Different Parts of Supplemental Nutrition Assistance Program (SNAP) Leads to Modest Changes in Food Security of Low-income Households

Ronglu Jiang, Xuanlong Liang, Hanbing Lu, Yumeng Tian, Shanrong Zhou, Ruizhe Zhu

Part 1: Research Proposal

Executive Summary / Abstract

Authors (Names and Percentages): Ronglu Jiang 100%

The U.S. Department of Agriculture (USDA) introduced the Supplemental Nutrition Assistance Program (SNAP) which gives vulnerable households a monetary subsidy to improve food insecurity and reduce household food stress. The extension of SNAP, SNAP-Ed, provides low-cost, and easy-to-prepare recipes and nutrition education materials that teach people to plan, shop, cook, and save to improve their health. This research examined whether SNAP's food subsidy improves the food security status for program participants and whether SNAP-Ed leads to improvement of food security in low-income households. The objective of the research is to determine the impact of SNAP on low-income households and hence the need to involve eligible non-SNAP participants for comparison purposes.

In this study, participants are required to first take the SNAP participation surveys that ask: whether the participant or any member of the household **receives SNAP food subsidy** (without any healthy diet or education on food preparation and nutrition), **receive SNAP food subsidy and education on food preparation and nutrition** (without any healthy diet), and **receive SNAP food subsidy and a healthy diet** (without education on food preparation and nutrition) for six months consecutively in the recent past. Participants in the control group sample did not receive any SNAP benefits but lived below the 130% poverty level. The sample, in terms of race and age, was representative of the US population.

Participants who had been in the program over the prior 6-7 months and new-entrant households involved the following four cohorts: (1) A control group: non-SNAP participants who have experienced persistent poverty; **SNAP participants will be randomized to one of the other three conditions:** (2) Participants receive only food subsidies without any healthy diet or education on food preparation and nutrition; (3) Participants receive food subsidies and education on food preparation and nutrition, without any healthy diet; (4) Participants in the fourth condition will receive food subsidies and a healthy diet without education on food preparation and nutrition.

Participants completed a suggested survey module developed by the U.S. Department of Agriculture, which are measured as indicators of their food security status and well-being.

The research can help USDA to have a preliminary understanding of the effectiveness of SNAP and SNAP-Ed, and provides support for the general direction of which subprojects to invest or reduce in the future. Under increasing food prices, future decision makers of the SNAP program can find insights from the result of the research and develop more effective policies supporting low-income families to reach food security and food sufficiency after the COVID-19 pandemic with limited resources.

Statement of the Problem

Authors (Names and Percentages): Yumeng Tian 100%

Food security is an important safeguard for public health and social development. According to Coleman-Jensen (2012), 14.5% of American households do not have access to enough food to live a healthy life, which means that these households are food insecure. In addition, food insecurity is particularly acute among African American households, and the incidence of food insecurity is higher among households below the income poverty line and those with more females and children. Food insecurity is important because it can have a negative impact on the physical and mental health and future development of people, especially children and adolescents, such as undernutrition and increased risk of chronic disease. Holben (2010) identifies poverty, education and the inability to find work as factors contributing to food insecurity.

To improve food insecurity and reduce household food stress, the US government introduced the Supplemental Nutrition Assistance Program (SNAP) which gives vulnerable households a monetary subsidy, thereby improving the food security, health and well-being of the most vulnerable. A study in North Carolina suggests a possible positive correlation between food insecurity and obesity, and they found that households receiving less than $150.10 per capita in SNAP food assistance had a higher body mass index than those receiving more than $150.10. They explained that the cheapest foods tend to be low in nutrients and thus have a negative impact on the body. On the other hand, it is not enough to rely closely on monetary subsidies; diet education is also needed to help vulnerable groups develop healthy eating habits. This is why there is the SNAP-Ed project, which helps the families being supported to maximize the use of food dollars to buy healthy food, thus reducing the likelihood of food insecurity.

Therefore, based on the context of the SNAP project, we need to address whether and to what extent SNAP and the SNAP-Ed project have improved the food security status of participants and what other factors influence these relationships.

Research Questions and Hypotheses

Authors (Names and Percentages): Ronglu Jiang 100%

The study is designed to answer the following questions:

1. Does **SNAP's food subsidy** improve the food security status and well-being for program participants?

   Null Hypothesis: H0: $\mu T - \mu C \leq 0$. The SNAP food subsidy **does not improve** the food security status and well-being of program participants.

   Alternative Hypothesis: Ha: $\mu T - \mu C > 0$.The SNAP benefits can **improve** the food security status and well-being of program participants.

   Possible Recommendation: If the study result rejects the null hypothesis, the program can spend more budget on providing subsidy. If the research rejects the alternative hypothesis, future decision makers of SNAP can consider reducing the partial budget in giving subsidies to other more efficient supporting policies.

2. Do the **SNAP-Ed recipes** recommended diet **have a positive effect** on the food security status and well-being of program participants?

   Null Hypothesis: H0: $\mu T - \mu C \leq 0$.The SNAP-Ed recipes recommended diet **does not improve** the food security status and well-being of program participants.

Alternative Hypothesis: Ha: μT – μC > 0. The SNAP-Ed recipes recommended diet **improves** the food security status and well-being of program participants.

Possible Recommendation: If the study rejects the null hypothesis, SNAP can consider creating free online/offline cooking sessions to tutor the participants to cook with healthy recipes. If the results reject the alternative hypothesis, SNAP can take further investigation on this subproject and seek improvement to better help participants.

3. Do the **nutrition education materials** offered by SNAP-Ed **improve** participating households' food security and well-being?

   Null Hypothesis: H0: μT – μC ≤ 0. The nutrition education materials offered by SNAP-Ed **couldn't improve** participating households' food security and well-being.

   Alternative Hypothesis: Ha: μT – μC > 0. The nutrition education materials offered by SNAP-Ed **improve** participating households' food security and well-being.

   Possible Recommendation: If the research results reject the null hypothesis, SNAP can increase budget in providing free nutrition education, including advertising the materials to participants by email and increase the frequency of opening nutrition education. If the results reject the alternative, SNAP can improve the quality of these nutrition materials, and invite specialists to contribute to the materials' modification.

In order to compare which benefit improved food security most efficiently, the study separated the participants into four different groups, receiving varied benefits. In the study, the author hypothesized that SNAP can improve participating households' food security by offering food subsidy monthly, providing SNAP-Ed, and offering free training opportunities. Besides, the author predicted that monthly subsidies can affect food security, since it restricted the category of foods participants can buy using the SNAP benefits. Finally, the author hypothesized that education on food preparation and nutrition can help improve food security and increase household perceptions to buy staple food, which includes fruits or vegetables; meat, poultry, or fish; dairy products; and breads or cereals.

The study can provide insights on the efficiency of SNAP helping low-income households access enough food to live a healthy life. The results of the study are intended to identify and quantify SNAP's effectiveness, and offer relevant recommendations to SNAP policy management in different states to improve the households' food security more efficiently.

