

National, Regional, and State-Level Estimates of Returns to Scale in the Expanded Food and Nutrition Education Program

Ranju Baral, George C. Davis, and Wen You

The effectiveness of the Expanded Food and Nutrition Education Program in achieving its goals at the national, regional, and state level is unknown. Using US Department of Agriculture (USDA) data from all states and territories for the years 2000–2006, the impact of program and participant characteristics and returns to scale on the three outcome indicators used by the USDA are estimated. Program and participant characteristics do not seem to be as important as the amount of money spent on the program. Generally speaking, there are constant and increasing returns to scale for two of the three federal outcome indices for most states but not all.

Key Words: EFNEP, evaluation, program effectiveness, returns to scale

JEL Classifications: I12, I18, I38

Nutrition education programs are a common policy tool for improving nutrition and consequently public health. Lower socioeconomic status (SES) populations are usually targeted because of the well-established positive relationship between SES and health quality (e.g., Marmot and Wadsworth, 1999). 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 US Department of Agriculture (USDA) (General Accounting Office, 2004). In operation for more than 40 years

and now in all 50 states and six territories, the EFNEP has become a cornerstone in US nutrition education (USDA, 2009a).

The aim of the EFNEP is to assist limited-resource audiences to “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, 2009a). The EFNEP is administered by the USDA Cooperative Extension Service to youths and adults. The Adult EFNEP serves adults 19 years and older and is the focus in this article.

Within each state, the State Cooperative Extension staffs provide leadership in the program and choose and/or develop the curriculum. The curriculum primarily focuses on improving various aspects of nutrition ranging from participants’ ability to make healthy food choices, skills in food preparation, food safety and sanitation to managing a food budget. The State Cooperative Extension staffs provide training

Ranju Baral is a PhD candidate, and Wen You is an assistant professor, Department of Agricultural and Applied Economics, Virginia Tech, Blacksburg, Virginia. George C. Davis is a professor, Department of Agricultural and Applied Economics, and Department of Human Nutrition, Foods, and Exercise, Virginia Tech, Blacksburg, Virginia.

Appreciation is extended to Stephanie Blake for data assistance and to Prashanta Bhattarai for his help in making figures.

to the paraprofessionals (also called program assistants [PAs]) and volunteers, who are usually local to the community they serve, and are responsible for recruiting the participants and delivering the EFNEP curriculum. The curriculum is delivered as a series of lessons, at least six as required by the USDA, over several months. The number of lessons and the topics covered in the lessons are usually tailored to the needs of the program participants and usually vary across the states. The program is delivered to individuals or to a group of individuals depending on the need and convenience.

Although the USDA spends approximately \$66 million per year on the EFNEP (USDA, 2009b), very little is known about the impact of this money on achieving its stated objectives at the national level. There are a few single-state, single-year analyses for subsets of EFNEP participants (e.g., Arnold and Sobal, 2000; Dickin, Dollahite, and Habicht, 2005; Dollahite and Scott-Pierce, 2003) and the general finding is that the EFNEP improves nutrition knowledge and behavior, but these analyses cannot be generalized to a national level. At the national level, the only information provided is by the USDA's EFNEP "impact data" annual report, in which impact is measured as the percentage of participants showing improvement from a pretest to a posttest on questions related to food resource management practices, nutrition practices, and food safety practices (USDA, 2009a). The impact data suggest most participants tend to improve in these domains. However, the impact data are only summary statistics on the outcomes with no multivariate analysis associating these results with dollar expenditures and program/participant characteristics.

Although the few state-level results are encouraging and suggestive, there are still several outstanding questions related to the effectiveness of the money spent on the EFNEP program at the national level: 1) Are there program or participant characteristics that can be identified as contributing to the stated objectives? 2) Does the money spent on the EFNEP contribute to the stated objectives of the program? 3) What are the returns to scale associated with the EFNEP money in achieving its objectives at the national level as well as at the state/territory

level? 4) Are there geographical differences across states and regions in the returns to scale measures?

Using the USDA impact data for all states and US territories for 7 consecutive years (2000–2006), but in conjunction with several available covariates, we answer these questions using a nonlinear multivariate regression model. Results show that only a few program and participant characteristics are significant in explaining changes in the outcome indices used by the USDA. The money spent on the EFNEP is the most important factor and has a positive and significant impact on two of the three outcome indices. Using a purely statistical significance criterion, there is constant returns to scale at the national level for two of the three indices (food resource management practices and nutrition practices), whereas three of the five regions show increasing returns in the food resource management index. Approximately 40% of states/territories show increasing returns to scale in one of the two indices, and the remaining states mainly show constant returns to scale in both the indices. The third index (i.e., food safety practices) shows decreasing returns to scale regardless of the level of analysis. A broader perspective based on economic significance vs. statistical significance suggests that a majority of the states have increasing returns to scale in two of the three indices. These findings are the first to quantify the budget effects at the national level and also help in identifying potential sources in which cost-effective information may be gleaned. Limitations are discussed in the conclusions.

Conceptual Framework

When evaluating publicly funded projects, Shephard (1974) advocates the use of an indirect production function approach. The indirect production function comes from maximizing output subject to a cost constraint as opposed to the common dual problem of minimizing expenditures subject to a production constraint that generates the cost function. In a public funding setting such as the EFNEP, the constraint is the amount of money provided to

operate the program and given that amount of money, the goal is to maximize the output, which in this case would be positive outcome measures of the program. Thus, the indirect production function can be written as $Y = f(B, \mathbf{Z})$, where Y is a measure of output, B is the budget amount, and \mathbf{Z} is a vector of other covariates that may affect the outcome such as participant and program characteristics. Given the functional form that will be used, the marginal effect of the budget on the output turns out to be a measure of returns to scale, as will be shown.

