Three Essays on the Economics of Food, Health, and Consumer Behavior

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

in the Graduate School of The Ohio State University

By

Thadchaigeni Panchalingam

Graduate Program in Agricultural, Environmental, and Developmental Economics

The Ohio State University

2021

Dissertation Committee

Brian Roe, Advisor

Allen Klaiber

Zoë Plakias

Wuyang Hu

Abstract

There are many determinants of health such as individual dietary and health-related habits, constraints such as money and time, as well as market goods and services such as medical care, access to health insurance, and environmental conditions. In this dissertation, I focus on three key elements of household and individual consumption behaviors that are tied to economics of health and nutrition—policy, preferences, and consumption self-control.

In the first essay, I demonstrate how receiving subsidized health care services can lead to new patterns of household consumption, specifically, undertaking fewer preventative health measures by the targeted households. This topic has received less attention in the literature. To do this, I investigate the effects of recent Medicaid expansions on eligible households' quarterly food and non-food expenditures using state and time variation in Medicaid expansion. Using an event-study design, and a triple difference-in-differences framework, I find that the Medicaid eligible households from expansion states spent less on fresh produce per adult and more on over-the-counter medications and remedies while not changing their expenses on frozen fruits and vegetables which have similar nutritional value as fresh fruits and vegetables. The robust reduction in fresh produce expenditures and increase in expenditures on over-the-counter medications and remedies suggest that while expanded public health insurance increases formal healthcare

activity, it decreases informal preventative non-healthcare expenditures. These findings may begin to shift the focus in the literature on the unintended consequences of Medicaid expansion from sins of commission, i.e., moral hazard responses such as increased smoking, alcohol use and junk food consumption, to sins of omission, i.e., responses in which preventative health habits erode.

In the second essay, I focus on healthy eating in institutions such as schools and colleges, which is promoted in the US through programs such as improving access to local foods in cafeterias. Using a nationwide choice experiment survey, I model both individual and joint preferences of parents and students for locally sourced foods in school lunches. Results indicate that students and parents would prefer that locally produced items be added to school lunch menus. However, while parent and student preferences align on some aspects of locally sourced meal elements, their preferences are not identical, with parents displaying a higher willingness to pay for locally sourced vegetables and students displaying a higher willingness to pay for locally sourced fruit. Joint choices are influenced by both parties. Parents dominate the joint outcomes when the household income is lower, when students eat school lunch more frequently and in dyads featuring a female parent and female student compared to male parent-male student dyads. These findings may hold implications for efforts to promote locally sourced food elements in school lunches and the role of parent engagement in that process.

In the third essay, I investigate what characteristics of households, if any, that predict purchase of portion-controlled sizes of full calorie carbonated beverages (i.e., soda sold in less than 12 oz containers) and whether this behavior is associated with other healthy dietary habits. I find that household demographics including income, education, and presence of children or elderly are not associated with the purchasing behavior of full calorie carbonated beverages that are less than 12 oz. However, this behavior is negatively associated with the share of carbonated beverages that are diet and positively associated with the share of food expenditure dedicated to fresh produce, which are proxies used to capture healthy dietary habits. Overall, the findings suggest that there is an association between purchases of less than 12 oz of regular carbonated beverages (i.e., the portion-controlled sizes) and portion control behavior.

Dedication

To my parents Kulamaniy & Panchalingam, aunt Nirmala & uncle Mohanthas

Acknowledgments

I am extremely grateful to my advisor Professor Brian Roe who believed in my ideas, encouraged me to pursue them, guided me to ask the right questions, and patiently and promptly read through everything I wrote as a PhD student. He was always kind, willing to help, and never gave up on me. His good cheer made the many conversations shaping my research ideas and overcoming hurdles genuinely enjoyable. I thank him for finding and providing funding for me to use the NielsenIQ datasets which were immensely important for writing this dissertation. Dr. Roe guided me through the tough job market. Without his encouragement and constructive feedback, I would not have sustained through the tough times. The advisor-advisee match is crucial to a PhD student's success and I am grateful that I got a great advisor!

I am grateful to my dissertation committee members Professors Allen Klaiber, Zoë Plakias, and Wuyang Hu for their extensive constructive feedback and suggestions. I learned a great deal from each one of them and am very fortunate to have had them on my guidance committee. I thank them all for their flexibility in scheduling both my candidacy and defense exams in a pandemic despite their workloads. As an emerging female scholar, I draw a lot of inspirations from Dr. Plakias as a role model and I thank her for uplifting the voices of scholars like me.

I thank Professors Jeevika Weerahewa, Mariah Ehmke, Chian Jones Ritten, Dannele Peck, Benjamin Rashford, and Chris Bastian for encouraging me to pursue a PhD and for their close guidance during my MS program at the University of Wyoming. I was fortunate to meet these great people at UW who were so generous in sharing their knowledge and experience.

I thank my friends Asanka, Osama, Sachintha, Patalee, Kate, Christina, Sahar, Eric Davis, and Ziqian for making my time as a graduate student a memorable one beyond the office doors. They were there during both happy and sad moments. I survived my first year in the program because of my friends and all the nights we spent solving microeconomics problems in classrooms all over the campus. They also gave helpful feedback on my research. I thank Sandeep who helped me when I first came to the US and introduced me to the small Sri Lankan community in Laramie. I sincerely thank all these beautiful kind-hearted people and will cherish these memories forever.

I thank my wonderful husband Eric Swanson for being a supportive partner both professionally and personally. I thank him for reading and editing every draft I wrote for this dissertation. I could not have found the motivation to continue writing during this pandemic without his cheerful support despite being stuck with me in our living-room office for almost an entire year.

I am incredibly grateful to my parents Kulamaniy, Panchalingam, and my little brother Ariram for believing in me and being with me throughout my academic endeavors. My parents sacrificed a lot for me to achieve my academic dreams and I am forever indebted to them. I want to offer my sincere thanks to my aunt Nirmala and uncle Mohan in Canada who financially supported most of my primary and secondary education. I would also like to thank my in-laws Lou and Carol who were incredibly encouraging and supportive. They were sincerely interested in my progress and offered valuable advice from their lives.

I am very grateful for attending AEDE-OSU. I enjoyed my classes and learned a great deal from the excellent AEDE faculty. I also sincerely thank Sarah for being so helpful assisting with administrative matters and for being proactive in helping the international students. She is always cheerful and willing to help students. I also thank GSA at AEDE and their sincere efforts to maintain a feeling of community even during the pandemic. My sincere gratitude goes to all of them.

Vita

B.S. in Agricultural Technology and Management 2012

University of Peradeniya

M.S. in Agricultural and Applied Economics..2016

University of Wyoming

Publications

Panchalingam, T., Ritten, C. J., Shogren, J. F., Ehmke, M. D., Bastian, C. T., & Parkhurst, G. M. (2019). Adding realism to the Agglomeration Bonus: How endogenous land returns affect habitat fragmentation. *Ecological Economics*, *164*, 106371.

Jones Ritten, C., Bastian, C., Shogren, J. F., **Panchalingam, T.,** Ehmke, M. D., & Parkhurst, G. (2017). Understanding pollinator habitat conservation under current policy using economic experiments. *Land*, *6*(3), 57.

Ehmke, M., Jones-Ritten, C., Shogren, J., & **Panchalingam, T.** (2015). Integrating ecological and economic considerations for pollinator habitat policy. *Choices*, *30* (316-2016-7771).

Fields of Study

Major Field: Agricultural, Environmental and Development Economics

Table of Contents

List of Tables

List of Figures

Researcher(s) own analyses calculated (or derived) based in part on data from Nielsen Consumer LLC and marketing databases provided through the NielsenIQ Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business.

