for State, Local and Indian Tribal Governments" states that for buildings owned by local government agencies, the cost of space should be computed as an annual rate not exceeding 2% of total costs of acquisition. Following the guideline, only 2% of the current value of the building is used as the cost of office space for a given year.

The cost of office space attributed to the youth EFNEP program should be reflective of the office space actually used by the youth program only. Further, the value of office space should only be considered for months active in HWHK. Thus, the market value of the office building in which the program is housed is adjusted by the share of office space used by the program, and for months active in the program.

Cost of Owned Office Space[7] = *Current market value of building where EFNEP office is housed *annual usage fee as established by the OMB * the percentage of building space*

---

[7] Source: Unit offices (cost survey done with Unit Office Incharge); and PAs bimonthly time allocation survey

*occupied by EFNEP * percentage of EFNEP space occupied by youth EFNEP *Proportion of*

months active in the HWHK program

2.4.2. a. 2 Rented Office Space

The cost of rented office space for the program is derived from the value of annual rent paid for the building in which the program is housed. Similar to the owned office space, the cost of office space attributed to the program should reflect the office space used by the youth EFNEP program. Also, the cost of office space should be reflective of months active in the HWHK program. The value of monthly rent paid for the entire office building in which the program is housed, for the months active in the HWHK, is adjusted by the share of office space used by the program.

Cost of Rented Office Space[8] *= Annual rent of the building where EFNEP is housed * the percentage of building space occupied by EFNEP * percentage of EFNEP space occupied by youth EFNEP * Proportion of months active in the HWHK program*

2.4.2. b. Equipment

Equipment cost includes the value of equipment such as computers, printers, fax machines, copiers, VCRs, cameras used by the youth EFNEP staffs for the HWHK program. Only the values of equipment at the unit offices are considered. The OMB guideline (Circular A 87) allows 6.67% of the total cost of equipment to be used each year. Following the guideline, only 6.67 % of the cost of equipment is used as annual cost of equipment.

---

[8] Source: Unit offices (cost survey done with Unit Office Incharge); and PAs bimonthly time allocation survey

Some of the equipment is solely used by the youth EFNEP, whereas other equipment is shared by the adult and the youth program. For equipment that is not solely used by the youth EFNEP, the cost attributed to the program is adjusted by the share of equipment used for the youth EFNEP. Further, the cost of equipment is adjusted for months active in the HWHK.

Cost of Equipment[9] = *Sum over all equipment type (total cost of equipment * Share of equipment used for the youth EFNEP *Annual usage fee as established by the OMB* Proportion of months active in HWHK)*

2.4.3. Material Cost

Material costs are costs associated with any material resources used in support of delivering the HWHK program. Two types of materials used for the HWHK are identified: supplies (for example, food) and printed materials.

2.4.3. a. Supplies

The cost of supplies includes the cost of food items purchased by the PAs for demonstration use in support for the delivery of the HWHK curriculum. PA's are reimbursed for the expenses incurred in purchasing the supplies for the program.

Cost of Supplies [10] = *Cost of supplies purchased by PA for the demonstration purpose for the HWHK for months active in HWHK.*

2.4.3. b. Printing

---

[9] Source: Unit offices (cost survey done with Unit Office Incharge); and PAs bimonthly time allocation survey
[10] Source: State Office (As reported in the "green system")

Printing cost includes the cost of printing the evaluation instruments, handouts and any other materials used in support of implementing the HWHK program. The number of students who receive each type of printed materials multiplied by the unit cost of printing those materials yields the cost of printing.

Cost of Printing[11] = *Number of youth EFNEP participants * unit cost of printing*

2.4.4. Energy Cost

Energy costs are costs associated with the usage of energy in the EFNEP unit offices, in support of the HWHK program. It mainly consists of the cost of utilities.

2.4.4. a. Utilities

The cost of utilities includes the value of money spent on utilities such as phone, electricity, water, sewage, garbage, heating oil/gas and janitorial or maintenance services. The cost of utilities is derived from the monthly bills paid for each utility type. Utility bills are usually available for the entire building in which the youth EFNEP is housed. Energy costs attributed to the program should reflect the share of utilities consumed by the youth EFNEP program. Because the share of energy use is difficult to measure, the share of office space occupied by the youth EFNEP program is used as a proxy for the share of utility usage. The costs of utilities are also adjusted for months active in the HWHK program. The adjusted monthly utility bills for each utility type are summed over months active in the HWHK to get the cost of utilities.

---

[11] Source: State Office

Cost of Utilities[12] = *Sum of the monthly cost of utilities of each type, for months active in HWHK, for the entire building in which youth EFNEP is housed * percentage of building space occupied by EFNEP* percentage of EFNEP space occupied by youth EFNEP*

Measuring Total Costs of the Program

Costs under each category are measured separately for individual unit offices (county offices from where the program is delivered) and added to get the total cost of program for each county. Total cost for each county is then summed across all unit offices to get the total cost of the youth program in Virginia.

Total Cost of the Program = *Sum over all counties (Labor Cost + Capital Cost + Material Cost + Energy Cost)*

3. Results and Discussion

To demonstrate the application of CEA tools and methods developed in section 2, data from youth program in Virginia are used. The results discussed below are derived from the costs and effects data collected from the following 15 counties in Virginia: Appomattox, Carroll, Charles City, Culpeper, Hampton, King William, Loudoun, Lynchburg, Norfolk City, Patrick, Richmond City, Rockingham, Scott, Westmoreland, and Wise.