Data

The USDA collects data on the EFNEP by state/territory through its Nutrition Education Evaluation and Reporting System (NEERS) (USDA, 2009c). The NEERS is a software program that allows administrators at the county and state level to collect and report data to the USDA in aggregate form on adult participation. Given the structure of the program, the budget effects potentially interact with participant and program characteristics so the analysis must incorporate these potential confounding factors. In general, the NEERS collects data on participant characteristics (e.g., age, education, and income), program characteristics (e.g., type of lesson, type of instructor), and most importantly responses to 10 required behavioral checklist questions for adults. The total EFNEP budget for each state/territory is available as well. The data covers all 50 states and six US territories for the fiscal years 2000–2006.

Dependent Variables

We use the same measurements developed by the USDA as our dependent variables. Specifically, the USDA uses 10 behavioral checklist questions to form the basis for evaluating the success of the EFNEP. The 10 behavioral checklist questions were developed over a span of 5 years from 1993–1997 and involved several phases of development that assured content and face validity (see Anliker, Willis, and Montgomery, 1998, for details). The major phases were 1) a committee formed the questions

based on reviews of other existing and submitted instruments; 2) feedback was solicited on the instrument from the EFNEP coordinators in all 50 states; 3) more feedback on the instrument was solicited from a larger pool of specialists; 4) an Expert Panel met in Washington, DC, to revise the instrument; 5) the revised instrument was subjected to focus group testing and tested for reading level (determined it was at the 6th-grade reading level); 6) based on feedback from the focus group, the instrument was revised and pilot-tested in seven states; 7) statistical analysis was conducted on the pilot-tested instrument and the instrument was checked for validity and internal reliability; and 8) additional minor revisions were made based on the validity and reliability analysis to improve and finalize the instrument.

The 10 behavioral questions are given in Table 1 and responses are based on a 5-point Likert-type scale (e.g., 1 = do not do, 2 = seldom, 3 = sometimes, 4 = most of the time, 5 = almost always). Participants answer these behavioral questions both pre- and post-EFNEP participation. USDA uses subsets of the 10 questions to create three composite measures of different aspects of nutrition knowledge: the Food Resource Management Practices (FRMP) composite is constructed from answers to questions 1–4, the Nutrition Practices (NP) composite is constructed from answers to questions 1 and 7–10, and the Food Safety Practices (FSP) composite is constructed from answers to questions 5 and 6. Improvement in a composite occurs when the score on at least one of the composite questions increases from the pretest to the posttest.

Table 2 lists the variables' definitions and summary statistics. Following the approach implemented by the USDA, the EFNEP impacts are measured in terms of the percentage of participants showing an improvement for each composite, so consequently these percentages by definition range from 0 to 100. We will refer to the percentages as indices: the FRMP index, the NP index, and the FSP index. On average, the value of the FRMP index is 83% (i.e., 83% of the participants improved in the FRMP composite), the value of the NP index is 88%, and the value of the FSP index is

Table 1. Behavior Checklist Question Components of Nutrition Knowledge Indices

	Behavior Checklist Questions	Indices
1	How often do you plan meals ahead of time?	FRMP, NP
2	How often do you compare prices before you buy food?	FRMP
3	How often do you run out of food before the end of the month?	FRMP
4	How often do you shop with a grocery list?	FRMP
5	How often do you let these foods (meat and dairy) sit out for more than 2 hours?	FSP
6	How often do you thaw frozen foods at room temperature?	FSP
7	When deciding what to feed your family, how often do you think about healthy food choices?	NP
8	How often have you prepared foods without adding salt?	NP
9	How often do you use "Nutrition Facts" on the label to make food choices?	NP
10	How often do your children eat something in the morning within 2 hours of waking up?	NP

FRMP, Food Resource Management Practices; FSP, Food Safety Practices; NP, Nutrition Practices.

64%. The corresponding standard deviations are FRMP 9%, NP 6%, and FSP 11%.

Covariate Explanatory Variables

Budget. This variable is defined as the annual budget allocation (in real US dollars) for each state and US territory. The USDA's annual nominal budget allocation for EFNEP is based on the following exogenous rule¹: 4% is used for federal administration, 10% is equally distributed among the states, and the remainder is distributed based on states'/territories' population under 125% of income poverty as determined by the last preceding decennial census. The allocations (2000–2006) were converted into real 2005 dollar values using the Consumer Price Index. The average amount over all observations (i.e., states/territories and years) is \$1,091,583 with a standard deviation of \$961,058 (Table 2). We hypothesize that controlling for the other covariates, as the budget increases, the index scores will improve.

Number of Participants and Participant Characteristics. We include in the analysis the number of adult participants to control for the size of the program. The average number of participants is approximately 2,880 with a standard deviation of 4,200 (Table 2). To

control for the effects of the youth EFNEP program on the outcomes of the adult program, we also included the number of youth participants in the analysis. The average number of youth participants is approximately 7,645 with a standard deviation of 9,100.

We include several available sociodemographic categorical characteristic percentages of the participants that are collected by the USDA: gender, race/ethnicity, age, educational level, household income, and place of residence as defined in Table 2. These socioeconomic variables are typical of those found in the literature on health behavior (e.g., Pollard et al., 2009). Including these participant characteristics will allow us to determine if, conditional on the amount of money being spent, the program is more effective for one participant characteristic subgroup over another. A finding of differential participant characteristic effects would suggest that targeting certain participant subgroups could lead to a more effective program. When disaggregated by gender, females constitute 89%, and by race, whites and blacks account for approximately 65% of the participants on average. More than two-thirds of the participants ($\approx 68\%$) fall in the 20- to 39-year age range and approximately 63% have income below 100% of the poverty line. Unfortunately, approximately 72% of participants did not provide information on their level of education, but for those who did, the largest percentage is for those with education only through some high school ($\approx 23\%$). Less

¹This rule was revised in 2008 but is relevant for the time for which the data are used in this study.