The conclusions drawn from the NielsenIQ data are those of the researcher(s) and do not reflect the views of NielsenIQ. NielsenIQ is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

1. Chapter: Effects of Public Health Insurance Expansions on the Non-health care Consumption Expenditures of Low-Income Households

1.1. Introduction

In this essay I interrogate whether public health insurance expansion leads to new patterns of non-health care consumption among beneficiaries. In doing so, first I examine whether there are overall spending changes in recurring total food and non-food expenditures (i.e., income effects), and then examine whether there are changes in preventative non-health care consumption expenditures (i.e., substitution effects). I define preventative non-health care as purchases that improve health or prevent ill-health but are not generally classified as health care, e.g., spending on fresh fruits and vegetables and vitamin supplements. Theoretically, it is possible that health insurance coverage may reduce preventative non-health care activities, but empirical research has provided mixed results. The new patterns of preventative non-health care consumption may occur because of improved finances due to subsidized health care (Baicker et al. 2013; Finkelstein et al. 2012; Mazumder and Miller 2016), changes in relative costs of health care and non-health care consumption (Dave and Kaestner 2009; Einav and Finkelstein 2018), increased access to preventive health care and primary care in general (Courtemanche et al. 2018; Miller, Johnson and Wherry 2019; Simon, Soni and Cawley 2017), or a combination of these factors.

Whether public health insurance expansion leads to new patterns of non-health care consumption among beneficiaries is an important policy question, firstly, because preventative non-health care consumption of healthy foods, vitamins, other diet aids, and contraceptives affects overall individual health and can either reinforce or undermine total health. An individual can maintain better health by eating a healthy diet (i.e., a preventative non-health care) or by taking prescription medications (i.e., health care). Greater consumption of certain types of food such as fruits and vegetables reduce the risk of cardiovascular disease, type-2 diabetes, some cancers, and obesity (Millen et al. 2016; Rolls, Ello-Martin and Tohill 2004), while greater consumption of foods with too much saturated fat, refined grains, added sugar, and sodium exacerbate these conditions (Binkley, Eales and Jekanowski 2000; Nicklas et al. 2001). Second, it is unclear whether the households with positive income effects due to obtaining public health insurance (Baicker et al. 2013; Finkelstein et al. 2012; Mazumder and Miller 2016) would spend the extra income on more non-health care consumption, including preventative non-health care consumption. The empirical evidence on changes in overall non-health care consumption (e.g., food, transportation, education) is mixed (Dillender 2017; Gruber and Yelowitz 1999; Leininger, Levy and Schanzenbach 2010; Levy, Buchmueller and Nikpay 2019).

This essay connects two strands of literature. First, it connects to the literature on preventive behavioral changes due to expanded health insurance coverage, which focuses on ex-ante moral hazard via sins of commission, i.e., increasing risky health behaviors such as smoking, drinking, or consuming unhealthy food and drink (Barbaresco, Courtemanche and Qi 2015; Brook et al. 1983; Cotti, Nesson and Tefft 2019; He, Lopez and Boehm 2020; Simon, Soni and Cawley 2017). However, the provision of more health care need not only spur increases in risky behaviors but could also affect preventative behaviors—health sins of omission, i.e., a lack of exercise or diets with fewer fruits and vegetables. The healthfulness of the overall diet could be affected by the proportions of healthy and unhealthy food items. The question of whether health care provision induces health sins of omission has received less attention in the literature with the exception of studies assessing the effects of expanded coverage on the frequency of exercise, where results are mixed (Brook et al. 1983; Courtemanche and Zapata 2014; Dave and Kaestner 2009; De Preux 2011; Simon, Soni and Cawley 2017). This is because in addition to the increased monetary costs (e.g., health care costs and lost wages) of poor health behaviors, there are also non-monetary costs such as physical pain and suffering which are difficult to measure, and most of the negative health consequences do not occur until many years after the behaviors take place. Even though changes in food consumption behavior have received less attention, Nguyen et al. (2016) assess diet quality via two-sided *t*-tests to compare health risk factors between expansion and non-expansion states before the Affordable Care Act public health insurance expansions began. In the related literature on food security, Himmelstein (2019) finds that Medicaid expansion was associated with a significant adjusted 2.2-percentage-point decline in rates of very low food security in the states that expanded Medicaid between 2014 and 2016.

The second strand of related literature is on the relationship between health insurance and changes in non-health care consumption expenses such as food, housing, and education. These studies found mixed results due to differences in programs they studied and methods used, and did not focus on preventative non-health care consumption expenditures. Gruber and Yelowitz, (1999) found that the Medicaid expansions in the late

1980s and early 1990s lowered the asset holding of eligible households by a net 16.3 percent while increasing consumption of non-durable goods by $538 in 1987 dollars. Leininger, Levy and Schanzenbach (2010) found strong positive effects of expansions on transportation expenditures in state Child Health Insurance Programs between 1996 and 2002. In contrast, Dillender (2017) found no evidence of non-health spending changes due to an additional person in the family being eligible for Medicaid eligibility in the 2000s. Levy, Buchmueller and Nikpay (2019) found no evidence of changes in non-health care consumption expenditures in response to recent Medicaid expansion using Consumer Expenditure Survey data collected by the Census Bureau.

This study contributes to the limited literature on the relationship between public health insurance and non-health care consumption with a specific focus on preventative non-health care consumption. I show that provision of more health care could also affect preventative behaviors in addition to risky health behaviors by exploring expenditures on goods known to help prevent disease, such as fruits and vegetables, and supplementary nutrition and preventative care items which have received less attention in the literature. This study also contributes to the literature on the effects of public health insurance coverage expansions on financial outcomes by providing new evidence of changes in the range of recurring consumption expenditures.

I examine how subsidizing health care expenditures may affect the preventative non-health care expenditures of targeted households by exploiting a quasi-experimental change in the Medicaid expansion under the Affordable Care Act (ACA). To identify the effects of Medicaid expansion on preventative non-health care expenditures, I use an event-study design, and a triple difference-in-differences approach, which are appropriate to exploit the variations in the time of adoption by states and the eligibility criterion under the ACA Medicaid expansions. The expenditure data come from NielsenIQ Homescan Consumer Panel, which uses at-home scanner technology to track regular grocery and other household purchases from stores the households visit for their regular groceries and for other recurring purchases from pharmacies, large retailers, and supercenters.[1] I analyze data from 2011 to 2017. Compared to data from the U.S. Census Bureau's Consumer Expenditure Survey (CEX), which collects data on US households' expenditure on housing, food, education, transportation, and health care, the data used in this study are less nationally representative and do to not capture all consumption categories captured in CEX such as housing, transportation, education, and health care.[2] Therefore, assessing the changes in overall financial wellbeing of the households is out of scope for this study.

My findings show that there is a significant decrease in fresh produce expenditures (18.6% decrease from pre-expansion levels) and an increase in frozen food expenditures, but no significant increase in frozen vegetables and fruits expenditures (4.4% increase from pre-expansion levels) per adult equivalent unit per quarter among eligible households in response to Medicaid expansion. Among non-food categories, there is a significant increase

[1] Some examples of products purchased and recorded in the data include most at home food purchases, toys and sporting goods, some apparel, electrical items such as mobile phones, household appliances such as microwaves, household cleaning supplies, diet aids, non-prescription medications, laundry supplies etc.