3.1. Costs of Implementing Youth EFNEP- HWHK in Virginia

The total costs of implementing the youth EFNEP- HWHK program in 15 counties of Virginia during the school calendar year 2011/2012 was estimated to be $134,333 (Labor =

---

[12] Source: Unit offices (cost survey done with Unit Office In-charge); and PAs bimonthly time allocation survey

$112,558; Capital = $7,250; Material = $5163; Energy = $9362). Cost of labor (salaries and

benefits including travel and training) had the largest share of expenditure for the program

(84%), followed by energy cost (7%), capital cost (5.4%), and material cost (3.8%). The costs of

implementing the program in each of the 15 counties are summarized in Table 1. The average

total cost of implementing a youth EFNEP program in one county is $8,956. The range of total

costs across programs varied from $1,982 to $16,825. Similarly, the average cost related to labor

was $7,504 (range $1,806 - $16,673); the average cost related to capital was $604 (range $30 -

$2,283); the average cost related to material was $344 (range $27 - $1,651); the average cost

related to energy was $780 (range $4 - $2,119).

[Place Table 1 Approximately Here]

Cost data for some of the cost headings were not available for some counties, therefore

not included in total cost calculation. Richmond and Carroll counties did not provide cost data

for office space, equipment and utilities. Hampton County did not provide cost data on

equipment. The youth EFNEP office at Norfolk is housed in some other governmental building,

and the EFNEP is exempt from paying money for office space and utilities. So, youth EFNEP

program in Norfolk incurs zero cost related to office space and utilities (rather than missing

information). Loudoun County reported that the equipment used for the youth EFNEP program is

provided by the State and that the program does not spend on equipment. Missing data on cost is

expected not to significantly change the total cost estimates because the cost share of capital cost

and utility cost is relatively small for all programs in general. The impact of missing cost data in

part will be handled in the sensitivity analysis.

3.2. Effects of Youth EFNEP-HWHK in Virginia

The effects of youth EFNEP-HWHK program were measured by two outcome indicators: change in nutrition related behavior, and change in nutrition related self-efficacy. Changes in behavior were measured by the total number of individuals who improve in at least one of 19 behavior items, from pretest to posttest. Similarly, changes in self-efficacy were measured by the total number of individuals who improve in at least one of 16 self-efficacy items, from pretest to posttest. These measures of effects are consistent with the current practice of the USDA in measuring total effect of the EFNEP.

A total of 2,566 students were reached by the program. Of students who were reached by the program, 73% of them (1,864) were evaluated. The rest of the students were excluded from the study because either they did not complete one or both of the pre/post surveys, or their pre/post surveys did not match based on the identifiers used.

One of the key assumptions while using change in behavior and change in self-efficacy as measures of effects is that no factors other than the program itself affects change in participants' behavior or self-efficacy. Given that the design of the program is not experimental, the influence of factors other than the youth EFNEP program on the outcome measures cannot be ruled out. In an attempt to control for some of these factors, the probability of a change in outcome (probability of participants improving in at least one item) was modeled as a function of participant characteristics (gender, age, ethnicity, family type, geographic location), and program characteristics (class size, class type, number of volunteers assisting the class, duration of program, number of handouts distributed, PA's age and experience) (Table 2). Predicted probabilities for each participant only associated with the program characteristics were then generated by setting the coefficients on variables related to student characteristics equal to zero,

which is equivalent to predicting the probability just associated with the program effects. The sum of these predicted probabilities across all individuals would then give the predicted number of people who improved controlling for the other confounding effects, at least those included in the model and these are called the predicted effects. Results using the predicted effects are compared with the results using the observed effects (actual improvements in behavior).

[Place Table 2 Approximately Here]

Change in Behavior and Self-efficacy:

The effects for each of the 15 counties are summarized in Table 3. On average, 171 participants were reached by the HWHK program in each county (range 26 - 434) of which about 124 students were evaluated (range 11 – 296). The average observed improvements in behavior was 119 (range 11 – 29), and the average observed improvements in self-efficacy was 119 (range 10 – 285). The predicted improvements were slightly higher. Of 1,864 students evaluated, 1,786 students (96%) improved in at least one behavior (observed effects) from pretest to posttest. About 1,782 students (96%) improved in at least one self-efficacy (observed effects) items from pretest to posttest. Accounting for the factors outside of program's control, the predicted effects are higher indicating that confounding factors may be offsetting program effects for some individuals. For behavior the predicted effects (improvements) are 1819 and for self-efficacy 1853.

[Place Table 3 Approximately Here]

3.3. The Cost Effectiveness Ratio

Using the data on costs and effects collected from 15 counties, the CERs for youth EFNEP-HWHK in Virginia were computed and are reported in Table 4. The total CER (TCER)

per participant reached by the program in Virginia was $52. For both outcomes, improvements in behavior and self-efficacy, the TCER per outcomes (actual effects) were on average $75. Using the predicted effects, the TCER per behavior improvement was estimated to be $75, and the TCER per self-efficacy improvement was $73.

The average CER (ACER) per participant reached is $89 (range $20 – $243). The ACER per improvement in behavior is $149 (range $24 – $537), and the ACER per improvement in self-efficacy is $153 (range $25 – $591). The values of ACER estimated by using the predicted improvements are slightly less. Estimates of ACER are higher than the estimates of TCER. The condition in which ACER is equal to TCER is only when the effects are equal for all counties (conditions derived in section 2), which is not the case here. The difference between the estimates of ACER and TCER basically reflects the variability in effects across counties.