Table 2. Variable Descriptions and Summary Statistics

Variable	Description	Mean (SD)
Dependent variables		
FRMP index	Percentage of participants who improved in one or more Food Resource Management questions	82.72 (8.62)
NP index	Percentage of participants who improved in one or more Nutritional Practice questions	87.91 (5.85)
FSP index	Percentage of participants who improved in one or more Food Safety Practice questions	64.44 (11.40)
Covariates		
Budget	Annual federal budget allocation per state/territory (in real dollar units)	1,091,583 (961,058.10)
Participant	Number of participants in a given state for a given year (1,000s)	2.88 (4.20)
Youths	Number of youth participants in a given state for a given year (1000s)	7.64 (9.10)
Female	Percentage of female participants	89.29 (8.20)
Male ^a	Percentage of male participants	10.70 (8.21)
White	Percentage of white participants	40.34 (26.21)
Black	Percentage of black participants	25.02 (25.79)
Hispanic	Percentage of Hispanic participants	19.65 (23.01)
Other race ^a	Percentage of participants from other race	14.97 (27.30)
Age Under 20 ^a	Percentage of participants with age younger than 20 years	13.11 (7.54)
Age 20–29	Percentage of participants with age 20–29 years	45.56 (18.34)
Age 30–39	Percentage of participants with age 30–39 years	22.71 (11.18)
Age 40–49	Percentage of participants with age 40–49 years	10.58 (6.67)
Age 50–59	Percentage of participants with age 50–59 years	3.62 (3.52)
Age 60+	Percentage of participants with age older than 60 years	2.26 (4.51)
Age NA	Percentage of participants without age information	2.16 (3.79)
Income 50% of poverty level ^a	Percentage of participants whose household income is less than or equal to 50% of poverty level	36.75 (19.22)
Income 100% of poverty level	Percentage of participants whose household income is between 51–100% of poverty level	25.69 (10.98)
Income 150% of poverty level	Percentage of participants whose household income is between 101–150% of poverty level	8.82 (5.96)
Income 150+% of poverty level	Percentage of participants whose household income is more than 150% of poverty level	4.34 (4.85)
Income NA	Percentage of participants for whom household income information is not available	24.39 (22.37)
Edu NA	Percentage of participants without information on education status	71.82 (36.99)
High school ^a	Percentage of participants with education up to high school	22.67 (30.21)
Some college	Percentage of participants with some college-level education	4.56 (6.94)
College graduate and up	Percentage of participants who graduated from college or above	1.70 (2.92)
Rural farm	Percentage of participants who reside in rural farm area	2.98 (8.70)

Table 2. Continued

Variable	Description	Mean (SD)
Rural nonfarm	Percentage of participants who reside in town with population under 10,000 and rural nonfarm area	26.82 (24.89)
Small town	Percentage of participants who reside in town and cities with population 10,000–50,000 and their suburbs	21.34 (16.04)
Suburb of cities	Percentage of participants who reside in suburb of cities with population over 50,000	5.34 (6.51)
Central cities ^a	Percentage of participants who reside in central cities with population above 50,000	43.52 (29.07)
Group lesson ^a	Percentage of participants who were delivered group lessons	72.70 (23.25)
Individual lesson	Percentage of participants who were delivered individual lessons	20.53 (18.63)
Both group and individual lesson	Percentage of participants who were delivered both group and individual lessons	5.59 (8.20)
Other lesson	Percentage of participants for whom other types of instruction were used	1.17 (3.69)
Professionals	Percentage of professionals among the instructors	2.05 (4.91)
Paraprofessionals	Percentage of paraprofessionals among the instructors	12.67 (17.88)
Volunteers ^a	Percentage of volunteers among the instructors	85.27 (19.77)

^a Variables used as base in the model. Mean is calculated as the overall average across states and territories for all 7 years. SD, standard deviation; FRMP, Food Resource Management Practices; NP, Nutrition Practices; FSP, Food Safety Practices.

than half of participants ($\approx 43\%$) reside in central cities with populations over 50,000.

Program Characteristics. One may expect different levels of key program characteristics to have different impacts on learning outcomes; consequently, the analysis includes information collected by the USDA regarding types of lessons and instructors. EFNEP offers four lesson types: 1) individual lessons; 2) group lessons; 3) both individual and group lessons; and 4) other types. A state-level analysis by Dollahite and Scott-Pierce (2003) shows that the participants who took individual lessons showed more positive behavior change than those who took group lessons. This seems a very intuitive finding and one that has important implications for how the program should be structured to achieve the largest impact if the result holds at a national level. Although individual lessons may be more effective in terms of individual outcomes, they are not necessarily more cost-effective. The tradeoff that must be considered is the implied decline in

effectiveness as students are added to a lesson vs. the decrease in the cost per student of the lesson. If the decline in the effectiveness is offset by a greater decline in cost, then group lessons would be more cost-effective. The largest percentage of participants attend group lessons ($\approx 73\%$) followed by individual lessons ($\approx 21\%$), group and individual lessons ($\approx 6\%$), and other lesson types ($\approx 1\%$).

Three types of instructors are involved in the program: professionals, paraprofessionals, and volunteers. The professionals train the paraprofessionals and volunteers and provide technical support. Paraprofessionals are the key personnel responsible for delivering the curriculum. Volunteers primarily assist in program delivery and program management. Being mostly from the local area, the paraprofessionals and volunteers are expected to be best suited to deliver lessons to local participants. However, if paraprofessionals and volunteers are not as effective as professionals, then this would imply a reallocation of resources toward more professionals may be more effective in achieving

the goals of the program. Volunteers constitute the largest percentage of instructors ($\approx 85\%$) followed by paraprofessionals ($\approx 13\%$) and professionals ($\approx 2\%$).