[2] Levy, Buchmueller and Nikpay (2019) used the CEX data and found no effects on total non-healthcare consumption in response to the recent Medicaid expansions. While conceptually and methodologically different from Levy, Buchmueller and Nikpay (2019), this paper accounts for changes over the time and more importantly, staggered implementation of the Medicaid expansion and overcomes data limitations in Levy, Buchmueller and Nikpay (2019) that are due to several missing states.

in expenditures on health and beauty products (10.3% increase from pre-expansion levels) of which more than 90% comes from increases in over-the-counter medications and remedies, with no changes in vitamins and diet supplements. The robust reduction in fresh produce expenditures and increase in expenditures on over-the-counter medications and remedies are a novel finding in the literature. This linkage suggests that while expanded public health insurance increases formal health care activity (i.e., over-the-counter medications), it decreases informal preventative non-health care expenditures (i.e., fresh produce expenditure). Consistent with previous literature (Dillender 2017; Levy, Buchmueller and Nikpay 2019), no significant change in total expenditure is documented, suggesting that even if there are positive income effects, they do not manifest in recurring household food and other consumption expenditures.

1.2. Data

The debate about universal health coverage is among the top health-related policy discourses in the United States. Currently, employer-sponsored health insurance covers about 55 percent of the US population, while public insurance covers the elderly, the disabled, and low-income children, and adults (U.S. Census Bureau 2018). The Medicaid program, which was created by the Social Security Amendments of 1965, provides health insurance to 1 in 5 Americans.[3] There have been substantial changes to the eligibility for Medicaid since then. The Patient Protection and Affordable Care Act (ACA) of 2010

[3] Kaiser Family Foundation estimates based on the Census Bureau's American Community Survey, 2008-2017.

expanded Medicaid for all non-Medicare eligible individuals under 65 with incomes up to 138% of the Federal Poverty Line (FPL) defined by Unites States Health and Human Services (HHS) (The Patient Protection and Affordable Care Act 2010; Sommers and Rosenbaum 2011). Under the ACA, all eligible adults are guaranteed a basic benefit package that meets essential health care benefits (Kaiser Family Foundation 2013). However, the Supreme Court ruling on the constitutionality of the ACA made the decision to expand Medicaid optional for states. Following the Supreme Court decision in 2012, only 32 states and the District of Columbia (DC) had expanded Medicaid from 2014 to 2016, which is the timeframe considered for Medicaid expansion in this study (Appendix Figure A.1). Only 36 states and DC have adopted the Medicaid expansion in the period of 2014-2019 (Kaiser Family Foundation 2019).

I use NielsenIQ Homescan Consumer Panel (NHCP) data on purchases made by consumers from 2011 to 2017.[4] The NHCP consists of a panel of households who scan their purchases after all grocery and other shopping trips from stores they usually visit using an at-home scanner technology. The data captures a variety of store types; this includes grocery, drug, mass merchandise, superstore, club, convenience, and dollar stores.[5] The stores that are visited by the panelists are not restricted to the stores where NielsenIQ receives point of sale (POS) data for the NielsenIQ's retailer scanner data (not used in this study). However, if a panelist visited a store that is covered in retailer scanner data, then the panelist is not required to enter the prices paid. This is an attempt to minimize

[4] The data were obtained from Kilts Center for Marketing at the University of Chicago.
[5] The store types are further subdivided into channel types such as apparel, footwear, automotive, computer, etc. For example, in 2017 about 0.1% of all items purchased are from apparel stores and about 0.01% of items are from computer stores.

the data entry burden of panelists. Instead, NielsenIQ imputes the price as an average weighted price for the item that week in that store. The imputed prices are indistinguishable by the researcher from the prices recorded by the panelists in the data set.

The dataset comprises a representative panel of 40,000-60,000 active panelist households in each panel year with a retention rate of 80% from one year to the next. The sampling of panelists follows a proportionate random sampling approach in which the key demographic characteristics of the panelists are matched to the demographics of the continental US population and regular checks are made to ensure the representativeness.[6] NielsenIQ samples all states except Alaska and Hawaii. The are 8,819 Medicaid eligible households in the sample which is 10.6% of the total number of unique households in the sample. Nationally about 18% of the U.S. population is enrolled in Medicaid. These figures are not directly comparable since national level statistics are at the individual level rather than at the household level. The incentives for NHCP participation include monthly prize drawings, gift points redeemable for merchandise and gift cards, sweepstakes, and contests; the incentives are designed not to influence purchasing habits. No account-specific coupons are provided to avoid any impacts on selection of outlets and products and incentives are regularly tested to check if they are correlated with retention rates.[7]

The Universal Product Code (UPC) of each purchased item is recorded, and consumers provide information on the price (if required by the NielsenIQ), quantity, store

[6] NielsenIQ uses nine demographic characteristics to balance the sample which are household size, income, head's age, female and male heads' education levels, presence of children, race, ethnicity, and head's occupation.

[7] He, Lopez and Boehm (2020) provide evidence that Medicaid expansion in 2014 did not affect household program attrition in the NielsenIQ panel.

information and package details.[8] The data set also includes information on various demographic characteristics of the households including household size, income, education of the household heads, age composition, employment, race/ethnicity, and zip code of residence. A previous study tested the accuracy of NHCP data found that it captures about 80% of the total calories reported by National Health and Nutrition Examination Survey (NHANES) (Oster 2018).[9] The NHCP is not without measurement errors due to households failing to report trips or products and misreporting the information on stores and dates. However, several validation studies show that the reported data have a high level of accuracy while data misreporting is comparable to other commonly used datasets in economics and unlikely to affect the averages calculated (Einav, Leibtag and Nevo 2008; Zhen et al. 2009).

While NHCP data is available at a weekly reporting level, many items are not purchased on a weekly basis. So, I aggregate the data for each household across calendar quarters to ensure purchasing patterns that are representative of each household's normal patterns while still being able to precisely exploit the timing of state expansion of coverage (Appendix Table A.1). Further, data on consumer expenditure surveys such as CEX use quarterly time units, so using this interval makes comparison with other studies easier. Also, research calculating the healthfulness of grocery purchases use quarterly shopping

[8] The products that do not use standard UPC codes (e.g., random weight product) are called Magnet data and only a subset of households regularly report Magnet Product purchases. These products can still be grouped into different product categories.
[9] NHANES is an annual survey of a nationally representative sample of about 5,000 persons in the US and it is used to assess the health and nutritional status of adults and children. The survey includes demographic, socioeconomic, dietary, medical conditions, and health behavior-related questions.

baskets of total expenditures (Volpe and Okrent 2012).[10] Twenty-five states and DC expanded Medicaid in January 2014 while seven states adopted after January 2014. Appendix Table A.1 shows the state expansion dates.[11]

1.3. Methods

To measure the impacts of Medicaid expansion on various expenditure outcomes, I estimate the following event-study model:

$$Y_{hsyq} = \beta_0 + \sum_{j=-8}^{j=8} \beta_{0j} I(Expand_{syq} = j) + \sum_{j=-8}^{j=8} \beta_{1j} I(Expand_{syq} = j) \times Eligible_{hsyq} + \beta_E Eligible_{hsyq} + \beta_x X_{hsyq} + \alpha_h + \tau_y + \gamma_q + \delta_{syq} + \varepsilon_{hsyq}. \quad \text{(Eq.1.1)}$$

In the above equation above, Y_{hsyq} is the outcome for household h in state s, in year y, and quarter q. Outcome variables are explained in detail in section 3.1. $I(Expand_{syq} = j)$ are indicator variables for quarters leading up to the expansion and quarters following the expansion in state s. The omitted expansion quarter is $j = -1$. Quarters ≥ 8 together and quarters ≤ -8 together are combined into two single indicator variables. $Eligible_{hsyq}$ is an indicator equal to one if a household has an annual income below 100% of the federal poverty line (FPL). I include controls for time varying

[10] Two of the USDA's diet healthfulness measures are calculated in terms of various food group expenditure shares. The food groups' recommended expenditure shares are calculated using those food group prices from quarterly food at home price database (QFAHPD). This is one of the reasons that USDA uses quarterly food baskets.
[11] Although Wisconsin did not adopt the ACA expansion, it offers Medicaid to adults below 100% of FPL; therefore, it is considered an adoption state for the purpose of this study.

household characteristics[12] (X_{hsyq}); household (α_h), year (τ_y), and quarter (γ_q) fixed effects; and state specific linear time trends $\left(\delta_{syq}\right)$.