[Place Table 4 Approximately Here]

3.4. Sensitivity Analysis

The estimates of the CER presented above are the point estimates and may suffer uncertainties. The uncertainties in the estimates of costs and effects could arise due to missing data for some cost components, and also due to possible measurement errors, which must be accounted for. It is a common practice to exhibit the uncertainty of point estimates by constructing confidence intervals around the estimates. Confidence interval captures the variability in the estimate, the likelihood that the sample drawn from the same population will yield the estimated value. Bootstrapping is a common method of choice to determine the variability of a ratio estimator (Chaudhary and Stearns 1996; Briggs et al. 1999; Polsky et al. 1997). The standard errors and the confidence intervals of the estimated TCER were derived by

34

using a non-parametric bootstrapping method. First a sample of 15 sets of costs and effects were drawn from all observations of costs and effects using a simple random sampling with replacement. The bootstrap replicate of the TCER was then computed from the drawn sample. The method was repeated for 10,000 times to obtain 10,000 independent replications of bootstrap TCER. The observed distribution of the standard error and the confidence interval of bootstrap TCER are presented in Table 5. The sampling distribution of the TCER estimates from each bootstrap replications are plotted in Figure 1.

The estimated range of 95% confidence interval is $56 to $115 and $55 to $114 for TCER per improvement in observed and predicted behavior, respectively. Similarly, the 95% confidence interval for TCER per improvement in observed and predicted self-efficacy is $56 to $116 and $53 to $111, respectively (Table 5). The observed values of TCER for both outcome measures lie within the estimated confidence range of TCER.

[Place Table 5 Approximately Here]

[Place Figure 1 Approximately Here]

3.5. Cross-County Comparison of CER

The point estimates of the CER by county serve as a starting point for making a comparative analysis in identifying the most cost effective program, and are given in Figure 2. Three measures of effects for each county are shown. The first bar (blue bar) measures the cost per student reached, which is simply the average cost per participant reached by the program. Second and third bars give the cost per improvement for behavior and self-efficacy, respectively.

For each county, the numerators for all three outcomes are same and what varies among the three is the measure of effects.

[Place Figure 2 Approximately Here]

The graph clearly shows a huge variation in the estimates of CER across counties. Counties with high CER per reach indicate that such counties (counties 5, 9, and 15) spend more dollars per participant. This could be because such counties either have high program cost or they recruit only a small number of participants. The within county gap in the estimates of CER per reach and CER per improvements is staggering for some counties, and this is an indication of program ineffectiveness. On average, the CER per improvement is about 1.7 times the CER per reach. For county 2 the CER per improvement is only 1.2 times the CER per reach. On the other hand, for county 9, the CER per improvement is about 2.4 times the CER per reach. For counties 1-4, 6-8, and 11 the cost per reach is not very different than cost per improvement. For counties 9 and 5 the cost per improvement is much higher than cost per reach. Large discrepancy between CER per reach and CER per improvements means that relatively few program participants actually improve their behavior. The smaller the discrepancy, the more effective the program is in achieving the goal of improving the behavior of its participants.

So far, in estimating the CER, what counted as an improvement for an individual was that they improved on *at least one item*. The USDA improvement threshold for the adult EFNEP using the 10 item behavioral checklist is 1, and is followed in the current study to be consistent with the USDA measure. One can argue that the improvement measure used by the USDA (and consequently this study) has a very low bar in improvement threshold, and it completely ignores the degree of improvements achieved by the program. The improvement thresholds can be set

36

higher by redefining *the number of items an individual improves on*. The improvement threshold could be any number between 0 and the total number of questions in the survey. For example, there are 19 items that measure behavior, and one could calculate the number of people who improved on 0 questions, 1 question, 2 questions, up to all 19 questions. Increasing the *improvement threshold* has implications on the estimates of CER. The total CER estimated by setting different thresholds on improvements are given in Table 6 (and Figure 3).

[Place Table 6 Approximately Here]

[Place Figure 3 Approximately Here]

The numerator of CER (total costs), used in estimating CER in Table 6, remains the same for all threshold values. Only the denominator (total effects) is changed by increasing the improvement thresholds. Obviously the CER estimate increases with higher threshold which sets more stringent criteria to be counted as improvements. The estimates of CER shows that rate of increase in CER is not linearly related to the increase in threshold criteria. The CER increased from $75 per improvement to $78 per improvement when threshold increased from 1 to 2, whereas the CER increased from $209 to $306 when threshold increased from 7 to 8. The increase in CER is small at smaller values of thresholds, however, it takes offs rapidly past threshold 5. For higher threshold values, the increase in CER is much higher for self-efficacy improvements compared to behavior improvements.

In a nutshell, we can conclude that setting more stringent thresholds on improvements rapidly increases the CERs, and is more so for higher thresholds. This finding has important policy implications for improving program effectiveness. By increasing threshold, only the individuals who improve in multiple dimensions related to nutrition are counted as program

37

effects (improvements). Also, the improvements in more than one dimension can be achieved with only a small increase in costs, for example, by increasing costs from $75 to $78, gains in outcomes or program effectiveness can be doubled (from 1 to 2 improvements). On the other extreme, if the program aims to improve the behavior in all 10 dimensions (threshold = 10), the program cost will increase to a very large extent ($736 per improvement). The problem the policymakers now face is the obvious tradeoffs between the cost and the desired program outcomes. If the policy objective is to secure improvements in multiple dimensions, then the program will certainly incur higher costs. Deliberation however is to be made regarding the choice of optimal threshold for improvements.