Empirical Model

The FRMP index, the NP index, and the FSP index are the percentages of individuals improving in the corresponding composite; so by definition, each is restricted to the range 0–100%. To account for this restricted range, we use a logistic functional form for each equation (Kmenta, 1997). Mathematically, the three equations to estimate have the form

$$(1) \quad Y_i = 100 / (1 + e^{-\alpha_{0i} - \alpha_{1i} \ln B - \alpha_{2i} Z}) + \varepsilon_i$$

$i = \text{FRMP, NP, FSP,}$

where Y denotes the nutrition index (percentage improvement), $\ln B$ is the natural log of the budget, Z the vector of control variables, ε the error term with an expected value of zero, and the α s are conformable parameters. Equation (1) corresponds to a simple indirect production function.

An important feature of this model is that it is nonlinear and allows for increasing, constant, or decreasing returns to scale in the budget. Specifically, the change in each dependent variable for a 1% change in the budget (i.e., a one-unit change in $\ln B$) is the marginal budget effect (MBE), or mathematically from equation (1),

$$(2) \quad MBE_i = \partial Y / \partial \ln B$$

$$= \alpha_{1i} \times \left[100 \times e^{-\alpha_{0i} - \alpha_{1i} \ln B - \alpha_{2i} Z} / (1 + e^{-\alpha_{0i} - \alpha_{1i} \ln B - \alpha_{2i} Z})^2 \right]$$

The sign of the MBE is determined by the sign of the parameter on the budget, α_{1i} , because the term in brackets is always positive. The magnitude of this effect will vary depending on the value of the budget and other covariates (i.e., the point of evaluation). This formulation will allow for varying returns to scale by state and territory and consequently by region. Furthermore, because the dependent variable is a percentage and the budget is expressed in natural logs, simple calculus reveals that this marginal effect is equal to the ratio of average cost (AC_i) to marginal cost (MC_i) or

$$(3) \quad MBE_i = AC_i / MC_i \quad i = \text{FRMP, NP, FSP.}$$

Basic economics indicates if $AC_i > MC_i$, or $MBE_i > 1$, there are increasing returns to scale; if $AC_i = MC_i$ ($MBE_i = 1$), there are constant returns to scale; and if $AC_i < MC_i$ ($MBE_i < 1$), there are decreasing returns to scale.

Because the error terms across equations are likely to be correlated, we use the non-linear seemingly unrelated regression method (Wooldridge, 2002) for model estimation using the statistical software STATA Version 11. Because many of the explanatory variables are categorical percentages, a reference set of categories must be chosen to avoid perfect collinearity. The reference case here is the percentage of participants younger than age 20 years, with income less than 50% of the poverty level, education up to the high school level, residing in a central city with a population over 50,000, participated in a group lesson, and instructed by a volunteer. Gender is excluded from the estimation model as $\approx 90\%$ of the sample is female, so the variation is quite limited. Finally, because of repeated observations by states/territories, we use a cluster-robust covariance matrix because there is likely correlation in the errors over time for the same state/territory (Wooldridge, 2002).

Results

The parameter estimates and their corresponding p values are given in Table 3 and are only briefly discussed before turning to the returns to scale estimates, because they are the central focus of the analysis. Although we recognize the choice of a significance level for inference is somewhat arbitrary, for ease in communication, we use a 0.10 significance level cutoff for deciding significance but report p values such that readers may choose a different cutoff. Following the result presentation is a discussion of the results.

Budget

Consistent with our hypothesis, the parameter estimate for the budget is positive and statistically significant at conventional levels in the FRMP and the NP index equations (FRMP:

Table 3. Nonlinear Seemingly Unrelated Regression Parameter Estimates

Variable	System Equation Results					
	FRMP		NP		FSP	
	Parameter	<i>p</i> value	Parameter	<i>p</i> value	Parameter	<i>p</i> value
Log of budget	0.146	0.01**	0.116	0.04**	0.001	0.98
Number of participants	0.015	0.35	0.012	0.42	0.003	0.84
Number of youth participants	0.000	0.99	-0.005	0.47	0.003	0.68
Income 100% of poverty level	-0.011	0.17	-0.008	0.27	-0.001	0.94
Income 150% of poverty level	-0.034	0.16	-0.030	0.15	-0.027	0.12
Income 150+% of poverty level	0.039	0.11	0.058	0.01**	0.026	0.09*
Income NA	-0.002	0.50	0.000	0.87	0.001	0.82
White	-0.003	0.52	0.000	0.95	0.003	0.39
Black	-0.005	0.28	-0.002	0.58	0.003	0.41
Hispanic	0.004	0.25	0.003	0.45	0.010	0.02**
Age 20–29	-0.005	0.46	0.000	0.98	0.002	0.71
Age 30–39	0.012	0.29	0.015	0.20	0.027	0.01**
Age 40–49	-0.008	0.67	0.000	0.99	-0.009	0.54
Age 50–59	-0.109	0.25	-0.031	0.42	-0.071	0.20
Age 60+	0.111	0.19	0.025	0.32	0.030	0.40
Age NA	-0.023	0.07*	-0.021	0.04**	-0.017	0.08*
Some college	-0.005	0.80	0.002	0.90	-0.014	0.22
College grad and up	-0.014	0.71	-0.015	0.67	-0.001	0.97
Education NA	-0.003	0.48	0.000	0.89	-0.005	0.07*
Rural farm	0.003	0.84	0.011	0.52	-0.017	0.16
Rural nonfarm	0.012	0.00**	0.010	0.00**	0.010	0.00**
Small town	0.006	0.22	0.006	0.19	0.004	0.29
Suburb of cities	-0.005	0.41	-0.015	0.08*	-0.002	0.74
Individual lesson	0.006	0.12	0.005	0.18	0.000	0.90
Both group and individual lesson	0.007	0.45	-0.005	0.47	0.004	0.64
Other lesson	-0.010	0.34	-0.003	0.82	-0.014	0.06*
Professionals	0.033	0.15	0.035	0.09*	0.024	0.08*
Paraprofessionals	0.006	0.25	0.002	0.74	0.002	0.46

** Significant at 5%; * significant at 10%.

Goodness-of-fit statistics is given by the correlation coefficient between the predicted values and the actual values: FRMP (0.46), NP (0.48), and FSP (0.56).