The identifying variation of this model comes from comparing households who are eligible for Medicaid expansion with households who would be ineligible during quarters leading up to and during quarters after Medicaid expansions. The identifying assumption is that the households that are eligible and ineligible did not have differential expenditure trends in expanding and non-expanding states. I verify this by graphing the trends in outcomes in expanded and non-expanded states between eligible and ineligible households. I test the null hypothesis of identical pre-trends using a joint F-test of pre-expansion coefficients.

I do not directly observe whether a household receives Medicaid, thus, I use the eligibility for Medicaid as a proxy. A household in an expansion state is eligible for Medicaid if its income is below 100% of FPL. This cutoff of 100% of FPL is used instead of 138% of FPL (which is the official cutoff) because the households that have incomes between 100% and 138% of FPL are still eligible for health insurance marketplace subsidies in non-expansion states (Simon, Soni and Cawley 2017). The households between 100% and 138% of FPL are excluded from the main analysis.[13] The income recorded in NHCP has a lag of approximately two years, therefore the FPL that I used,

[12] All models include controls for age, marital status, years of education of the household head, race, ethnicity of the household, number of household heads, hours of employment, and presence of children in the household.

[13] The Affordable Care Act of 2014's individual mandate requires all Americans to obtain health insurance or pay a tax penalty. As this mandate affects both expansion and non-expansion states similarly, it does not affect the results from this study. The individual mandate was repealed in 2019.

based on the Department of Health and Human Services federal poverty guidelines, is matched to the year corresponding to when the income was reported.[14] Further, there are five states with significant prior expansions. Persons who are 65 and older are eligible for Medicare and persons under 26 are eligible for dependent coverage provision. Thus, the sample used in the main analysis does not include the following: households with income between 100 and 138% of the FPL, households from states with significant prior expansions, and households with a head of age 65 and above or below age 26. If the household has more than one head, both heads' ages are considered. In a robustness check I relax income as well as age eligibility.

Following the event-study model, I also estimate the following triple difference-in-differences model to get the average effects (Eq.1.2). The event study analysis estimates dynamic policy effects and also functions as a test for the parallel trend assumption while the triple difference model gives an average estimate. The notation follows Equation 1.1 except for $Expand_{syq}$, which is an indicator variable that denotes whether state *s* expanded Medicaid in year y and quarter q. The triple difference model's three differences come from the state (expanded or not), the timing (before and after the expansion), and the eligibility within the state (eligible and ineligible households). The coefficient of interest is β_I.

$$\begin{aligned} Y_{hsyq} = \beta_0 + \beta_M Expand_{syq} + \beta_E Eligible_{hsyq} + \beta_I Expand_{syq} \\ * Eligible_{hsyq} + \beta_x X_{hsyq} + \alpha_h + \tau_y + \gamma_q + \delta_{syq} + \varepsilon_{hsyq} \end{aligned} \qquad \text{(Eq.1.2)}$$

[14] This potentially introduces measurement error regarding Medicaid eligibility; however, this avoids a potential source of bias arising from household incomes responding to Medicaid expansions since the incomes considered for eligibility precede the income changes that potentially result from changes in Medicaid status.

Table 1.1 shows summary statistics of the household characteristics of the sample used in this study. Except for income, marriage, and employment characteristics, households in the eligible and ineligible groups are similar in other characteristics such as education, race, and household size. Medicaid enrollment by race/ethnicity shows that 40% of the enrollees are white, 21% are Black, and 25% are Hispanic (Kaiser Family Foundation 2019). While education level is not a perfect predictor of the eligibility, other studies have used having less than a high school degree as a proxy for eligibility cutoff for Medicaid. In the NielsenIQ sample the education attainment of eligible households is higher than national averages.

Table 1.1. Summary Statistics of the Household Characteristics for Years 2011-2017 of the NielsenIQ sample used

	Eligible households		Ineligible households		
Variable	Mean	Standard Deviation	Mean	Standard Deviation	Difference in means
Income (2011$)	21,207	22,404	71,810	35,481	-51,603***
Household size	2.581	1.665	2.623	1.348	-0.041***
Years of Education	13.830	1.962	14.800	1.838	-0.975***
Age	50.670	9.525	50.570	9.379	0.103**
White	0.788	0.409	0.799	0.401	-0.011***
Black	0.124	0.329	0.110	0.313	0.014***
Asian	0.024	0.154	0.042	0.200	-0.017***
Hispanic Origin	0.066	0.249	0.071	0.256	-0.005**
Presence of Children	0.321	0.467	0.313	0.464	0.008**
Married	0.431	0.495	0.694	0.461	-0.263***
Employed	0.485	0.446	0.802	0.314	-0.318***
Weekly Work Hours	16.280	15.440	30.320	12.210	-14.040***
Number of households	8,819		74,219		

Notes: Author's calculations from NHCP. The Consumer Price Index is used to deflate income. The number of households shown are the unique number of households across all years and each household contributes to more than one household-quarter observation. A t-test was conducted to compare the means. *, **, *** indicate significance at the 10%, 5%, 1% level, respectively.

1.3.1. Outcome Variables

The primary outcome variable Y_{hsyq} is expenditure (2011$) per adult equivalent[15] of household h from state s in quarter q in year y on a given category. I use expenditure per adult equivalent unit so that the expenditures are comparable across households with different sizes and compositions. All outcomes are adjusted for inflation. The major categories are food, non-food, and total expenditure and the sub-categories derived from departments as defined by NHCP. The major food categories are dairy, deli, dry grocery, frozen food, fresh produce, and packaged meat. The major non-food categories are non-food grocery, general merchandise, health and beauty products, and alcohol.[16] Health and beauty products were disaggregated as vitamin and diet aids, hygiene and sanitary products, over-the-counter medications/remedies, and beauty products. The sub-category of frozen fruits and vegetables are also analyzed as they are considered to contain similar nutrition as fresh fruits and vegetables. Table 1.2 shows summary statistics of the major outcome categories. The total food expenditures for the eligible households on the lower bound of the income strata is about 79% of what is reported in the Consumer Expenditure Survey. It appears under-reporting is greater among the eligible households, but this does not affect the estimates as households are observed both before and after the expansion so similar underreporting is assumed to cancel. The expenditure in the non-food category is not directly comparable as it is not defined similarly across the two datasets.

[15] Adult-equivalent units are a measure of household size that accounts for ages and relative calorie needs of household members (Oster 2018). The Consumer Price Index is used to deflate expenditures.

[16] If a panelist, e.g., buys a printer or a cellphone along with groceries, both the printer and the cellphone are scanned and included under the general merchandise category.

Table 1.2. Summary Statistics of Household Level Quarterly Expenditures per Adult Equivalent Unit for Years 2011-2017 of the NielsenIQ Sample Used

	Eligible households		Ineligible households		
Variable	Mean ($)	Standard Deviation ($)	Mean ($)	Standard Deviation ($)	Difference in means
Total Expenditure	509.50	419.30	528.50	391.50	-18.98***
Total food expenditure	320.30	234.80	324.40	220.70	-4.05***
Dairy	37.39	32.84	39.61	31.29	-2.22***
Deli	22.17	54.65	24.57	57.99	-2.40***
Dry Grocery	173.30	134.4	167.90	120.20	5.36***
Fresh Produce	17.87	26.36	23.72	29.78	-5.85***
Frozen Food	52.67	53.52	52.00	49.50	0.67
Packaged Meat	16.94	19.59	16.55	18.02	0.39
Total non-food expenditure	175.80	234.70	186.40	205.60	-10.58***
General Merchandise	67.38	130.40	78.83	130.20	-11.45***
Health and Beauty	50.64	115.30	55.23	77.44	-4.58***
Non-Food Grocery	57.82	93.07	52.37	70.22	5.45***
Alcohol	13.32	54.66	17.67	53.03	-4.35***
Number of households	8,819		74,219		

Notes: Author's calculations from NHCP, all expenditures are adjusted for inflation. *, **, *** indicate significance at the 10%, 5%, 1% level, respectively.