Importance of the Study

Authors (Names and Percentages): Ronglu Jiang 100%

Despite the fact that average incomes and employment in the U.S. have improved since 2020, according to a report from the U.S. Department of Agriculture (USDA), "some U.S. households continue to face difficulties obtaining adequate food, particularly in the face of increasing food prices." The research of examining whether different parts of the SNAP program, like the SNAP-Ed, leads to improvement of food security in low-income households helps USDA to have a preliminary understanding of the effectiveness of each subproject in SNAP, and provides support for the general direction of which subproject to invest or reduce in the future. Under increasing food prices, future decision makers of the SNAP program can find insights from the result of the research and develop more effective policies supporting low-income families to reach food security and food sufficiency after the COVID-19 pandemic with limited resources. More efficient policy can provide effective support to females, children, infants, and adolescents who are in extreme need of high quality and nutritious food, and can decrease likelihood of health and mental issues caused by food insecurity.

Reviewing the findings of the research will help USDA and other federal agencies related to national food security generate ideas on how to efficiently help low-income families reduce food poverty, increase food sufficiency, and improve nutrition health. Based on research evidence, adjustments made by policymakers, practitioners, and intermediaries in SNAP related policies can reduce national poverty and further benefit national security and stability. Research study in this article can play a vital social role, assisting the U.S. government and businesses to develop services, policies, and products that are responsive to better food sufficiency in the future.

Literature Review

Authors (Names and Percentages): Shanrong Zhou 100%

Food security is always a concerning topic in the US. According to the United States Department of Agriculture (USDA), the prevalence of food insecurity (including low and very low food security) increased to a peak in 2011 at 14.9%. Although the percentage of food insecurity has decreased since then, 10.2% of US households were food insecure in 2021. Furthermore, among the households with different characteristics, the households whose income is below the poverty line had the significantly highest food insecurity rate (32%) in 2021, which is more than the tribble of the average food insecurity rate (10.2%) over the US. According to data provided by US Poverty Statistics, the number of poor people in the United States in 2021 will reach 37.9 million, accounting for 11.6% of the total population of the United States. This shows that in 2021, about 12.128 million people in the United States had food security problems.

SNAP has been promoting food security and a healthy diet, providing nutritional support for low-income households, from both financial and educational perspectives. Divided into different demographic groups, scholars have validated the effectiveness of SNAP in improving food security for children and national households. Research shows that, on average, SNAP helps more than 41 million low-income Americans access a nutritionally balanced diet each month. Previous research also showed that the educational resources provided by SNAP-Ed have a significant effect in improving the participants' food resource management skills and food safety practices. However, whether the specific parts of SNAP-Ed would have different effects on food security has not been widely studied yet. For example, studies have shown that improving knowledge, skills and attitudes related to healthy eating and food preparation through relevant education is the key to improving the eating habits of low-income groups. We know that SNAP-Ed includes education on food preparation and nutrition. Then we can judge whether the education provided by the SNAP-Ed program is effective by studying the perceptions and behaviors of low-income people on food safety based on this kind of education in the program. This result will inform possible improvements to the SNAP program. Therefore, we are going to focus on these multiple parts of SNAP-Ed and their association with households' food security improvement.

Research Plan

Authors (Names and Percentages): Hanbing Lu (25%), Yumeng Tian (25%), Ruizhe Zhu (25%), Shanrong Zhou (25%)

Population of Interest

Authors (Names and Percentages): Ruizhe Zhu 100%

The population of interest is persons who are low-income earners or live below the federal poverty line threshold (130 percent). Access to food is a significant problem for the majority of people from low-income families. In order to ascertain whether the supplemental nutrition aid program helps families from low-income households to receive food benefits, instruction on nutrition and food preparation, and a balanced diet, it is important to study these groups. Additionally, the population's increase in food security is a pivotal component and anticipated outcome of the SNAP project.

The study populations included all households eligible for SNAP benefits based on the income and resources criterion. These include households whose income is at or below 130 percent of the poverty line before any program deductions are made (Anders & Rafkin, 2022). Gross monthly income and net income both comprise earned and unearned cash income in the form of unemployment insurance, child support, cash assistance, and Social Security. Second, a household whose income is at the poverty line or less after program deductions are made. Thirdly, a household whose assets are below specific limits, including those without a person who is 60 years of age or higher or who is disabled and those with such a 60 year old member and assets of $2,750 or less and $4,250 or less, respectively (Anders & Rafkin, 2022). Additionally, the study participants in the control group must have been person's eligible for SNAP benefits but who had not received the benefits.

Sample Selection

Authors (Names and Percentages): Ruizhe Zhu 100%

In this current study, the researchers will evaluate the sample's baseline characteristics. The race is defined using self-identification using US Census categories. Age, gender, the presence of children in the home (with children or no children), marital status (married or living with a partner and or not married or not living with a partner), disability (disabled or not disabled), self-reported employment status (full-time, part-time, or unemployed), educational level (high school or college graduate) via self-report, home ownership (rent, own, or live with family or friends), access to a vehicle (access or no access), and chronic and serious health conditions (a checklist consisting of obesity, high blood pressure, heart disease, kidney disease, diabetes or high blood sugar, high cholesterol, arthritis, and cancer) will be all assessed among participants based on 2021 income brackets.

We will collect data on SNAP participants in a survey. In the study survey, three questions will be asked to determine sample participants (the inclusion criteria) (1) "Did any member of your household receive Supplemental Nutrition Assistance Program (SNAP) food subsidy (without any healthy diet or education on food preparation and nutrition) for at least 6 months consecutively in the recent past?" (2) "Did you or any of your family members receive SNAP food subsidy and education on food preparation and nutrition (without any healthy diet) for six months consecutively in the recent past?" (3)"Did you or any of your family members receive SNAP food subsidy and a healthy diet (without any education about food preparation and nutrition) for six months consecutively in the recent past?" The control group sample is participants who did not receive any SNAP benefits but lived below the 130% poverty level. The sample, in terms of race and age, will be representative of the US population (5-year estimates) (Pollard & Jacobsen, 2020).

These inclusion criterias will help to reduce the study population groups by eliminating participants not fit for inclusion in each study cohort. Consequently, participants who had been in the program over the prior 6-7 months and new-entrant households thus involved four cohorts:

1. Participants in the first condition, a control group, are non-SNAP participants who have experienced persistent poverty. SNAP participants will be randomized to one of the other three conditions.

2. Participants in the second condition will receive only food subsidies and will not receive any education on food preparation and nutrition or healthy diet.

3. Participants in the third condition will receive food subsidies and education on food preparation and nutrition, but nothing on a healthy diet.

4. Participants in the fourth condition will receive food subsidies and a healthy diet, but without any education about food preparation and nutrition.

The first cohort that entails those eligible for SNAP but have not received SNAP benefits or participated in SNAP. Their selections eliminated those in the study population that have participated in SNAP before and those that are not eligible for the program benefits. The second group cohort eliminated sample participants eliminated by the group 3 and 4 selection criterion.

The selection is critically based on income. The objective of the research is to determine the impact of SNAP on low-income households and hence the need to involve eligible non-SNAP participants for comparison purposes

Sample Size

Authors (Names and Percentages): Ruizhe Zhu 100%

To calculate the appropriate sample size, H0: $\mu T - \mu C \leq 0$. The study hypothesized that SNAP benefits improve the food security status of program participants (null hypothesis).

The power of a statistical test is the likelihood of rejecting the null hypothesis when it is false. For statistical power, we set it to 90% in this study.