FRMP, Food Resource Management Practices; NP, Nutrition Practices; FSP, Food Safety Practices.

0.15, *p* value = 0.01 and NP: 0.12, *p* value = 0.04). The parameter estimate for the budget is not significantly different from zero in the FSP index equation (0.00, *p* value = 0.98).

Other Covariates

Very few of the participant characteristics are significant. We focus on those that are significant in more than one equation. As the percentage of participants with incomes over the 150% of the poverty line increases, the NP index (0.06, *p* value = 0.01) and FSP index (0.03,

p value = 0.09) both increase. Although still positive, this effect is not statistically significant in the FRMP index (0.04, *p* value = 0.11). When the percentage of participants who did not report their age increases relative to the percentage of younger participants (i.e., those of age younger than 20 years), all indices decrease (FRMP: -0.02, *p* value = 0.07; NP: -0.02, *p* value = 0.04; FSP: -0.02, *p* value = 0.08). The other control variable that is statistically significant in more than one equation is the place of residence variable (rural nonfarm). Relative to the base place of residence

(central cities), rural nonfarm scored significantly higher on all three indices (FRMP: 0.01, p value = < 0.001; NP: 0.01, p value = 0.001; FSP: 0.01, p value = < 0.001).

With respect to the program characteristics, there is no apparent significant difference in the index scores associated with the lesson types with the exception of the “other” lesson type in the FSP index equation. As the percentage of participants that did not participate in a group and/or individual lesson increases, the FSP index score decreases (-0.01 , p value = 0.06). The variable percentage of professional instructors is significant in NP and FSP index equations. An increase in the percentage of professional instructors increases both the NP index (0.03, p value = 0.09) and the FSP index (0.02, p value = 0.08).

Returns to Scale

As indicated, the marginal budget effect is nonlinear and varies by evaluation point. The marginal budget effect for each index is evaluated at the average value of all covariates at the 1) national level (all states/territories), at the average value for all covariates 2) for the corresponding regional² level, and at the average value of all covariates 3) for the corresponding state/territory level. The variance and associated p values for these marginal budget effects are based on the delta method and account for clustering (Wooldridge, 2002). The relevant hypotheses tested are for decreasing, constant, and increasing returns to scale, which amounts to testing if the marginal budget effects are less than one, one, and greater than one, respectively, and so these are the results discussed. To avoid redundancy, we only discuss the returns to scale for the FRMP and NP indices because the marginal budget effect for the FSP index is not significantly different from zero, regardless of level of aggregation (national, regional, state), implying decreasing returns to scale at all levels. Although the

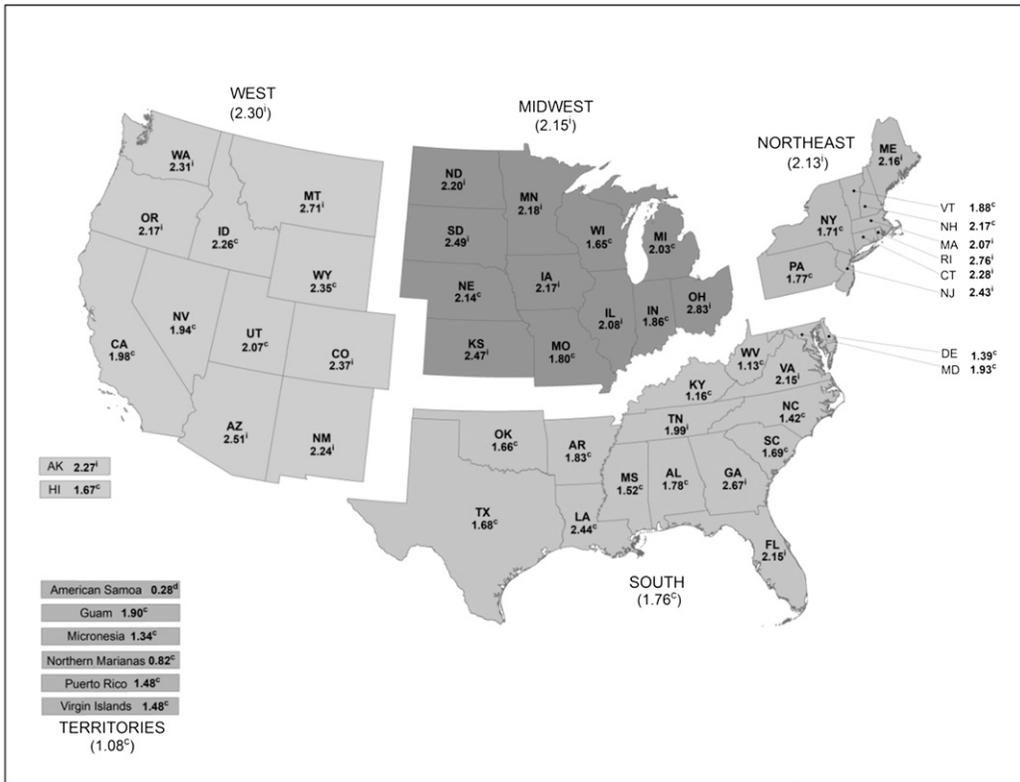
emphasis in this section is on statistical significance, the final section incorporates a broader analytical perspective by also discussing the economic significance of the results (McCloskey, 1985).