1.3.2. Intent-to-treat (ITT) Effects

I use eligibility rather than actual treatment assignment to estimate the effects of Medicaid expansion on preventative non-health care consumption because the NHCP data does not include information on who received Medicaid under the ACA expansion. Therefore, the estimates are the effects of intent-to-treat (i.e., those who are eligible are intended to receive the Medicaid) and not the average treatment effects as not all eligible households would participate in Medicaid.

Intent-to-treat effects underestimate the true treatment effects (Angrist and Pischke 2014). ITT is equivalent to estimating the effect of the ability to apply for Medicaid on the outcomes of interest rather than receiving Medicaid itself (Taubman et al. 2014). Not participating in Medicaid even though one were eligible would be same as when only a fraction of the eligible population who are offered the treatment take it up in a randomized control trial.

In designs where the actual treatment (i.e., Medicaid receipt) is distinct from the variable that is randomly manipulated to assign the treatment (i.e., income eligibility), the correct design has to compare all those who are eligible to all those who are ineligible (Duflo, Glennerster and Kremer 2007). This is because the households who are eligible and actually participate in Medicaid could be different due to selection compared to the households who are eligible and do not participate.[17] Dividing the ITT by the difference in compliance rates between treatment and control groups captures the causal effect of Medicaid expansion on those who actually participate in Medicaid (Angrist and Pischke 2014). Currently about 25% of those who are eligible for Medicaid and Children's Health Insurance Program (CHIP) are uninsured (Kaiser Family Foundation 2019). A conservative assumption would give roughly 75% of compliance rate hence:

$$\text{Effect of treatment on the treated} = \text{Intent} - \text{to} - \text{treat}/0.75 \quad \text{(Eq.1.3)}$$

[17] The concern that some households in the control group being treated can be mitigated since the eligibility criteria is based on income. To make sure this is a valid assumption, later in the robustness tests, I test whether the household incomes react to the treatment and show that this is not the case. This eliminates the concern that income is manipulated to get Medicaid by the control group.

The effects identified in this study are ITT as shown in Equation 1.3. Here we assume that there are no always-takers (i.e., those who take the treatment regardless of treatment assignment).

1.4. Results and Discussion

1.4.1. Parallel Trends

Figures 1.1, 1.2, and 1.3 plot average quarterly expenditures on food, non-food, and total expenditures for Medicaid eligible and ineligible households from expansion and non-expansion states over the 2011-2017 period. The plots suggest that the relative pre-expansion trends between eligible and ineligible households are similar in expansion and non-expansion states. Event study estimates are used to test for the presence of differential pre-trends between Medicaid eligible and non-eligible households in expansion and non-expansion states before expansion (Appendix Tables A.2, A.3, A.4). All models include controls for age, marital status, years of education of the household head, race, ethnicity of the household head, number of household heads, hours of employment, and presence of children in the household. Additionally, all models include household, year and quarter fixed-effects, and state-specific time trends. Robust standard errors clustered at the state-level are in parentheses.

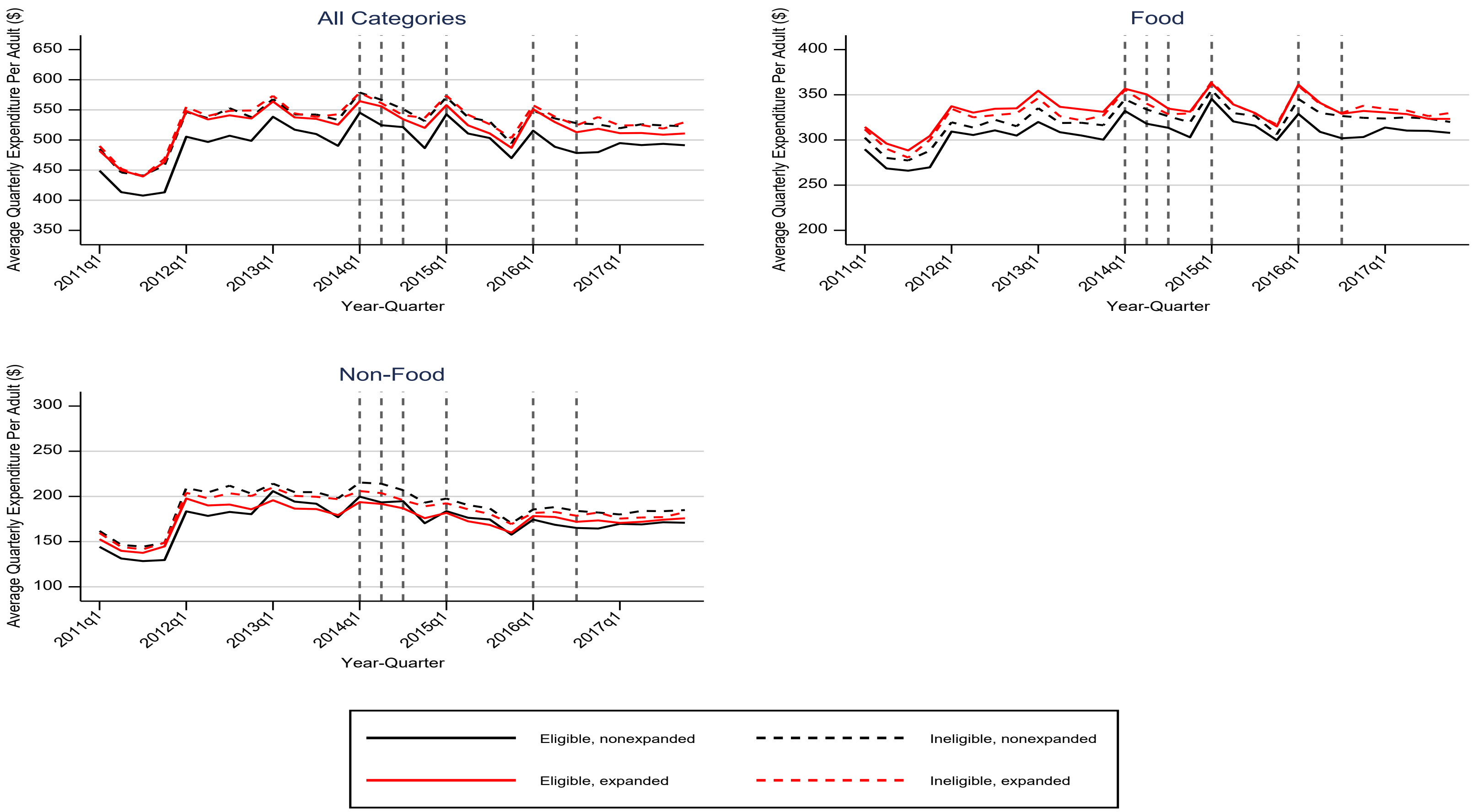


Figure 1.1. Trends of average quarterly expenditures (2011$) per adult equivalent unit for eligible and ineligible households in expansion and non-expansion states, 2011-2017

Note: Vertical dashed lines indicate the year and quarter that different states expanded Medicaid.