We will hold the assumption that an effect size in a power or sample size calculation is the lowest improvement that the SNAP program benefits have on its participant's food security. We assumed that effect size =2/10(which is a meaningful improvement of 2 points), sample standard deviation =10 and level of significance = 95%.

The two assumptions on statistics on the level of power, and the standard deviation we can calculate the sample size.

$Z = > \bar{X}-0\ \sigma/\sqrt{n}\ z1-\alpha$

Where;

$P(Z >) = P (> z1-\alpha)= P(X-0\ \sigma/\sqrt{n} >z1-\alpha)$

Given that the α =0.05 we will reject the null hypothesis id Z >Z(1-0.05) = 1.645

For a normal distribution with standard deviation=10

$1- \beta = 1 - P (W \leq -)\ z1-\alpha\ \mu A\ \sigma/\sqrt{n}$

$\beta = P (W \leq z1-\alpha-\mu A\ \sigma/\sqrt{n})$

The power calculation: Φ (1-0.9) = Φ (0.10)=−1.28 ;

$n = \{[z1-\alpha-\Phi\text{-}1\ (\beta)]/\ \mu A/\sigma\}\text{^}2$

Then n= (1.645-(-1.28)/(2/10))^2 = 855.5625

The smallest sample size is 856 which ensures a 90% statistical power on the perceived impact of SNAP benefits on food security among its participants from low households.

Operational Procedures

Authors (Names and Percentages): Hanbing Lu 100%

We will request the U.S. Department of Agriculture approval for this study and work with the SNAP program director to recruit subjects and to conduct the study. Recruited subjects will be randomly assigned to treatment groups and will receive different treatments for 6 consecutive months. Participants will receive a variety of rewards, including cash, food vouchers and lottery tickets, to motivate individuals to participate.

Given the sensitivity of the topic, the reluctance of many people to give a negative response in person, and the fact that they are more likely to tell the truth in writing than in conversation, the study will conduct an online survey to assess the food safety status and well-being of participants after the six months. Subjects who have not received SNAP benefits or participated in SNAP but lived below the 130% poverty level in the same past 6 consecutive months will be recruited via Survey Sampling International. Current SNAP participants from each household will take the survey via official SNAP email.

The email will invite individuals to take the survey and does not include other details, thereby minimizing response bias. Participants who successfully completed the 3 survey questions to determine their inclusion criteria will complete the rest of the survey after giving their informed permission.

Note, however, that participants in the control group who are most affected by food insecurity will be more likely to complete the survey. In the treatment group, participants with strongly positive or negative opinions about the program will be more likely to complete the survey. As a result, survey responses are likely to be quite polarized, with some relative moderation due to which program participants choose to complete the survey. Unwillingness to participate in the study will result in their exclusion from the study with no penalty. The expected average completion time is 15 minutes.

Prior to data collection, the data collection team will receive two consecutive days of training. The training includes an introduction and background to the study, protocols to ensure that the treatment is applied correctly to each group, how to monitor participant performance, how to deliver the survey, and group discussion of each survey question. To assure the quality and dependability of the data collection process, the data collection supervisor, who has prior expertise in survey data collection, will oversee and assess the work of the data collectors.

Brief Schedule

Authors (Names and Percentages): Hanbing Lu 100%

The estimated maximum duration of the project is approximately 359 days, and the estimated minimum duration of the project is approximately 302 days. The proposed schedule is as follows:

Stage	Activity	Estimated duration	Deliverable
Research planning and set-up	Identify research problem	1-3 days	Confirmed research problem
	Formulate hypothesis	1 day	Confirmed hypothesis
	Review published literature	3-5 days	Notes from the review of relevant literature
	Prepare research design	1-2 weeks	Draft research proposal
	Gather peer/reciprocal review	1-2 weeks	Confirmed research proposal
Data collection	Develop sampling plan	2-3 days	Sampling plan
	Develop data collection Instrumentation plan	1 weeks	Draft data collection instrumentation plan
	Conduct a pilot study on a small sample	1-2 weeks	Confirmed data collection instrumentation plan
	Provide training for data collection personnel	2 days	Instructions or training for those who are collecting information
	Conduct the experiment	6 months	Confirmed experiment
	Conduct data collection	1 month	Raw survey data
Data analysis	Prepare data for analysis	2-3 weeks	Data ready for analysis

	Analyze data	1-2 weeks	Draft outcomes/results from analysis
	Draw conclusions	2-3 days	Draft conclusions
Writing up	Final report	3-4 weeks	Final draft
	Review draft with supervisor	3-5 days	Feedback notes
	Final editing	1-2 weeks	Final report
	Final submission	1 day	Final submission of report

Data Collection

Authors (Names and Percentages): Hanbing Lu 100%

The independent variables intended to be examined in this study are SNAP program with food subsidies (treatment group 1), SNAP program with food subsidies and food preparation and nutrition education (treatment group 2), SNAP program with food subsidies and healthy eating (treatment group 3), and no SNAP program (control group). The independent variables will be measured by food security status and well-being indicators. Other variables to control for in the study may be other food-related benefits that participants receive from other programs. We want to make sure that they are only receiving food-related benefits from the SNAP program.

Household food security status within the past 12 months will be assessed using the six-item short form of the food security survey module developed by the U.S. Department of Agriculture. It is a condensed version of the 18-item food security survey module to reduce the response burden of households experiencing food insecurity. These questions are designed to determine whether a household has access to adequate food, the ability to purchase and prepare food, and the ability to consume a balanced diet**.** Responses of "often" or "sometimes", "almost every month" or "some months but not every month", or "yes" are coded as affirmative responses. The household's raw score is calculated as the total of affirmative responses. Households can be divided into three groups using the following cut-off points. 'Households with high or marginal food security' (0-1 affirmative responses); 'households with low food security' (2-4 affirmative responses); and 'households with very low food security' (5-6 affirmative responses).

An additional 3 questions on a scale of 0 to 6 to measure well-being indicators including healthy life, diabetes and hypertension (food insecure people have a 2.4 times greater risk of diabetes and hypertension), and weight loss, will also be included in the survey.

Meanwhile, data on participants' sociodemographic characteristics, including gender, age, height, weight, ethnicity, race, education level, occupation, history of illness in the household, and the state or region in which they reside in the United States will be collected at the end of the survey because we do not want to start the survey with more sensitive questions.

Data Security

Authors (Names and Percentages): Yumeng Tian 100%

Since this is an experiment-based study and the experimental data are obtained using surveys, there are some data security concerns about subject safety, researcher training, security aspects of the partnering organization, and data storage.

The first is the participants' confidentiality and health. This study examined the difference in food safety and health between the treatment group and the control group after the treatment group received food subsidies and nutrition instruction. Although there are no physical experiments performed on the participants, it is crucial to recruit healthy individuals to guarantee that their health does not deteriorate during the trial. Second, we must collect basic information about the respondents, such as their gender, age, and ethnicity, for the study, which may include information such as their income and whether they suffer from various chronic diseases, in order to determine whether they are members of socially disadvantaged groups and their physical health status. This information may be utilized in the study and shared with other researchers with the consent of the participants. The names and addresses of respondents will not be collected or published in order to prevent their identification.