At the national level, the marginal budget effect for the FRMP index is 2.01 and for the NP index is 1.20. Although the FRMP marginal budget effect is 2.01, the returns to scales test indicates it is not statistically different from one at the 0.10 significance level, although it is greater than one (increasing returns to scale) at the 0.11 level (p value = 0.11). Similarly, the marginal budget effect of 1.20 for NP is also not statistically different from one, indicating constant returns to scale.

There is however a great deal of variability by region and state/territory in the marginal budget effects and hence returns to scale. Figures 1 and 2 provide the marginal budget effects evaluated at the regional level and at the state/territory level for the FRMP index and the NP index, respectively.

For the FRMP index (Figure 1), the marginal budget effect is greater than 1.00 for all regions. The highest returns to scale are in the West ($MBE_{FRMP} = 2.30$) followed by the Midwest ($MBE_{FRMP} = 2.15$), the Northeast ($MBE_{FRMP} = 2.13$), the South ($MBE_{FRMP} = 1.76$), and the Territories ($MBE_{FRMP} = 1.08$). The returns to scale tests indicate that three of the five regions: the West (p value = 0.08); the Midwest (p value = 0.09); and the Northeast (p value = 0.09) have statistically increasing returns to scale, and the two regions South and Territories have statistically constant returns to scale, yet even within each region, there is some variability across states. With exception of the Territories, each region has some states that have statistically significant increasing returns to scale in the FRMP index. Seven of 13 states in the West had statistically increasing returns to scale: Washington (2.31), Montana (2.71), Arizona (2.51), Colorado (2.37), New Mexico (2.24), Alaska (2.27), and Oregon (2.17). Seven of the 12 states in the Midwest had statistically increasing returns to scale: South Dakota (2.49), Kansas (2.47), Iowa (2.17), Ohio (2.83), Minnesota (2.18), North Dakota (2.20), and Illinois (2.08). Five of the nine states in

²We divided the United States into four regions based on the regional divisions used by the US Census Bureau. For completeness of the analysis, all US territories were categorized into a fifth region.



Note: Statistical test indicating *d* = decreasing returns to scale; *c* = constant returns to scale; *i* = increasing returns to scale at 0.10 significance level.

Figure 1. Returns to Scale on Food Resource Management Practices by State/Territory

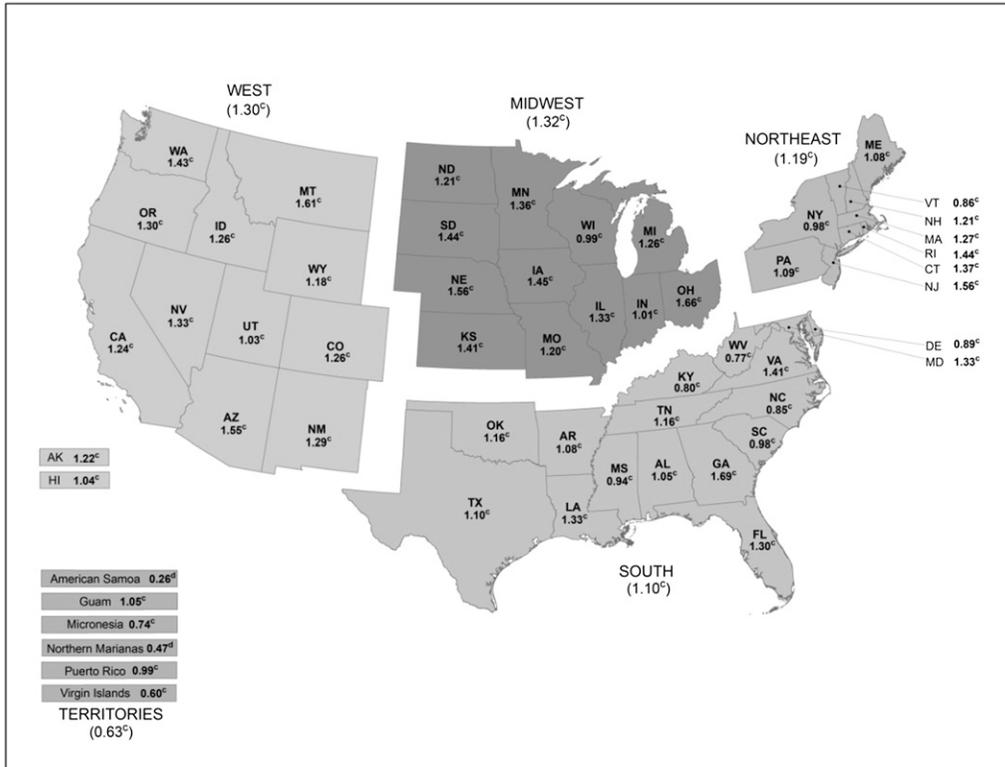
the Northeast had statistically significant increasing returns to scale: Massachusetts (2.07), Rhode Island (2.76), Connecticut (2.28), Maine (2.16), and New Jersey (2.43). Only four of the 16 states in the South had statistically increasing returns to scale: Georgia (2.67), Tennessee (1.99), Virginia (2.15), and Florida (2.15). All other states have statistically constant returns to scale in the FRMP index.

For the NP index (Figure 2), the marginal budget effect is numerically greater than 1.00 for four of the five regions. The highest numerical returns to scale are in the Midwest ($MBE_{NP} = 1.32$) followed by the West ($MBE_{NP} = 1.30$), the Northeast ($MBE_{NP} = 1.19$), the South ($MBE_{NP} = 1.10$), and the Territories ($MBE_{NP} = 0.63$), but the statistical tests indicate all regions have statistically constant returns to scale in the NP index. In contrast to

the FRMP index, although there is some numerical variability across states within each region, all states have statistically constant returns to scale with the exception of American Samoa and Northern Marianas, who have decreasing returns to scale in the NP index.