4. Conclusions

A model for evaluating the cost effectiveness analysis (CEA) for youth EFNEP program is developed in this study. The challenges due to lack of a general, valid, and reliable tools for measuring outcomes of the youth EFNEP was surpassed with the help of the advisory board members. The instrument developed in this study is general enough to capture the various aspects of nutrition education and is based on the Dietary Guidelines for Americans. This makes it readily usable for evaluating other youth EFNEP programs across the nation. The procedures and methods developed to measure both costs and effects of youth EFNEP in this study opens up avenues for other youth nutrition education programs in other states to conduct cost effectiveness analysis of their respective programs. The procedure for conducting CEA is applied to Virginia youth EFNEP program for demonstration.

The cost effectiveness analysis of youth EFNEP program (HWHK curriculum) in

Virginia was done for one school year (2011/12). From a program managers' perspective, data

on cost related to the management and implementation of the program were collected from 15

cooperative extension unit offices under which youth EFNEP was implemented, in the given

time frame. Data on effects were collected from all participants participating in the youth EFNEP

program in the same timeframe from all 15 counties. Total costs and total effects of the program

were compared. The total costs of the program were estimated to be $134,333, and the total

effects of the program measured by the number of participants whose behavior improved was

1786 (predicted 1819), and that measured by the number of participants whose self-efficacy

improved was 1782 (predicted 1853).

The total CER (TCER) for the youth EFNEP program in Virginia was estimated to be

around $75 per improvement in behavior and about the same for improvement in self-efficacy.

Now, how good the estimated figures of CER are is unknown because these are the first

estimates of its kind for youth EFNEP. Cross county comparisons reveal a wide variation in the

estimated CER across counties. The choices of thresholds on improvements are also shown to

have important implications on the estimates of CER.

In a separate paper (Baral et al. 2013), we estimated the maximum average cost of

improving *adult* EFNEP participant's behavior on three outcome indices: food resource

management practices (FRMP), nutrition practices (NP), and food safety practices (FSP). We

found that, for Virginia, the maximum average CER to be $579, $544, $718 per improvement in

FRMP, NP, and FSP, respectively. While these were the maximum average cost estimates for

adult program for Virginia, these are the best available information that could be used for

comparison. It is safe to say that the youth EFNEP program in Virginia with much lower

estimates of CER ($75 per improvement) are highly cost effective compared to the adult program in Virginia.

With a tightening federal budget, there is an increasing pressure to identify and use cost effective methods to achieve the stated goals of the youth EFNEP. There can be no discussion and comparison of cost effectiveness without estimates. This study developed the tools and procedures to estimate the cost effectiveness of youth EFNEP and estimated the CER for Virginia. The tool and procedures developed in this study can be utilized by other youth EFNEP programs to generate information on the cost effectiveness for respective programs. Results then can be used by the state and national EFNEP administrators to demonstrate program success, for comparing the cost effectiveness across states and ultimately to improve EFNEP efficiency.

References

Ajzen, I. 1991. Theory of Planned Behavior. *Organizational Behavior and Human Decision Process,* 50: 179-211.

Baranowski, T., M. Davis., K. Resnicow., J. Baranowski., C. Doyle., L.S. Lin., M. Smith., and D.T. Wang. 2000. Gimme 5 Fruit, Juice, and Vegetables for Fun and Health: Outcome Evaluation. *Health Education & Behavior*, 27 (1): 96-111.

Baral, R., G. C. Davis., S. Blake., W. You., and E. Serrano. 2013. Using National Data to Estimate Average Cost Effectiveness of EFNEP Outcomes by State/Territory. *Journal of Nutrition Education and Behavior*: 45(2): 183-187.

Briggs, A.H., C.Z. Mooney., and D.E.Wonderling. 1999. Constructing Confidence Intervals for Cost-Effectiveness Ratios: An Evaluation of Parametric and Non-Parametric Techniques Using Monte Carlo Simulation. *Statistics in Medicine*, 18: 3245-3262.

Cerin, E., A. Barnett., and T. Baranowski. 2009. Testing Theories of Dietary Behavior Changes in Youth Using Mediating Variable Model with Intervention Programs. *Journal of Nutrition Education and Behavior*, 41:309-318.

Chaudhary, M.A., and S.C. Stearns. 1996. Estimating Confidence Intervals for Cost-Effectiveness Ratios: An Example from a Randomized Trial. *Statistics in Medicine*, 15: 1447-1458.

Congress of the United States, Congressional Budget Office. 1987. Educational Achievement: Explanations and Implications of Recent Trends. Available at http://www.cbo.gov/sites/default/files/cbofiles/ftpdocs/62xx/doc6204/doc13b-entire.pdf Accessed September 25, 2012.

Contento, I., G.I. Balch., Y.L. Bronner., L.A.. Lytle., S.K. Maloney., C.M. Oslon., and S.S. Swadener. 1995. The Effectiveness of Nutrition Education and Implications for Nutrition Education Policy, Programs, and Research: A Review of Research. *Journal of Nutrition Education*, 27:277-418.

Dollahite, J. S., D. Kenkel, and C.S. Thompson. 2008. An Economic Evaluation of the Expanded Food and Nutrition Education Program. *Journal of Nutrition Education and Behavior*, 40:134-43.

Embretson, S.E. 2010. Ed. Measuring Psychological Constructs: Advances in Model-Based

    Approaches. 1st edition. *American Psychological Association*. Washington, DC.