The second factor is researcher training and the security of the firms that collaborate. The goal of the training is to assure the quality and dependability of data collecting. After training, they should also be granted a certificate to establish their identification when collecting data from participants, hence lowering the risk that respondents may decline to participate in the survey. Moreover, in order to identify suitable participants, this study involved two partner organizations, the USDA and the SNAP project, from which we may receive permission to obtain personal information on a large number of people, particularly the poor; therefore, we must ensure that these data will not be leaked and avoid affecting the interests of other organizations.

The final step is the storage of survey data. The acquired data will be sorted into a useful format, and the raw and processed data will be stored in separate files that can be shared via encrypted links, used only for debate among group members, or disseminated with the consent of all group members. Also, if anyone uses our data in their research, they must attribute the source of the data.

Outcomes (Dependent Variables)

Authors (Names and Percentages): Shanrong Zhou 100%

The food security status and welling-being indicators of each survey's participants are the dependent variables of this research.

To be specific, this study takes the suggested survey module developed by the U.S. Department of Agriculture as the measurement of participants' food security. There are six questions in the survey: in the last 12 months, 1) how often they could buy enough food and last, 2) how often they could afford balanced meals, 3) whether they cut the size of or skipped meals because of not having enough money, 4) if yes for question 3, how often it happened, 5) whether they ate less because of not having enough money; 6) whether they felt hungry because of not having enough money. For questions 1 and 2, the options of the answer are "often", "sometimes", "never", and "don't know or refused". For questions 3, 5, and 6, the options of the answer are "yes", "no" and "don't know". For question 4, the options of the answer are "almost every month", "some months but not every month", "only 1 or 2 months", and "don't know". The sum of affirmative responses, as explained in Data Collection, will be the raw scores (ST) of each participant to measure the food security status. The lower the raw score, the higher the food security status. The module suggests dividing participant's food security status into several categories. However, compared to the categorical variables, the numeric variables can reflect more detailed changes in the measurement of food security status. Therefore, we directly take the numeric raw score (ST) to be a dependent variable of this research.

Additionally, the survey includes another 3 questions to measure the well-being status of participants. It asks the participants to scale from 0-6 to measure whether they can maintain a healthy life, diabetes and hypertension, and weight loss. The average score of 3 questions will be the raw scores (S2) of each participant to measure the well-being status. This numeric raw score (S2) will be another dependent variable of this research.

To align the raw scores for food security measurement and well-being indicator, SF(Low (6) to High(0)), the raw scores for food security is transformed to S1(Low(0) to High(6)). Thus, the final dependent variable will be: ST = S1 + S2 (S1: Transformed Food Security Score; S2: Well-being Score) for each participant.

Treatments (Independent Variables)

Authors (Names and Percentages): Shanrong Zhou 100%

We will randomly assign the survey's participants into four groups based on their SNAP participation status: 1) treatment group 1: SNAP program participants who only received food subsidy, 2) treatment group 2: SNAP participants who received food subsidy and education on food preparation and nutrition (without any healthy diet), 3) treatment group 3: SNAP participants who received food subsidy and a healthy diet (without education on food preparation and nutrition), and 4) control group: people who did not participate in the SNAP program. Each category of the survey participants' SNAP participation status will be recorded as the independent variable of this study. We will also record other features of the survey's participants (eg. age, gender) to explore the potential effects on food security for future study. However, this research focuses on the effectiveness of different parts of the SNAP program for improving people's food security and well-being. Whether participating in different parts of the SNAP program can lead to changes in survey participants' food security is the main research question to address. Therefore, SNAP participation status categories would be the only independent variables in this research. We will further discuss the other variables later.

Other Variables

Authors (Names and Percentages): Shanrong Zhou 100%

There are some demographic factors to be collected about the survey's participants, including gender, age, height, weight, ethnicity, race, education level, occupation, history of illness in the household, and the state or region in which they reside in the United States. These are the factors that also potentially influence people's food security status. We want to make sure that these other variables do not affect the observed relationship between the independent variable and the dependent variable. Thus, we need to balance the control and treatment groups to ensure that neither group has an unfair advantage or disadvantages related to the other variables. We will randomly select survey's participants for control and treatment groups to ensure that these groups are reasonably balanced, ruling out the possibility of other effects. With these demographic data collected, we can conduct ANOVA tests to compare the mean of control and treatment groups, in terms of these other variables. With the significance indicated by ANOVA's results, we can check whether these groups are reasonably balanced in terms of the other variables.

Statistical Analysis Plan

Authors (Names and Percentages): Xuanlong Liang 100%

To do the Statistical Analysis on our collected data, we are going to conduct multiple two sample t-tests and run two linear regressions each having different dependent variables and the same independent variables.

First, since there are two dependent variables that we can get from the result of each valid survey - score of food security status (SF) and score of welling-being status (S2), we need to transform and combine the values of the dependent variables in order to combine the information provided by the two dependent variables and form a new dependent variable which can represent both the food security status and the welling-being status while keeping the same measuring scale. We do this by transforming the score of food security. Originally, the higher the score of food security means the lower food security level. To reverse it, we change 0 to 6, 1 to 5, 2 to 4, 3 to 3, 4 to 2, 5 to 1, and 6 to 0. After doing this, the score from high food security level to low food security level becomes 6 to 0 instead of 0 to 6. We call this new score of food security states S1. Then we add this new score of food security states to the score of well-being status and get a new dependent variable (ST = S1 + S2). We will use this new dependent variable (ST) as our dependent variable in the following analysis with the higher this new dependent variable (ST) is the higher the overall food security status (food security status+well-being status).

In order to solve the first Research question - Does the SNAP\'s food subsidy improve the food security status for program participants - we conduct a one-sided two sample t-test on the mean of ST of the control group and treatment group 1 (receive only food subsidies and will not receive any education on food preparation and nutrition or healthy diet). If the mean of ST of the treatment group 1 is significantly higher than the control group, we have enough evidence to say that the SNAP\'s benefits improve the food security status for program participants.

In order to solve the second Research question - Do the SNAP-Ed recipes recommended dietimprove the food security status for program participants - we conduct a one-sided two sample t-test on the mean of ST of the treatment group 1 and treatment group 3 (receive food subsidies and a healthy diet, but without any education about food preparation and nutrition). If the mean of ST of the treatment group 3 is significantly higher than the treatment group 1, we have enough evidence to say that the SNAP-Ed recipes recommended dietimprove the food security status for program participants.

In order to solve the third Research question - Do the nutrition education materials offered by SNAP-Ed improve the food security status for program participants - we conduct a one-sided two sample t-test on the mean of ST of the treatment group 1 and treatment group 2 (receive food subsidies and education on food preparation and nutrition, but nothing on a healthy diet). If the mean of ST of the treatment group 2 is significantly higher than the treatment group 1, we have enough evidence to say that the nutrition education materials offered by SNAP-Ed improve the food security status for program participants.

Further, we can also compare the impact of the SNAP-Ed recipes recommended and that of the nutrition education materials offered by SNAP-Ed on the food security status for program participants by conducting a one-sided two sample t-test on the mean of ST of the treatment group 2 and treatment group 3. We can say that we have enough evidence that the action with significantly higher mean of ST has a larger positive effect on the food security status for program participants.