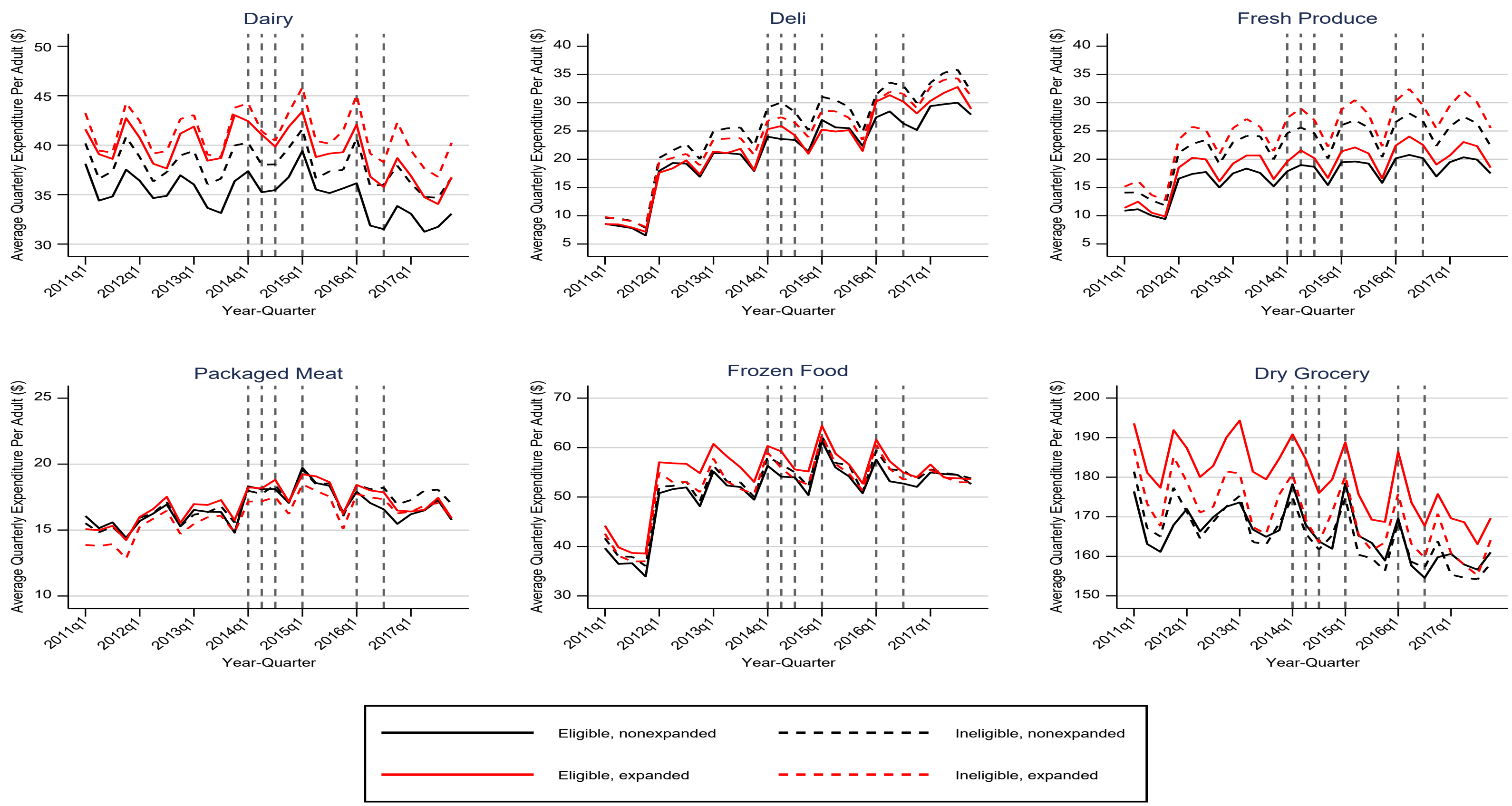


Figure 1.2. Trends of average quarterly food expenditures (2011$) per adult equivalent unit for eligible and ineligible households in expansion and non-expansion states, 2011-2017

Note: Vertical dashed lines indicate the year and quarter that different states expanded Medicaid.

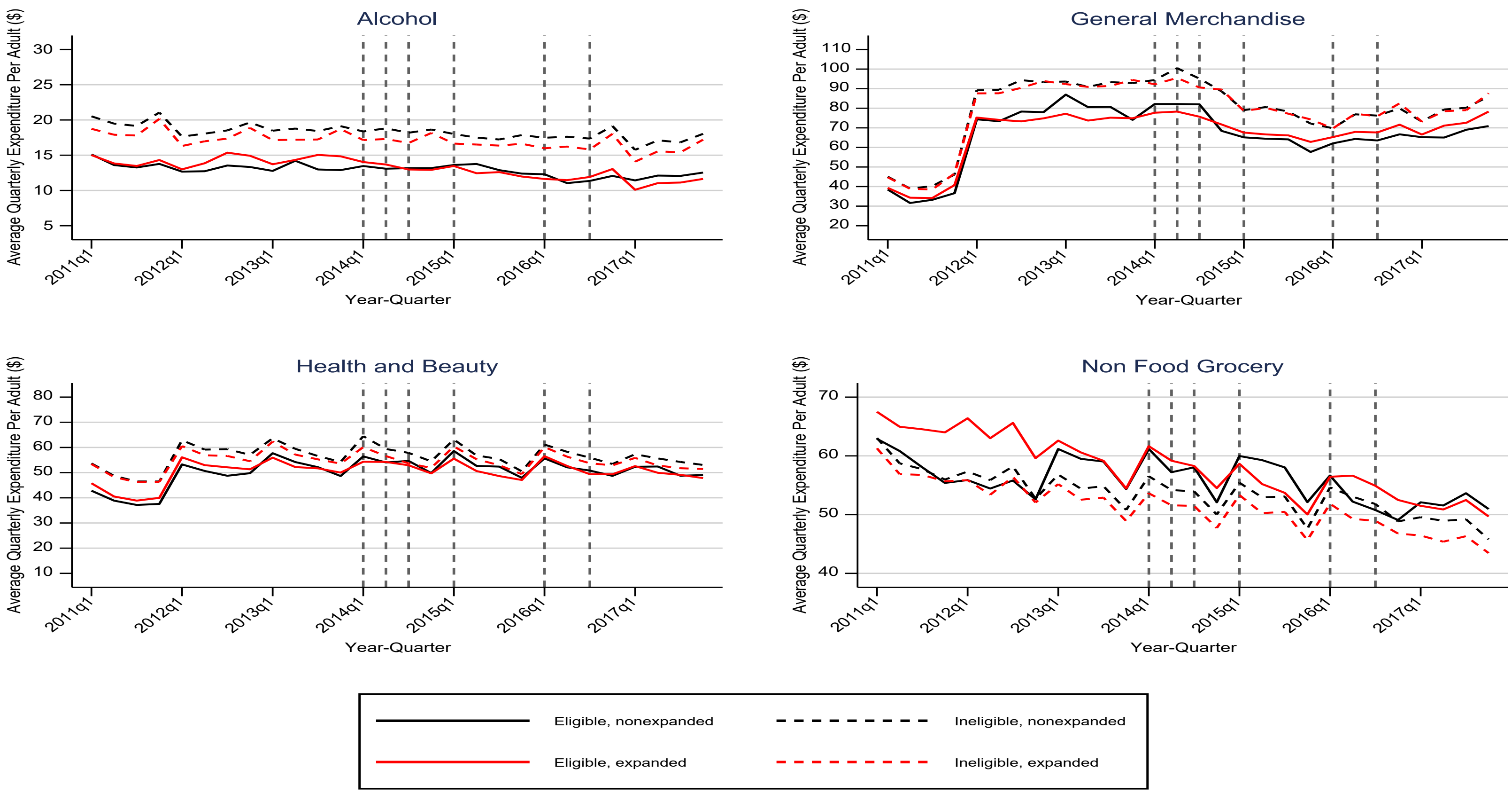


Figure 1.3. Trends of average quarterly non-food expenditures (2011$) per adult equivalent unit for eligible and ineligible households in expansion and non-expansion states, 2011-2017.
Note: Vertical dashed lines indicate the year and quarter that different states expanded Medicaid.