Field, A. 2005. Discovering Statistics Using SPSS, 2nd Ed. London: Sage Publications.

General Accounting Office (GAO).2004. Nutrition Education: USDA Provides Services through

    Multiple Programs, but Stronger Linkages Among Efforts are Needed. Washington, DC:

    United States General Accounting Office.

Glanz, K., B. K. Rimer., and K. Viswanath. 2008. Eds. Health Behavior and Health Education:

    Theory, Research, and Practice, 4th edition. San Francisco, CA: John Wiley and Sons.

Gold, M., J. Siegel, L. Russell, and M. Weinstein. 1996. Eds, Cost-effectiveness in health and

    medicine, ed. O.U. Press. New York.

Heck, R. H. 2006. Assessing School Achievement Progress: Comparing Alternative Approaches.

    Educational Administrative Quarterly.42: 667-699.

Hemandez-Garbanzo, Y., J. Brosh., E.L.Serrano., and R. Bhattarai. 2013. Psychosocial Measures

    Used to Assess the Effectiveness of School-based Nutrition Education Programs: Review and

    Analysis of Self-report Instruments for Children 8 to 12 Years Old. *Journal of Nutrition*

    Education and Behavior, (Article in Press).

Jolliffe, D., C. Gundersen., L. Tiehen., and J. Winicki. 2005. Food Stamp Benefits and Child

    Poverty. *American Journal of Agricultural Economics*, 87: 569-581.

Joy, A. B., V. Pradhan., and G. Goldman. 2006. Cost-Benefit Analysis Conducted For Nutrition

    Education in California" *California Agriculture*, 60:185-191.

Lien, N., L. A. Lytle., and K. A. Komro. 2002. Applying Theory of Planned Behavior to Fruit and Vegetable Consumption of Young Adolescents. *American Journal of Health Promotion*, 16(4):189-97.

Nunnally, JC., and I.H. Bernstein. 1994. Psychometric Theory, 3rd ed. New York: McGraw-Hill.

Office of Management and Budgeting. 1997. Cost Principles for State, Local and Indian Tribal Governments (Circular A 87). Washington, DC.

Pajeres, E.2010. Overview of Social Cognitive Theory and of Self-Efficacy. Available at: http://www.emory.edu/EDUCATION/mfp/eff.html . Accessed on 16 May 2010.

Parcel, G.S., E. Edmundson., C. L. Perry., H. A. Feldman., N. O'Hare-Tompkins., P.R. Nader., C. C. Johnson., and E.J.Stone.1995. Measurement of Self-Efficacy for Diet-Related Behaviors Among Elementary School Children. *Journal of School Health*, 65(1):3-40.

Polsky, D., H. A. Glick., R. Willke., and K. Schulman. 1997. Confidence Intervals for Cost-Effectiveness Ratios: A Comparison of Four Methods. *Health Economics*, 6: 243-252.

Rajgopal, R., R. H. Cox., M. Lambur., and E.C. Lewis. 2002. Cost-Benefit Analysis Indicates the Positive Economic Benefits of the Expanded Food and Nutrition Education Program Related to Chronic Disease Prevention. *Journal of Nutrition Education Behavior*, 34:26-37.

Reynolds, K.D., A.W. Hinton., R.M. Shewchuk., and C.A. Hickey. 1999. Social Cognitive Model of Fruit and Vegetable Consumption in Elementary School Children. *Journal of Nutrition Education*, 31(1): 23-30.

Schuster, E., Z. L Zimmerman., M. Engle., J. Smiley., E. Syversen., and J.Murray.2003.Investing in Oregon's Expanded Food and Nutrition Education Program (EFNEP): Documenting Costs and Benefits. *Journal of Nutrition Education Behavior,* 35: 200-6.

Serrano, E., S. Diehl., K, Jamison., J, Shelhamer., J, Wages., and K., Gehrt. 2007. Healthy Weights for Healthy Kids: How Virginia Cooperative Extension is Addressing Childhood Overweight, in Food, Nutrition, and Health. *Virginia Cooperative Extension,*.348-275

Serrano, E., M. McFerren., M. Lambur., M. Ellerbock., K. Hosig., N. Franz., M. Townsend., S. Baker., P. Muennig., and G. Davis. 2011.Cost-Effectiveness Model for Youth EFNEP Programs: What Do We Measure and How Do We Do It?" *Journal of Nutrition Education and Behavior*, 43 (4): 295-302.

Townsend, M. S., M. Johns., M. K. Shilts., and L. Farfan-Ramirez.2006. Evaluation of a USDA Nutrition Education Program for Low-Income Youth. *Journal of Nutrition Education and Behavior,* 38: 30-41.

United States Department of Agriculture (USDA). 2013.National Institute of Food and Agriculture. *Expanded Food and Nutrition Education Program*. Available at http://www.nifa.usda.gov/nea/food/efnep/efnep.html Accessed February 26, 2013.

USDA Cooperative State Research, E., and Extension Service. 2008. NC_TEMP211 Multi-State Team, EFNEP Related Research, Program Evaluation and Outreach: Statement of Issue(s) and Justification.

U.S. Department of Education. National Assessment if Title I Final Report. October 2007. Available at http://ies.ed.gov/ncee/pdf/20084012_rev.pdf Accessed September 25, 2012

Vries, H. D., E. Backbier., G. Kok., and M. Dijkstra. 1995. The Impact of Social Influences in the Context of Attitude, Self-Efficacy, Intention, and Previous Behavior as Predictors of Smoking Onset. *Journal of Applied Social Psychology*, 25(3): 237–257.