Although we get our sample in a quite random way, we can use an alternative approach that takes all the variables (including the potential confounding variables such as age, ethnicity, etc.) into account. This approach is to run a linear regression between ST and all the other variables. The coefficients can inform us the direction and magnitude of the impact each variable has on the food security status for program participants whenever they are considered as significant. The coefficients are considered as significant if the p-value of the coefficient is less than the significance level which is $\alpha = 0.05$ in our research. We can also do a ANOVA test to check whether the mean of ST of all the groups are equal to each other or not. If the p-value of the F Statistic is less than the significance level which is $\alpha = 0.05$ in our research, we can say we have enough evidence to reject the hypothesis that the mean of ST of all the groups are equal to each other. Then by comparing the significant coefficients, we can conclude the direction and magnitude of the impact each variable has on the food security status for program participants.

Limitations and Uncertainties

Authors (Names and Percentages): Yumeng Tian 100%

This study employed an experimental design in which diverse populations were assigned to control and treatment groups, followed by survey collection of feedback data. Given that our research question is whether this program affects the food security of low-income groups, an experiment is indeed the ideal technique to conduct a study that can produce estimates that are close to the genuine causal effect and, thus, draw meaningful conclusions. However, it is important to note that experiments have limitations and uncertainties that must be considered.

Initially, the expense of experimentation can be high. Even if the USDA and SNAP are willing to offer information about participants and are willing to assist us in funding the treatment group, we must endeavor to secure approval and financing from them. As indicated previously, we must also provide incentives for participants to complete post-experimental surveys. In contrast, the purpose of this study was to examine the impact of SNAP on food security in diverse groups and locations, necessitating the greatest sample size available; thus, the cost of this experiment was a significant limitation.

The second risk is sampling and measurement bias, which may be inescapable in this study since, despite the random assignment of participants to the control and treatment groups, the post-experimental survey is unavoidable. Low-income communities and those most affected by food safety issues may be more receptive to being questioned because they are interested in the widespread implementation of SNAP and are, therefore, more likely to contribute more information to aid in this component of the research. This has the potential to exaggerate the impact of SNAP because this study concentrated on the food safety impacts of SNAP funding on vulnerable populations, as opposed to all populations. Some respondents may not be able to provide precise information about themselves, which may affect their food security scores (the dependent variable) as well as other variables, and they may also attempt to obtain more government funding by answering, thereby introducing uncertainty into the study's findings.

Lastly, there are the variables and analysis restrictions. It is difficult to exclude the influence of other welfare programs or unofficial benefits on participants' food security, despite the fact that this study will collect as much information as feasible from participants. In addition, the basic statistical analysis methods do not always lead to the most accurate conclusions, necessitating controls for factors in the analysis methods, such as multiple regression to control for factors other than the core independent variables and individual effects models to exclude individual effects from confounding the results. Utilizing more complex analytical methods is challenging, but it is unnecessary for the current investigation.

Part 2: Simulated Studies

Authors (Names and Percentages): Xuanlong Liang 100%

Research Question 1:

Scenario 1: No Effect

Authors (Names and Percentages): Xuanlong Liang 100%

Simulation

```r
library(DT)
library(data.table)
library(truncnorm)
set.seed(4188)
analyze.experiment <- function(the.dat) {
  require(data.table)
  setDT(the.dat)

  the.test <- t.test(x = the.dat[Group == "Treatment1",
                                 S], y = the.dat[Group == "Control", S], alternative =
"greater")

  the.effect <- the.test$estimate[1] - the.test$estimate[2]
  upper.bound <- the.test$conf.int[2]
  p <- the.test$p.value

  result <- data.table(effect = the.effect, upper_ci = upper.bound,
                       p = p)

  return(result)
}

n <- 1800
simu1.dat <- data.table(Group = c(rep.int(x = "Treatment1", times = n/2), rep.int(x = "C
ontrol", times = n/2)))
simu1.dat[Group == "Control", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 6.5, s
d = 10), digits = 1)]
simu1.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 6.
5, sd = 10), digits = 1)]

analyze.experiment(the.dat = simu1.dat)
```

```
       effect upper_ci         p
1: 0.02733333      Inf 0.4318265
```

```r
#For simulations repeated 1000 times
B <- 1000
n <- 1800

set.seed(seed = 4188)
Experiment <- 1:B
Group <- c(rep.int(x = "Treatment1", times = n/2), rep.int(x = "Control", times = n/2))

sim.dat <- as.data.table(expand.grid(Experiment = Experiment, Group = Group))
setorderv(x = sim.dat, cols = c("Experiment", "Group"), order = c(1,1))
sim.dat[Group == "Control", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 6.5, sd
= 10), digits = 1)]
sim.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 6.5,
sd = 10), digits = 1)]
dim(sim.dat)
```

```
[1] 1800000       3
```

```
exp.results <- sim.dat[, analyze.experiment(the.dat = .SD),
                       keyby = "Experiment"]

DT::datatable(data = round(x = exp.results[1:100, ], digits = 3),
              rownames = F)
```

Show 10 entries Search:

Experiment	effect	upper_ci	p
1	-0.068		0.665
2	-0.215		0.908
3	0.01		0.474
4	0.003		0.494
5	0.041		0.4
6	-0.028		0.57
7	-0.06		0.647
8	-0.003		0.508
9	0.032		0.422
10	-0.004		0.509

Showing 1 to 10 of 100 entries Previous 1 2 3 4 5 … 10 Next

```
exp.results[, summary(effect)]
```

```
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max. 
-0.466444 -0.109056 -0.002778  0.001108  0.110917  0.466556 
```

```
exp.results[, sd(effect)]
```

```
[1] 0.1603741
```

```
exp.results[, mean(p < 0.05)]
```

```
[1] 0.051
```

```
exp.results[, mean(p > 0.05)]
```

```
[1] 0.949
```

```
1.96*exp.results[, sd(effect)]+0.001108
```

```
[1] 0.3154413
```

```
-1.96*exp.results[, sd(effect)]+0.001108
```

```
[1] -0.3132253
```

Analysis

For scenario 1 (no Effect), we assume that the average score for the treatment group 1 equals that of the control group which is 6.5. A single simulation shows that the effect size is 0.0273 which is less than 0.2 (the meaningful threshold we set).

Scenario 2: An Expected Effect

Authors (Names and Percentages): Xuanlong Liang 100%

Simulation

```
library(DT)
library(data.table)
library(truncnorm)
set.seed(4188)
analyze.experiment <- function(the.dat) {
  require(data.table)
  setDT(the.dat)

  the.test <- t.test(x = the.dat[Group == "Treatment1",
                                 S], y = the.dat[Group == "Control", S], alternative =
"greater")

  the.effect <- the.test$estimate[1] - the.test$estimate[2]
  upper.bound <- the.test$conf.int[2]
  p <- the.test$p.value

  result <- data.table(effect = the.effect, upper_ci = upper.bound,
                       p = p)

  return(result)
}


simu1.dat <- data.table(Group = c(rep.int(x = "Treatment1", times = n/2), rep.int(x = "C
ontrol", times = n/2)))
simu1.dat[Group == "Control", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 6.5, s
d = 10), digits = 1)]
simu1.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.
5, sd = 10), digits = 1)]

analyze.experiment(the.dat = simu1.dat)
```

```
      effect upper_ci          p
1: 0.2138889      Inf 0.08877162
```

```
#For simulations repeated 1000 times
B <- 1000
n <- 1800

set.seed(seed = 4188)
Experiment <- 1:B
Group <- c(rep.int(x = "Treatment1", times = n/2), rep.int(x = "Control", times = n/2))

sim.dat <- as.data.table(expand.grid(Experiment = Experiment, Group = Group))
setorderv(x = sim.dat, cols = c("Experiment", "Group"), order = c(1,1))
sim.dat[Group == "Control", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 6.5, sd
= 10), digits = 1)]
sim.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.5,
sd = 10), digits = 1)]
dim(sim.dat)
```

```
[1] 1800000       3
```

```
exp.results <- sim.dat[, analyze.experiment(the.dat = .SD),
                       keyby = "Experiment"]

DT::datatable(data = round(x = exp.results[1:100, ], digits = 3),
              rownames = F)
```