I estimate the effects of Medicaid expansion on the spending of low income households using Equations 1.1 and 1.2 and plot the coefficients against quarters relative to expansion as presented in Figures 1.4 through 1.8. Strong seasonal effects in the figures are justifiable given that most expansions happened at the same time. The quarter before the expansion is the omitted category. The triple-difference model coefficients and their standard errors are written on the event-study plots with their significant levels indicated. The detailed tables of event-study estimates are presented in appendix tables A.2, A.3, and A.4. The event-study estimates suggest that, overall, there is no strong evidence of differential pre-existing trends in quarterly expenditure between expansion and non-expansion states. The *p*-value of the *F*-test (to test whether the β_{1j} coefficients of Equation 1.1 where j = -8 to -2 are jointly zero) are presented on the same appendix tables. The *F*-test rejects the parallel trends for dry groceries and alcohol at the 5% significance level. Other categories' effects do not suggest existing pre-trends at conventional significant levels.[18]

1.4.2. Effects on Total Expenditures

In all estimations, the outcomes measured are the inflation-adjusted quarterly expenditures (2011$) per adult equivalent unit at the household-level. The eligible households' total expenditure summed across all categories, or total expenditure on food or non-food categories, do not show any impacts due to Medicaid expansion even though the effect on

[18] I also test for the parallel trends in the triple difference model by estimating the model for years before 2014, and the null of parallel trends could not be rejected in any expenditure category.

total and food expenditures is negative in magnitude (Figure 1.4). This suggests that even if there are positive income effects, they do not manifest in recurring household food and other consumption expenditures as captured in this data panel.

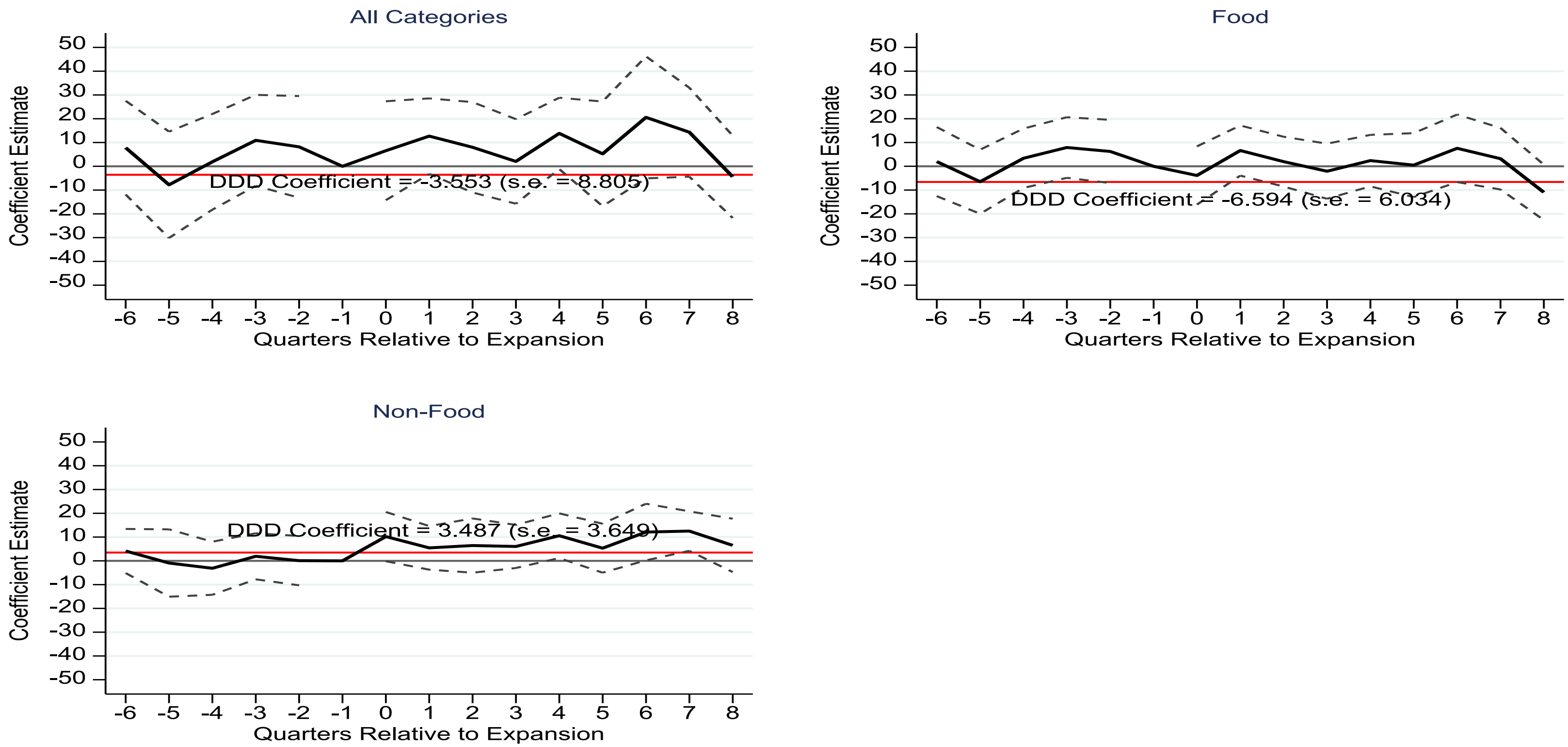


Figure 1.4. Effects of Medicaid expansion on eligible households' quarterly expenditure (2011$) per adult equivalent unit – total expenditures, 2011-2017

Notes: the red line denotes the triple difference model coefficient. The dark thick solid black line denotes the coefficients of interest from event study. The short, dashed lines are the confidence intervals of the event study coefficient estimates. The solid gray line denotes no change line (y-zero). The omitted expansion quarter is *j = -1*.

1.4.3. Effects on Food Expenditure

Among food categories, there is a significant negative effect ($2.88 per adult equivalent unit) on quarterly fresh produce expenditure, and a significant negative effect ($2.13 per adult equivalent unit) on quarterly frozen food expenditure (Figure 1.5). These respectively are about 18.6% and 4.4% from pre-expansion mean expenditures of the eligible households in expanded states with some variations over time as shown by the event study graphs.[19] The reduction in fresh produce expenditure is persistent across post-expansion quarters, while the reduction in frozen food expenditures is largely driven by a strong reduction eight or more quarters post-expansion. There are no effects on expenditure on other food categories.

The negative effect on the fresh produce expenditure shows that health insurance expansion may have unintentionally worsened diet habits. However, the negative effects on fresh and frozen food expenditures together could also mean that the households substitute foods at home with food away from home after the expansion due to an income effect as well as increase in confidence in their overall health and life quality. This claim is not testable without data on food-away-from-home purchases. Todd, Mancino and Lin (2010) show that food away from home is associated with an increase in daily caloric intake and a reduction in diet quality using dietary recall data from the 1994-96 Continuing Survey of Food Intakes by Individuals (CSFII) and the 2003-04 NHANES which covers a broader sample of US consumers. Consequently, a substitution of food-at-home

[19] The percentages are calculated based on the average expenditure of the eligible households live in expanded states before the expansion.

expenditure with food-away-from-home expenditure would still point to a reduction in overall diet quality. Further disaggregation of the expenditure on frozen food categories shows no change in the expenditures on frozen fruits and vegetables category (Figure 1.6).