Vries, H. D., M. Dijkstra., and P. Kuhlman. 1988. Self-efficacy: the Third Factor Besides Attitude and Subjective Norm as a Predictor of Behavioral Intentions. *Health Education Research*, 3(3):273-282.

Vereecken, C.A., W.V. Damme., and L. Maes. 2005. Measuring Attitudes, Self-Efficacy, and Social and Environmental Influences on Fruit and Vegetable Consumption of 11- and 12-Year-Old Children: *Reliability and Validity*. 105(2):257-261.

Virginia Department of Education. 2012. Adequate Yearly Progress Reports. Available at http://www.doe.virginia.gov/statistics_reports/accreditation_ayp_reports/ayp/index.html Accessed September 25, 2012.

Wessman, C. B., and H. Jensen. 2002. An Evaluation of the Costs and Benefits of Iowa's Expanded Food and Nutrition Education Program. Final Report., in PM1865a, *I.S.U. Extension*, Editor.

Wilson, M. 2010. Constructing Measures: An Item Response Modeling Approach. New York, NY: Routledge.

Table 1: Cost of Youth EFNEP by County

County ID	Labor Cost		Capital Cost		Material Cost		Energy Cost		Total Cost
	Cost (in $)	Cost share	Cost (in $)	Cost share	Cost (in $)	Cost share	Cost (in $)	Cost share	(in $)
1	13410	90	247	2	389	3	872	6	14919
2	11097	93	124	1	355	3	385	3	11961
3	9005	67	2283	17	482	4	1579	12	13348
4	5363	92	56	1	316	5	121	2	5855
5	6199	80	493	6	80	1	960	12	7732
6	6017	91	98	1	382	6	81	2	6578
7	1806	91	NA	NA	176	9	NA	NA	1982
8	8788	69	647	5	1651	13	1636	13	12722
9	1891	32	1869	32	27	0	2119	36	5905
10	6699	98	NA	NA	116	2	NA	NA	6816
11	9851	88	97	1	779	7	443	4	11169
12	2755	96	30	1	76	3	4	0	2865
13	16673	99	NA	NA	152	1	NA	NA	16825
15	4684	87	206	4	107	2	382	7	5379
16	8322	81	1100	11	74	1	781	8	10278
Total Cost	**112,558**	**84**	**7,250**	**5**	**5,163**	**4**	**9,362**	**7**	**134,333**
Average Cost	**7,504**		**604**		**344**		**780**		**8,956**

Table 2: Results from Logistic Regression Models Used to Generate Predicted Outcomes

Variables	Variable Definition	Improvements in Behavior		Improvements in Self-efficacy	
		Coef.	SE	Coef.	SE
Gender	0 = Female; 1 = Male	0.061	0.238	-0.425*	0.233
Age	Age in years	0.036	0.110	-0.121	0.102
Ethnicity	1 = African American; 2 = White; 3 = Others	0.077	0.163	0.120	0.153
Family Type	0 = Single Parent; 1 = Both Parent; 2 = Others	0.101	0.197	0.068	0.188
Location	1 = Rural; 2 = Subruban; 3 = Cities	-0.754***	0.211	-0.468***	0.171
Class size	Number of students in a class	-0.022***	0.008	0.000	0.005
Group Type	0 = Single Grade; 1 = Mixed Grade	-0.995**	0.470	-0.632*	0.363
Volunteers	Number of volunteers assisting the class	0.987***	0.341	0.134	0.217
Duration	Number of days between pretest and posttest	-0.007***	0.002	0.007**	0.003
Handouts	Number of handouts given to the participants	0.117***	0.044	-0.009	0.031
PA Age	Age of PA in years	-0.071***	0.020	-0.002	0.013
PA Experience	Experience of PA in years	-0.042	0.039	-0.017	0.047
Constant		7.137***	1.587	4.882***	1.352
LL		**-294.187**		**-313.256**	
Pseudo R2		**0.092**		**0.069**	

*Significant at 10%; ** Significant at 5%;*** Significant at 1% levels.*

Table 3: Effects of Youth EFNEP by County

County ID	Number of participants		% Evaluated	Change in Behavior		Change in Self-efficacy	
	Reached	Evaluated		Observed	Predicted	Observed	Predicted
1	378	296	78	295	294	285	294
2	325	274	84	263	265	264	273
3	314	260	83	247	251	259	259
4	102	86	84	80	84	81	85
5	33	21	64	20	21	20	21
6	321	230	72	207	226	209	228
7	97	82	85	82	82	80	82
8	120	103	86	103	102	101	102
9	26	11	42	11	11	10	11
10	60	38	63	38	38	36	38
11	434	246	57	242	241	239	245
12	52	30	58	30	29	29	30
13	148	93	63	83	91	79	92
15	104	55	53	46	50	52	55
16	52	39	75	39	35	38	39
Total Effect	**2,566**	**1,864**	**73**	**1,786**	**1,819**	**1,782**	**1,853**
Average Effect	**171**	**124**		**119**	**121**	**119**	**124**

Table 4: The Estimates of Cost Effectiveness Ratio's by County

| County ID | Student Reached | Cost Effectiveness Ratio (in $) per | | | |
| | | Change in Behavior | | Change in Self-efficacy | |
		Actual	Predicted	Actual	Predicted
1	39	51	51	52	51
2	37	45	45	45	44
3	43	54	53	52	52
4	57	73	70	72	69
5	234	387	370	387	372
6	20	32	29	31	29
7	20	24	24	25	24
8	106	124	125	126	124
9	227	537	537	591	540
10	114	179	181	189	180
11	26	46	46	47	46
12	55	95	97	99	96
13	114	203	186	213	183
15	52	117	108	103	99
16	198	264	293	270	264
Total CER	**52**	**75**	**74**	**75**	**73**
Average CER	**89**	**149**	**148**	**153**	**145**