Show 10 entries Search:

Experiment	effect	upper_ci	p
1	0.147		0.176
2	0.009		0.477
3	0.203		0.101
4	0.131		0.209
5	0.391		0.008
6	0.132		0.206
7	0.232		0.072
8	0.133		0.202
9	0.195		0.115
10	0.218		0.089

Showing 1 to 10 of 100 entries Previous 1 2 3 4 5 … 10 Next

```
exp.results[, summary(effect)]
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-0.2672  0.1156  0.2282  0.2282  0.3352  0.7996
```

```
exp.results[, sd(effect)]
```

```
[1] 0.1573903
```

```
exp.results[, mean(p < 0.05)]
```

```
[1] 0.432
```

```
exp.results[, mean(p > 0.05)]
```

```
[1] 0.568
```

```
1.96*exp.results[, sd(effect)]+0.2282
```

```
[1] 0.5366849
```

```
-1.96*exp.results[, sd(effect)]+0.2282
```

```
[1] -0.08028492
```

Analysis

First, we do the simulation for the one-sided two samples t-test of research question 1 which investigates whether the SNAP's food subsidy improves the food security status for program participants. From the calculation of the sample size we know that in order to have a power of 0.9, we need at least 856 participants in each group. Therefore, we are choosing a sample size of 900 for each group. Since the range of the score ST is 0~12, and the participants in the control group are people in poverty which might have severe food insecurity issues and unhealthy body status, we assume the average score for the control group is 6.5 and the standard deviation is 10. For the treatment group 1, since the participants received SNAP's food subsidy, we assume that the average score for the treatment group 1 is 2 points higher than the control group which is 8.5. We choose 2 points higher here because a 2 points higher leads to a higher level of food security status or well-being status. A single simulation shows that the effect size is 0.2139 which is higher than 0.2 (the meaningful threshold we set). The p-value is 0.0888 > 0.05, therefore we don't have enough evidence to reject the null hypothesis in favor of the alternative hypothesis based on this simulation. This is for scenario 2 (have effect).

Research Question 2:

Scenario 1: No Effect

Authors (Names and Percentages): Xuanlong Liang 100%

Simulation

```r
set.seed(4188)
analyze.experiment <- function(the.dat) {
  require(data.table)
  setDT(the.dat)

  the.test <- t.test(x = the.dat[Group == "Treatment2",
                                 S], y = the.dat[Group == "Treatment1", S], alternative
= "greater")

  the.effect <- the.test$estimate[1] - the.test$estimate[2]
  upper.bound <- the.test$conf.int[2]
  p <- the.test$p.value

  result <- data.table(effect = the.effect, upper_ci = upper.bound,
                       p = p)

  return(result)
}


simu2.dat <- data.table(Group = c(rep.int(x = "Treatment2", times = n/2), rep.int(x = "T
reatment1", times = n/2)))
simu2.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.
5, sd = 10), digits = 1)]
simu2.dat[Group == "Treatment2", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.
5, sd = 10), digits = 1)]

analyze.experiment(the.dat = simu2.dat)
```

```
       effect upper_ci         p
1: 0.02166667      Inf 0.4456868
```

```r
#For simulations repeated 1000 times
B <- 1000
n <- 1800

set.seed(seed = 4188)
Experiment <- 1:B
Group <- c(rep.int(x = "Treatment2", times = n/2), rep.int(x = "Treatment1", times = n/
2))

sim.dat <- as.data.table(expand.grid(Experiment = Experiment, Group = Group))
setorderv(x = sim.dat, cols = c("Experiment", "Group"), order = c(1,1))
sim.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.5,
sd = 10), digits = 1)]
sim.dat[Group == "Treatment2", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.5,
sd = 10), digits = 1)]
dim(sim.dat)
```

```
[1] 1800000       3
```

```
exp.results <- sim.dat[, analyze.experiment(the.dat = .SD),
                       keyby = "Experiment"]

DT::datatable(data = round(x = exp.results[1:100, ], digits = 3),
              rownames = F)
```

Show 10 entries   Search:

Experiment	effect	upper_ci	p
1	0.047		0.384
2	-0.053		0.632
3	0.192		0.115
4	-0.055		0.637
5	-0.088		0.706
6	-0.307		0.973
7	0.239		0.067
8	-0.175		0.863
9	0.172		0.144
10	-0.172		0.859

Showing 1 to 10 of 100 entries   Previous 1 2 3 4 5 … 10 Next

```
exp.results[, summary(effect)]
```

```
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max. 
-0.428000 -0.111000 -0.005944  0.002935  0.114639  0.534444 
```

```
exp.results[, sd(effect)]
```

```
[1] 0.1595711
```

```
exp.results[, mean(p < 0.05)]
```

```
[1] 0.054
```

```
exp.results[, mean(p > 0.05)]
```

```
[1] 0.946
```

```
1.96*exp.results[, sd(effect)]+0.002935
```

```
[1] 0.3156943
```

```
-1.96*exp.results[, sd(effect)]+0.002935
```

```
[1] -0.3098243
```

Analysis

For scenario 1 (no Effect), we assume that the average score for the treatment group 3 equals that of the treatment group 1 which is 8.5. A single simulation shows that the effect size is 0.0217 which is less than 0.2 (the meaningful threshold we set).

Scenario 2: An Expected Effect

Authors (Names and Percentages): Xuanlong Liang 100%

Simulation

```r
set.seed(4188)
analyze.experiment <- function(the.dat) {
  require(data.table)
  setDT(the.dat)

  the.test <- t.test(x = the.dat[Group == "Treatment2",
                                 S], y = the.dat[Group == "Treatment1", S], alternative
= "greater")

  the.effect <- the.test$estimate[1] - the.test$estimate[2]
  upper.bound <- the.test$conf.int[2]
  p <- the.test$p.value

  result <- data.table(effect = the.effect, upper_ci = upper.bound,
                       p = p)

  return(result)
}


simu2.dat <- data.table(Group = c(rep.int(x = "Treatment2", times = n/2), rep.int(x = "T
reatment1", times = n/2)))
simu2.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.
5, sd = 10), digits = 1)]
simu2.dat[Group == "Treatment2", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 9.
5, sd = 10), digits = 1)]

analyze.experiment(the.dat = simu2.dat)
```

```
   effect upper_ci         p
1:  0.117      Inf 0.2297679
```

```r
#For simulations repeated 1000 times
B <- 1000
n <- 1800

set.seed(seed = 4188)
Experiment <- 1:B
Group <- c(rep.int(x = "Treatment2", times = n/2), rep.int(x = "Treatment1", times = n/
2))

sim.dat <- as.data.table(expand.grid(Experiment = Experiment, Group = Group))
setorderv(x = sim.dat, cols = c("Experiment", "Group"), order = c(1,1))
sim.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.5,
sd = 10), digits = 1)]
sim.dat[Group == "Treatment2", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 9.5,
sd = 10), digits = 1)]
dim(sim.dat)
```

```
[1] 1800000       3
```

```
exp.results <- sim.dat[, analyze.experiment(the.dat = .SD),
                       keyby = "Experiment"]

DT::datatable(data = round(x = exp.results[1:100, ], digits = 3),
              rownames = F)
```