Discussion and Implications

The participant characteristics are measured in terms of categorical percentages (percentages of the sample falling within a category) so the effects (parameter estimates) associated with these variables are effectively measuring differences as a result of shifting the distribution of individuals on the characteristic domain. The closest literature to our analysis is on child education attainment (e.g., Gunn, Klebanov, and Duncan, 1996; Rumberger and Palardy, 2005).



Note: Statistical test indicating *d* = decreasing returns to scale; *c* = constant returns to scale; *i* = increasing returns to scale at 0.10 significance level.

Figure 2. Returns to Scale on Nutrition Practices by State/Territory

In these studies, the main demographic variables are those of the parents so any comparisons with our results would exceed the bounds of valid comparison. Furthermore, in these child education attainment studies, monetary inputs are usually not included as an explanatory variable, so by standard omitted variable analysis, the effects of parent or student characteristics include an omitted variable bias, which may lead to their significance. We explored this possibility and estimated all models without the budget variable and found that the same variables were still insignificant. Alternatively, one could argue that the statistical insignificance of many of the explanatory variables in our model could be the result of the categorical nature of the variables. It is well established that the categorical variables can mask the underlying significant relationships if the variable is better measured by a continuous variables such as income (e.g., MacCallum et al., 2002).

Keeping in mind these measurement caveats, the results of this analysis suggest there are generally no major inherent participant characteristic biases affecting the effectiveness of the EFNEP. In addition, there does not appear to be any general major program characteristic that is more effective than any other characteristic. These findings are actually useful and positive because they suggest the program appears to be uniformly effective across these different personal and program characteristics as a whole, so efforts to change the personal and program characteristic profiles would be predicted not to have a significant effect on the effectiveness of the program. However, our results show that states with a higher proportion of participants from rural areas generally tend to perform better than the states with a higher proportion of participants from the urban areas. What this implies is that other differences across states and regions also should be investigated

to help explain differences in outcomes not explained by these factors such as geographic, cultural, and other economic-based differences.

For two of the three indices used by the USDA to measure the success of the adult EFNEP, the FRMP, and the NP, the money being spent has a positive and significant effect. Three of five regions show statistically significant increasing returns to scale in the FRMP. Furthermore, 23 states show statistically significant increasing returns to scale in the FRMP index: Alaska, Arizona, Colorado, Connecticut, Florida, Georgia, Illinois, Iowa, Kansas, Maine, Massachusetts, Minnesota, Montana, New Jersey, New Mexico, North Dakota, Ohio, Oregon, Rhode Island, South Dakota, Tennessee, Virginia, and Washington. For the NP, almost all states and regions show constant returns to scale.

McCloskey (1985) makes a compelling case that often too much emphasis is placed on statistical significance and not enough on economic significance. She gives a vivid example of testing purchasing power parity, which boils down to a test of $\beta = \text{one}$ in a model (pp. 201–202). Just because someone has a very large sample and rejects the estimate that $\beta = 0.999$ is equal to one, from an economic significance or common sense perspective, purchasing power parity would not be rejected. It should be clear that her general argument would apply to tests of returns to scale and hence are very relevant here as well. Although from a statistical perspective the majority of states have constant returns to scale in the FRMP and NP indices, in terms of economic significance, more latitude should be applied in interpreting the numbers. For the FRMP, there are 29 states with return to scale estimates over 2.00 and another 18 with return to scale estimates between 1.50 and 2.00. For the NP, there are 42 states with return to scale estimates over 1.00. All of this suggests that, given the stated objectives and outcome indices for the EFNEP used by the USDA, the program is very effective within a broader evaluation perspective.

How can these results be used to improve the EFNEP? Much of the issue of improving the EFNEP comes down to first having objective valid information on effectiveness at the

state level. Although a lot of information is exchanged between EFNEP administrators, until we actually know which states are doing better relative to others, it will remain unclear who has the more effective program and hence more effective information. The model and returns to scale estimates here provide this first required piece of information. Much of the usefulness of this research is then in simply identifying outcome areas and states that are doing well relative to others and all that may be required is to share information on different practices, which can easily be done at the National EFNEP Conference each year.

The results shed light on two main areas where the EFNEP administrators and stakeholders should look to make improvements. First, it is clear that the money being spent on food safety is not having much of an impact on the FSP index. Re-evaluating the general EFNEP approach to food safety practices and considering alternative approaches may actually improve the food safety outcomes associated with EFNEP. Second, even for the FRMP and NP, there is a great deal of variability even within regions on returns to scale. EFNEP administrators and stakeholders in a relatively low-performing state need only look to the higher performing states in their region for insights into how their program operations differ and for suggestions for ways of improving their program operations.

Regarding limitations, two seem especially noteworthy. First, behavioral checklists and categorical-based surveys such as that used by USDA are used mainly for their high response rate and practicality in implementation. They are easy targets for criticism because they must navigate the often unclear tradeoffs between validity, reliability, response burden, and practicality in implementation. The procedures implemented by the USDA in developing the behavior checklist are in line with the recommendations of the literature (e.g., Contento, Randell, and Basch, 2002; Kristal et al., 1990) and there is some evidence that behavior checklist questions can correlate reasonably well with behavioral and some biological changes (e.g., Murphy et al., 2001). However, when there is interest in possibly multiple domains or

constructs, the improvement in a particular domain could simply depend on the number or type of questions asked, as pointed out by a reviewer. For example, the 10-item behavioral checklist used by the USDA contains only two questions for the FSP compared with five questions for the NP index. This implies that even if questions are answered randomly, the probability of improving on at least one question, as considered by the USDA, will be higher the more questions that are asked on a given domain. This is certainly an area where more research is needed on the quality of the USDA behavioral checklist and related indicators. Until this time, these existing national indicators are the obvious place to start in any type of national-level analysis and discussion of returns to scale and costs of the EFNEP.