Table 5: Derivation of Confidence Interval for the CERs using Bootstrap

(Replications = 10000)

| Outcomes | CER | | | | |
	Observed	Bootstrap	Bias	SE	[95% CI]*
Behavior (Actual)	75.21	77.79	2.58	15.29	$56.04 to $115.26
Behavior (Predicted)	73.83	76.37	2.53	15.07	$54.85 to $113.57
Self-efficacy (Actual)	75.38	78.01	2.62	15.55	$56.18 to $116.11
Self-efficacy (Predicted)	72.50	74.95	2.45	14.69	$53.99 to $111.35

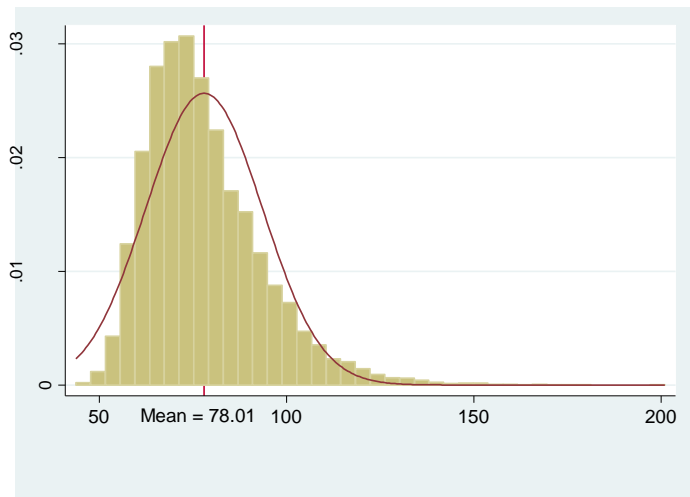**Table 6: CER by Improvement Threshold**