Show 10 entries Search:

Experiment	effect	upper_ci	p
1	0.157		0.161
2	0.092		0.28
3	0.263		0.05
4	0.023		0.442
5	-0.03		0.574
6	-0.038		0.596
7	0.284		0.038
8	-0.055		0.636
9	0.206		0.102
10	-0.007		0.518

Showing 1 to 10 of 100 entries Previous 1 2 3 4 5 … 10 Next

```
exp.results[, summary(effect)]
```

```
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.350444  0.005472  0.121722  0.116618  0.227917  0.639222
```

```
exp.results[, sd(effect)]
```

```
[1] 0.159128
```

```
exp.results[, mean(p < 0.05)]
```

```
[1] 0.186
```

```
exp.results[, mean(p > 0.05)]
```

```
[1] 0.814
```

```
1.96*exp.results[, sd(effect)]+0.116618
```

```
[1] 0.4285089
```

```
-1.96*exp.results[, sd(effect)]+0.116618
```

```
[1] -0.1952729
```

Analysis

To do the simulation for the one-sided two samples t-test of research question 2, we set the sample size of treatment group 1, 2, and 3 to be 900. The assumptions for the treatment groups (except average score) are the same as above. The assumptions of the average score for the treatment group 2, and 3 are 9.5, and standard deviation to be 10. The reason for a 1 point average score higher is we are assuming the education and recipe will improve the score but with a smaller number compared to the improvement caused by the subsidy (about half). A single simulation for research question 2 shows that the effect size is 0.117 which is lower than 0.2 (the meaningful threshold we set).

Research Question 3:

Scenario 1: No Effect

Authors (Names and Percentages): Xuanlong Liang 100%

Simulation

```
set.seed(5780)
analyze.experiment <- function(the.dat) {
  require(data.table)
  setDT(the.dat)

  the.test <- t.test(x = the.dat[Group == "Treatment3",
                                 S], y = the.dat[Group == "Treatment1", S], alternative
= "greater")

  the.effect <- the.test$estimate[1] - the.test$estimate[2]
  upper.bound <- the.test$conf.int[2]
  p <- the.test$p.value

  result <- data.table(effect = the.effect, upper_ci = upper.bound,
                       p = p)

  return(result)
}
simu3.dat <- data.table(Group = c(rep.int(x = "Treatment3", times = n/2), rep.int(x = "T
reatment1", times = n/2)))
simu3.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.
5, sd = 10), digits = 1)]
simu3.dat[Group == "Treatment3", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.
5, sd = 10), digits = 1)]

analyze.experiment(the.dat = simu3.dat)
```

```
       effect upper_ci         p
1: -0.1251111      Inf 0.7887632
```

```
#For simulations repeated 1000 times
B <- 1000
n <- 1800

set.seed(seed = 5780)
Experiment <- 1:B
Group <- c(rep.int(x = "Treatment3", times = n/2), rep.int(x = "Treatment1", times = n/
2))

sim.dat <- as.data.table(expand.grid(Experiment = Experiment, Group = Group))
setorderv(x = sim.dat, cols = c("Experiment", "Group"), order = c(1,1))
sim.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.5,
sd = 10), digits = 1)]
sim.dat[Group == "Treatment3", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.5,
sd = 10), digits = 1)]
dim(sim.dat)
```

```
[1] 1800000       3
```

```
exp.results <- sim.dat[, analyze.experiment(the.dat = .SD),
                       keyby = "Experiment"]

DT::datatable(data = round(x = exp.results[1:100, ], digits = 3),
              rownames = F)
```

Show 10 entries Search:

Experiment	effect	upper_ci	p
1	-0.032		0.583
2	0.05		0.375
3	0.104		0.256
4	0.01		0.474
5	-0.153		0.833
6	-0.168		0.856
7	0.13		0.201
8	0.296		0.032
9	0.161		0.157
10	0.099		0.268

Showing 1 to 10 of 100 entries Previous 1 2 3 4 5 … 10 Next

```
exp.results[, summary(effect)]
```

```
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.470444 -0.095111  0.010278  0.008035  0.116528  0.485556
```

```
exp.results[, sd(effect)]
```

```
[1] 0.1580486
```

```
exp.results[, mean(p < 0.05)]
```

```
[1] 0.053
```

```
exp.results[, mean(p > 0.05)]
```

```
[1] 0.947
```

```
1.96*exp.results[, sd(effect)]+0.008035
```

```
[1] 0.3178103
```

```
-1.96*exp.results[, sd(effect)]+0.008035
```

```
[1] -0.3017403
```

Analysis

For scenario 1 (no Effect), we assume that the average score for the treatment group 3 equals that of the treatment group 1 which is 8.5. A single simulation shows that the effect size is -0.1251 which is less than 0.2 (the meaningful threshold we set).

Scenario 2: An Expected Effect

Authors (Names and Percentages): Xuanlong Liang 100%

Simulation

```r
set.seed(5780)
analyze.experiment <- function(the.dat) {
  require(data.table)
  setDT(the.dat)

  the.test <- t.test(x = the.dat[Group == "Treatment3",
                                  S], y = the.dat[Group == "Treatment1", S], alternative
= "greater")

  the.effect <- the.test$estimate[1] - the.test$estimate[2]
  upper.bound <- the.test$conf.int[2]
  p <- the.test$p.value

  result <- data.table(effect = the.effect, upper_ci = upper.bound,
                       p = p)

  return(result)
}
simu3.dat <- data.table(Group = c(rep.int(x = "Treatment3", times = n/2), rep.int(x = "T
reatment1", times = n/2)))
simu3.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.
5, sd = 10), digits = 1)]
simu3.dat[Group == "Treatment3", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 9.
5, sd = 10), digits = 1)]

analyze.experiment(the.dat = simu3.dat)
```

```
        effect upper_ci         p
1: -0.02533333      Inf 0.5645105
```

```r
#For simulations repeated 1000 times
B <- 1000
n <- 1800

set.seed(seed = 5780)
Experiment <- 1:B
Group <- c(rep.int(x = "Treatment3", times = n/2), rep.int(x = "Treatment1", times = n/
2))

sim.dat <- as.data.table(expand.grid(Experiment = Experiment, Group = Group))
setorderv(x = sim.dat, cols = c("Experiment", "Group"), order = c(1,1))
sim.dat[Group == "Treatment1", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 8.5,
sd = 10), digits = 1)]
sim.dat[Group == "Treatment3", S := round(x = rtruncnorm(n = .N, a=0, b=12, mean = 9.5,
sd = 10), digits = 1)]
dim(sim.dat)
```

```
[1] 1800000       3
```

```
exp.results <- sim.dat[, analyze.experiment(the.dat = .SD),
                       keyby = "Experiment"]

DT::datatable(data = round(x = exp.results[1:100, ], digits = 3),
              rownames = F)
```