Second, the stated objectives of the EFNEP focus on education and behavior and it is well known that improved knowledge and behavior may not translate into health benefits. A few small state-level EFNEP cost-benefit studies do indicate the health benefits, in terms of reduced health spending, exceed the costs (Dollahite, Kenkel, and Thompson, 2008; Joy, Pradhan, and Goldman, 2006; Rajgopal et al., 2002; Schuster et al., 2003). At the national level, this is a daunting question. However, regardless of the theory used to support the notion of changing health through an education program, a central component is to change behavior and this is one of the main direct goals stated by EFNEP. Consequently, the type of analysis conducted here seems the appropriate starting point for any returns to scale and cost analysis. As usual, these two points imply further research is needed, but, at this time, the best available information indicates the EFNEP is an effective program in achieving its stated goals.

[Received December 2011; Accepted December 2012.]

References

- Anliker, J., W. Willis, and S. Montgomery. "The Development and Testing of the Behavior Checklist Questions for the EFNEP Evaluation/Reporting System." Internet site: www.csrees.usda.gov/nea/food/efnep/ers/documentation/checklist-development.pdf (Accessed August 12, 2011).
- Arnold, C.G., and J. Sobal. "Food Practices and Nutrition Knowledge after Graduation from the Expanded Food and Nutrition Education Program (EFNEP)." *The Journal of Nutrition* 32(2000):130-38.
- Contento, I.R., J.S. Randell, and C.E. Basch. "Review and Analysis of Evaluation Measures Used in Nutrition Education Intervention Research." *Journal of Nutrition Education and Behavior* 34(2002):2-25.
- Dickin, K.L., J.S. Dollahite, and J.P. Habicht. "Nutrition Behavior Change among EFNEP Participants Is Higher at Sites That Are Well Managed and Whose Front-Line Educators Value the Program." *The Journal of Nutrition* 135(2005):2199-205.
- Dollahite, J.S., D. Kenkel, and C.S. Thompson. "An Economic Evaluation of the Expanded Food and Nutrition Education Program." *Journal of Nutrition Education and Behavior* 40(2008):134-43.
- Dollahite, J.S., and M. Scott-Pierce. "Outcomes of Individuals vs. Group Instruction in EFNEP." *Journal of Extension* 41(2003). Internet site: www.joe.org/joe/2003april/a4.php (Accessed August 11, 2011).
- General Accounting Office (GAO). *Nutrition Education: USDA Provides Services through Multiple Programs, but Stronger Linkages among Efforts are Needed*. Washington, DC: United States General Accounting Office, 2004.
- Gunn, J.B., P.K. Klebanov, and G.J. Duncan. "Ethnic Differences in Children's Intelligence Test Scores: Role of Economic Deprivation, Home Environment, and Maternal Characteristics." *Child Development* 67(1996):396-408.
- Joy, A.B., V. Pradhan, and G. Goldman. "Cost-Benefit Analysis Conducted for Nutrition Education in California." *California Agriculture* 60(2006):185-91.
- Kmenta, J. *Elements of Econometrics*. 2nd ed. Ann Arbor, MI: University of Michigan, 1997.
- Kristal, A.R., B.F. Abrams, M.D. Thornquist, L. Disogra, R.T. Croyle, A.L. Shattuck, and H.J. Henry. "Development and Validation of a Food Use Checklist for Evaluation of Community Nutrition Interventions." *American Journal of Public Health* 80(1990):1318-22.
- MacCallum, C.R., S. Zhang, K.J. Preacher, and D.D. Rucker. "On the Practice of Dichotomization of Quantitative Variables." *Psychological Methods* 7(2002):19-40.

- Marmot, M., and M. Wadsworth, eds. *Social Determinants of Health*. Oxford, UK: Oxford University Press, 1999.
- McCloskey, D.N. "The Loss Function Has Been Misaid: The Rhetoric of Significance Tests." *The American Economic Review* 75(1985): 201–205.
- Murphy, A.P., L.L. Kaiser, M.S. Townsend, and L.H. Allen. "Evaluation of Validity of Items for a Food Behavior Checklist." *Journal of the American Dietetic Association* 101(2001): 751–61.
- Pollard, C., M. Miller, R. Woodman, R. Meng, and C. Binns. "Changes in Knowledge, Beliefs, and Behaviors Related to Fruit and Vegetable Consumption Among Western Australian Adults from 1995 to 2004." *American Journal of Public Health* 99(2009):355–61.
- Rajgopal, R., R.H. Cox, M. Lambur, and E.C. Lewis. "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 and Behavior* 34(2002): 26–37.
- Rumberger, R.W., and G.J. Palardy. "Test Scores, Dropout Rates, and Transfer Rates as Alternative Indicators of High School Performance." *American Educational Research Journal* 42(2005):3–42.
- Schuster, E., Z.L. Zimmerman, M. Engle, J. Smiley, E. Syversen, and J. Murray. "Investing in Oregon's Expanded Food and Nutrition Education Program (EFNEP): Documenting Costs and Benefits." *Journal of Nutrition Education and Behavior* 35(2003):200–206.
- Shephard, R.W. *Indirect Production Functions*. Meisenheim, Germany: Anton Hain Publisher, 1974.
- United States Department of Agriculture (USDA)—National Institute of Food and Agriculture (NIFA). *Expanded Food and Nutrition Education Program*. Internet site: www.nifa.usda.gov/nea/food/efnep/efnep.html (Accessed August 10, 2009a).
- . *NIFA Budget Information. Archived Extension Activities*. Internet site: www.csrees.usda.gov/about/offices/budget.html (Accessed August 11, 2009b).
- . Cooperative State Research, Education, and Extension Service. *Nutrition Education Evaluation and Reporting System*. Internet site: www.csrees.usda.gov/nea/food/efnep/neers5/neers5.html (Accessed August 11, 2009c).
- Wooldridge, J.M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press, 2002.