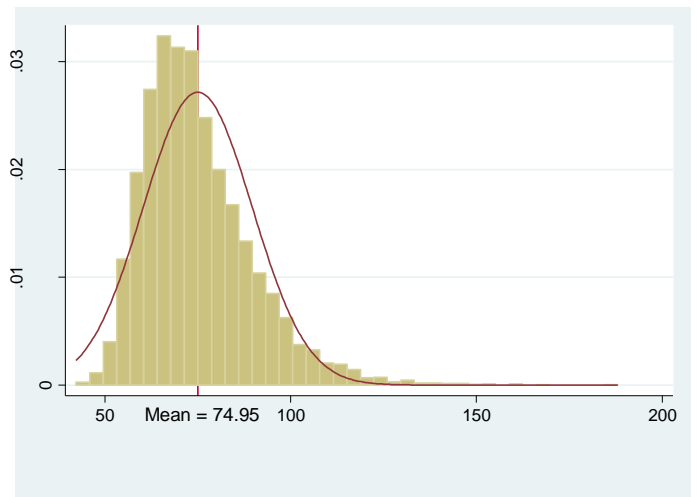| Thresholds on Improvements	CER per Behavior Improvement	CER per Self-efficacy Improvement
1	75	75
2	78	82
3	84	96
4	94	119
5	114	156
6	148	226
7	209	340
8	306	551
9	452	933
10	726	1658
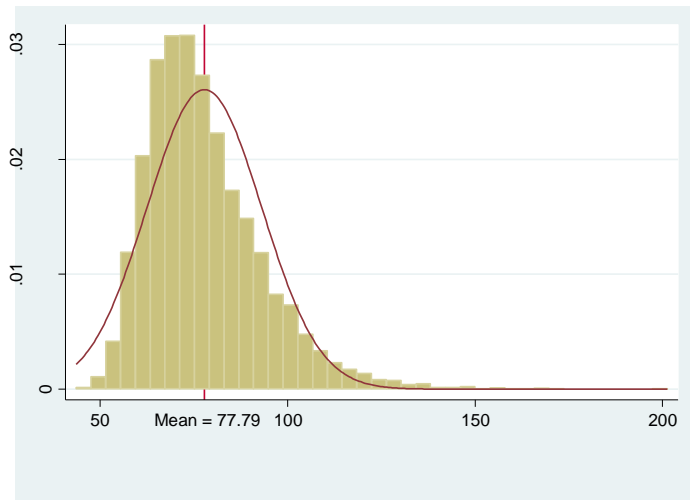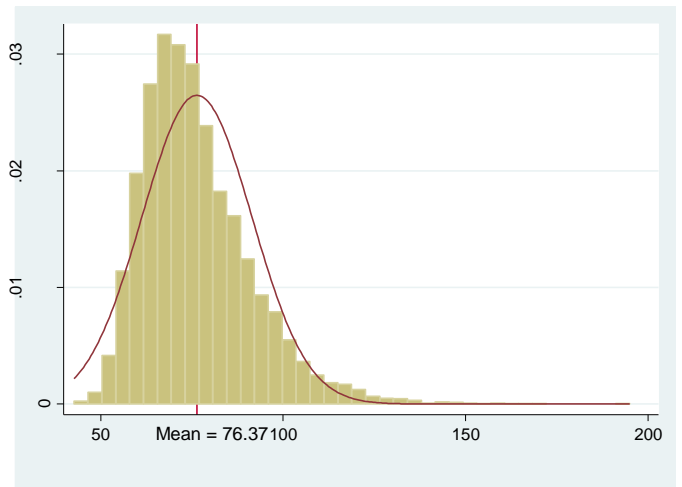
Figure 1. Histograms of 10000 bootstrap replicates of the cost effectiveness ratio estimates, overlaid by the normal density of same mean and variance. Figures (a) and (b) use actual and predicted improvements in behavior as effects measure, respectively. Figures (c) and (d) use actual and predicted improvements in self-efficacy as effects measure, respectively.
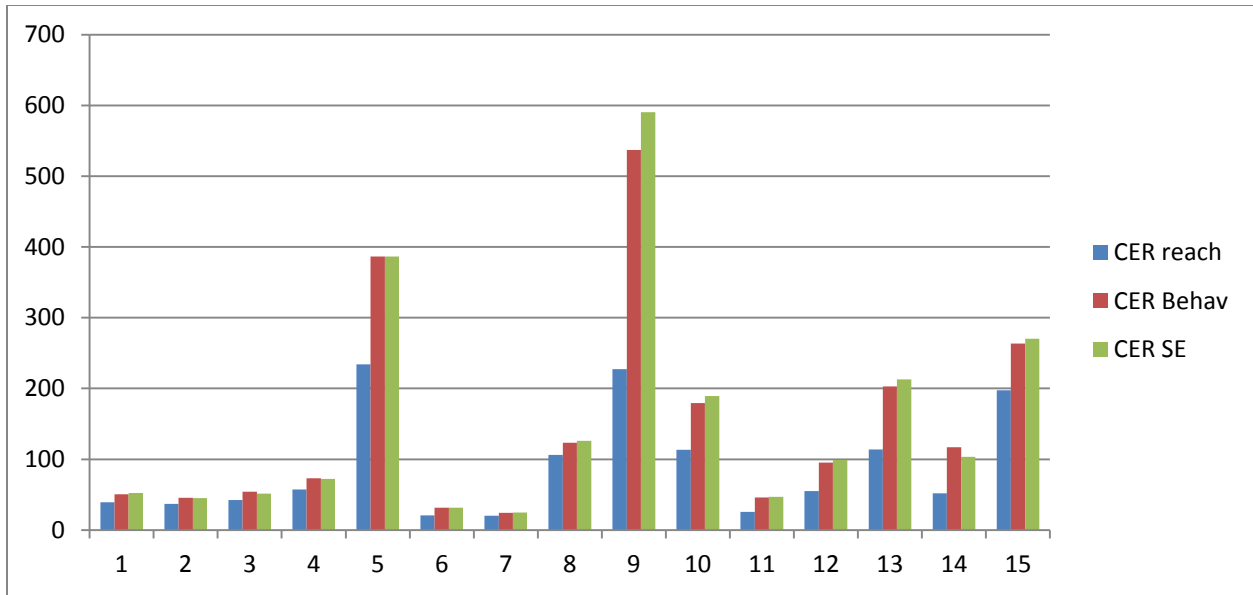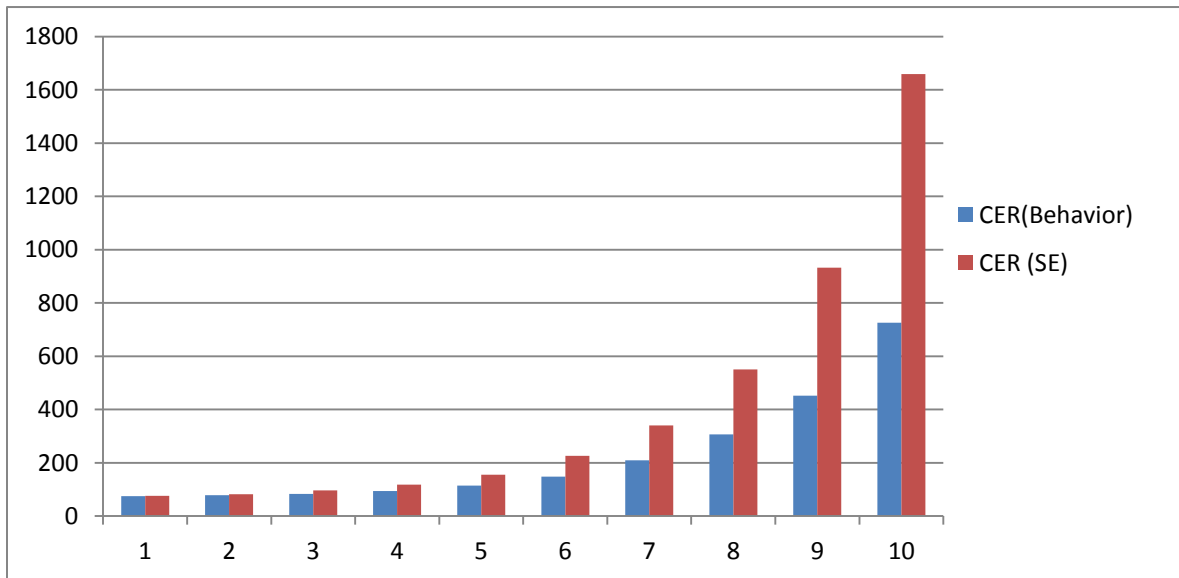
Figure 2. Cost Effectiveness Ratios by County

Figure 3. Cost Effectiveness Ratios by Threshold on Improvements