Show 10 entries Search:

Experiment	effect	upper_ci	p
1	0.048		0.379
2	0.167		0.146
3	0.172		0.14
4	0.128		0.211
5	-0.025		0.564
6	0.081		0.303
7	0.158		0.156
8	0.296		0.031
9	0.323		0.022
10	0.257		0.055

Showing 1 to 10 of 100 entries Previous 1 2 3 4 5 … 10 Next

```
exp.results[, summary(effect)]
```

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
-0.32189  0.01692  0.12150  0.12204  0.22956  0.60989
```

```
exp.results[, sd(effect)]
```

```
[1] 0.1559607
```

```
exp.results[, mean(p < 0.05)]
```

```
[1] 0.196
```

```
exp.results[, mean(p > 0.05)]
```

```
[1] 0.804
```

```
1.96*exp.results[, sd(effect)]+0.12204
```

```
[1] 0.4277231
```

```
-1.96*exp.results[, sd(effect)]+0.12204
```

```
[1] -0.1836431
```

Analysis

A single simulation for the research question 3 shows that the effect size is -0.0253 which is lower than 0.2 (the meaningful threshold we set).

Now, we repeat the simulation of each scenario for 1000 times and get the results in the following table:

	Research Question	Scenario	Mean Effect	95% CI of Mean Effect	% of False Positives	% of True Negatives	% of False Negatives	% of True Positives
1	Question 1	No Effect	0.001108	[-0.3132, 0.3154]	5.1%	94.9%	-	-
2	Question 1	Effect	0.22416	[-0.0803, 0.5367]	-	-	56.8%	43.2%
3	Question 2	No Effect	0.002935	[-0.3098, 0.3157]	5.4%	94.6%	-	-
4	Question 2	Effect	0.116618	[-0.1952, 0.4285]	-	-	81.4%	18.6%
5	Question 3	No Effect	0.008035	[-0.3017, 0.3178]	5.3%	94.7%	-	-
6	Question 3	Effect	0.11453	[-0.1836, 0.4277]	-	-	80.4%	19.6%

Sensitivity Test

Now we let's change the assumption of the standard error to be 6 and do the simulation again. Since we are shrinking the standard error, we do not have to increase our sample size in order to get the power of 0.9. Therefore, we will keep our sample size to be 900 participants per group. The meaningful effect size threshold is now: 2/6 = 0.333. The results of the simulation are as follows:

	Research Question	Scenario	Mean Effect	95% CI of Mean Effect	% of False Positives	% of True Negatives	% of False Negatives	% of True Positives
1	Question 1	No Effect	-0.0019454	[-0.2890, 0.2851]	4.5%	95.5%	-	-
2	Question 1	Effect	0.57412	[0.28130, 0.8669]	-	-	2%	98%
3	Question 2	No Effect	0.0003583	[-0.2934, 0.2941]	5.2%	94.8%	-	-
4	Question 2	Effect	0.2799	[-0.0218, 0.5816]	-	-	42%	58%
5	Question 3	No Effect	0.00838	[-0.2961, 0.3128]	6.6%	93.4%	-	-
6	Question 3	Effect	0.2846	[0.0055, 0.5747]	-	-	40.1%	59.9%

As we can see from the sensitivity test, by shrinking the assumed standard error, we can have a higher percentage of True Positives in the 1000 simulations. However, the true standard error depends on the actual data gathered from the surveys. We assume its range is about [6, 10].

References

Authors (Names and Percentages): Ronglu Jiang (17%), Xuanlong Liang (17%), Hanbing Lu (17%), Yumeng Tian (17%), Ruizhe Zhu (17%), Shanrong Zhou (17%)

1. Holben, D. (2010). Position of the American Dietetic Association: food insecurity in the United States. Journal of the American Dietetic Association, 110(9), 1368-1377.

2. Coleman-Jensen, A., Gregory, C., & Singh, A. (2014). Household food security in the United States in 2013. USDA-ERS Economic Research Report, (173).

3. Food Security in the United States. USDA ERS - Food Security in the United States. (n.d.). Retrieved October 15, 2022, from https://www.ers.usda.gov/data-products/food-security-in-the-united-states/ (https://www.ers.usda.gov/data-products/food-security-in-the-united-states/)

4. U.S. Poverty Statistics. Federal Safety Net. (2022, September 20). Retrieved October 15, 2022, from https://federalsafetynet.com/poverty-statistics/#:~:text=%E2%80%8BFull%20U.S.%20 (https://federalsafetynet.com/poverty-statistics/#:~:text=%E2%80%8BFull%20U.S.%20) Poverty%20Statistics,the%20year%202021%20was%2011.6%25.

5. Mabli, J., & Worthington, J. (2014). Supplemental Nutrition Assistance Program Participation and Child Food Security. Pediatrics, 133(4), 610–619. https://doi.org/10.1542/peds.2013-2823 (https://doi.org/10.1542/peds.2013-2823)

6. Mabli, J., & Ohls, J. (2014). Supplemental nutrition assistance program participation is associated with an increase in household food security in a national evaluation. The Journal of Nutrition, 145(2), 344–351. https://doi.org/10.3945/jn.114.198697 (https://doi.org/10.3945/jn.114.198697)

7. Policy basics: The Supplemental Nutrition Assistance Program (SNAP). Center on Budget and Policy Priorities. (n.d.). Retrieved October 15, 2022, from https://www.cbpp.org/research/food-assistance/the-supplemental-nutrition-assistance-program-snap (https://www.cbpp.org/research/food-assistance/the-supplemental-nutrition-assistance-program-snap)

8. Adedokun, O. A., Plonski, P., Jenkins-Howard, B., Cotterill, D. B., & Vail, A. (2018). Healthy choices for every body adult curriculum improves participants' Food Resource Management Skills and Food Safety Practices. Journal of Nutrition Education and Behavior, 50(6), 638–644. https://doi.org/10.1016/j.jneb.2018.02.005 (https://doi.org/10.1016/j.jneb.2018.02.005)

9. Mello, J. A., Gans, K. M., Risica, P. M., Kirtania, U., Strolla, L. O., & Fournier, L. (2010, December). How is food insecurity associated with dietary behaviors? an analysis with low-income, ethnically diverse participants in a nutrition intervention study. Journal of the American Dietetic Association. Retrieved October 15, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3005628/ (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3005628/)

10. Six-Item Short form food security survey module - USDA. (n.d.). Retrieved December 4, 2022, from https://www.ers.usda.gov/media/8282/short2012.pdf (https://www.ers.usda.gov/media/8282/short2012.pdf)

11. Anders, J., & Rafkin, C. (2022). The welfare effects of eligibility expansions: theory and evidence from SNAP. Available at SSRN 4140433.

12. Falk, G., & Aussenberg, R. A. (2014). The supplemental nutrition assistance program (SNAP): Categorical eligibility.

13. Pollard, K., & Jacobsen, L. A. (2020). The Appalachian Region: A Data Overview from the 2014-2018 American Community Survey. Chartbook. Appalachian Regional Commission.