Dry grocery, the category that would contain canned fruit and vegetables, had no significant changes associated with Medicaid expansion, and has an estimate around zero. This suggests that if eligible households increased their shelf-stable fruits and vegetable purchases, these increases were cancelled out by substitution away from other dry goods. This is unlikely. Furthermore, research shows that households with more income will normally substitute fresh for processed fruits and vegetables (Ferrier and Zhen 2014). Overall, these findings suggest that the expenditure reduction in high quality fruits and vegetables likely comes solely from the reduction in fresh produce expenditures and the reduction in frozen food expenditure comes from categories other than frozen fruits and vegetables.

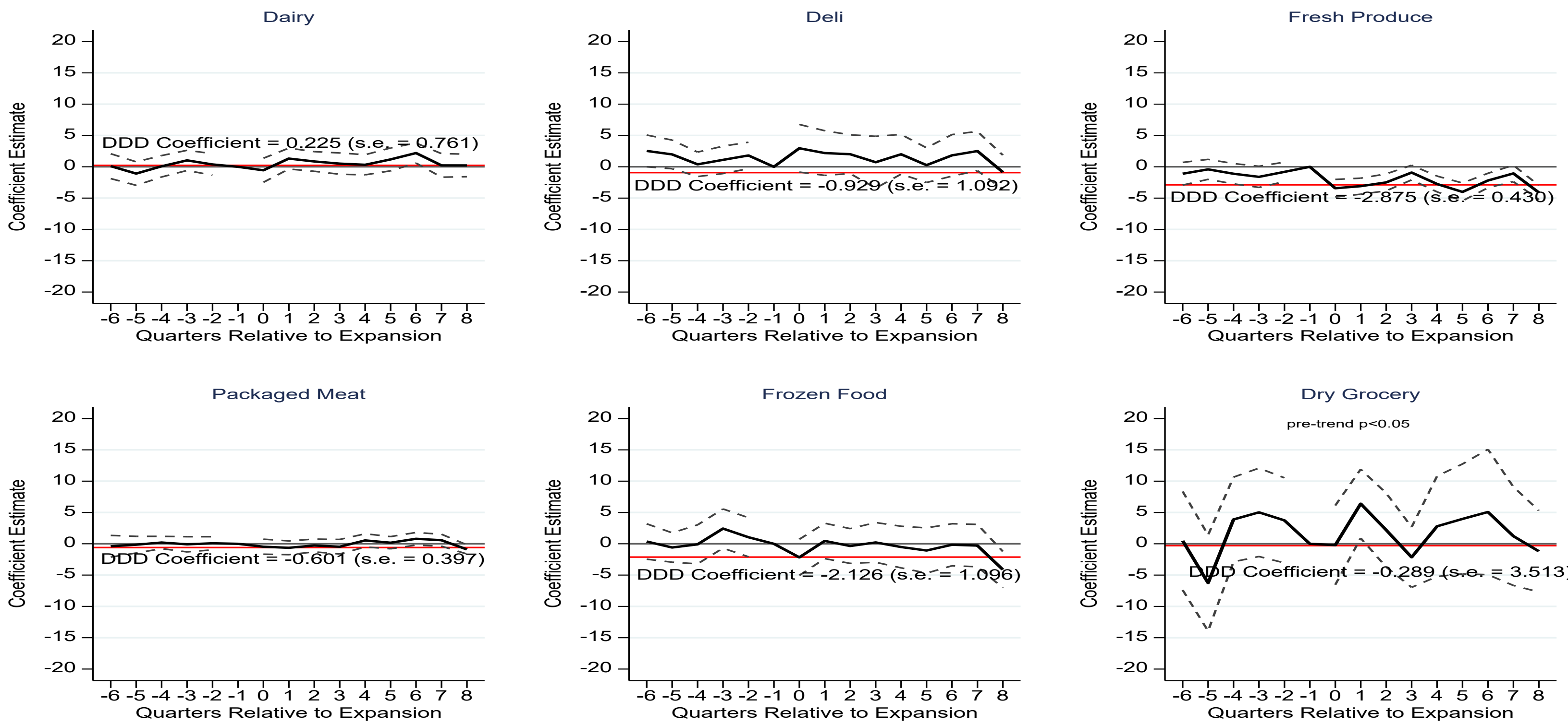


Figure 1.5. Effects of Medicaid expansion on eligible households' quarterly expenditure (2011$) per adult equivalent unit – food expenditures, 2011-2017

Notes: the red line denotes the triple difference model coefficient. The dark thick solid black line denotes the coefficients of interest from event study. The short, dashed lines are the confidence intervals of the event study coefficient estimates. The solid gray line denotes no change line (y-zero). The omitted expansion quarter is *j = -1*.

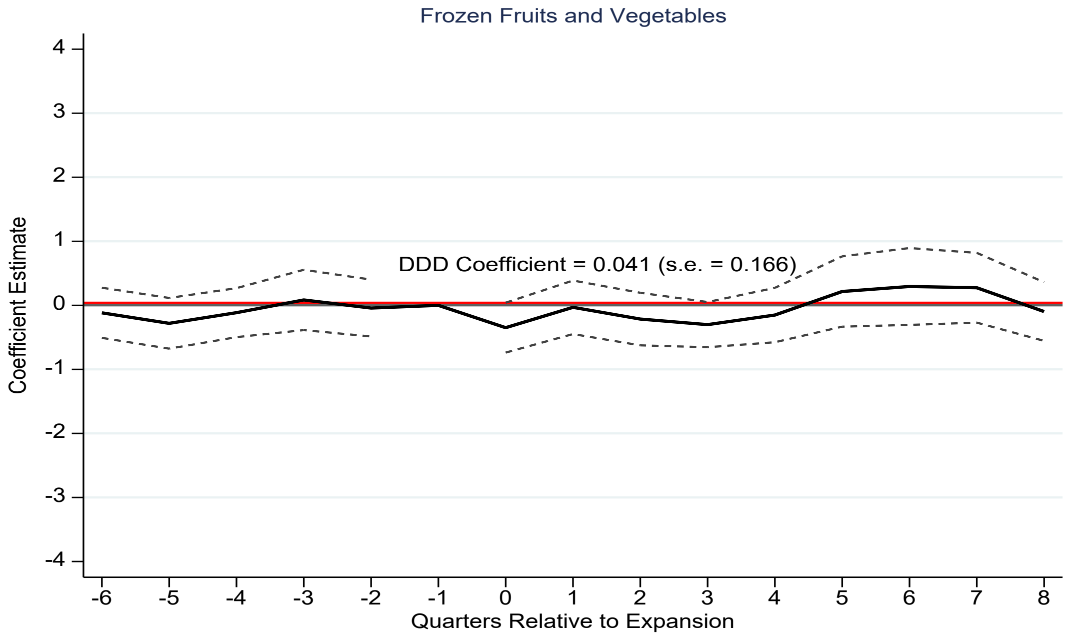


Figure 1.6. Effects of Medicaid expansion on eligible households' quarterly expenditure (2011$) per adult equivalent unit – sub-categories of frozen products – frozen fruits and vegetables, 2011-2017.

Notes: The red line denotes the triple difference model coefficient. The dark thick solid black line denotes the coefficients of interest from event study. The short, dashed lines are the confidence intervals of the event study coefficient estimates. The solid gray line denotes no change line (y-zero). The omitted expansion quarter is *j = -1*.

1.4.4. Effects of Non-Food Expenditures

Among non-food categories, there is a significant positive effect of $4.97 per adult equivalent unit on quarterly expenditure on health and beauty products of eligible households due to expansion (Figure 1.7). This is equivalent to about 10.3% from pre-expansion mean expenditure of the eligible households in expanded states. There are no effects on alcohol, general merchandise, or non-food grocery expenditures.

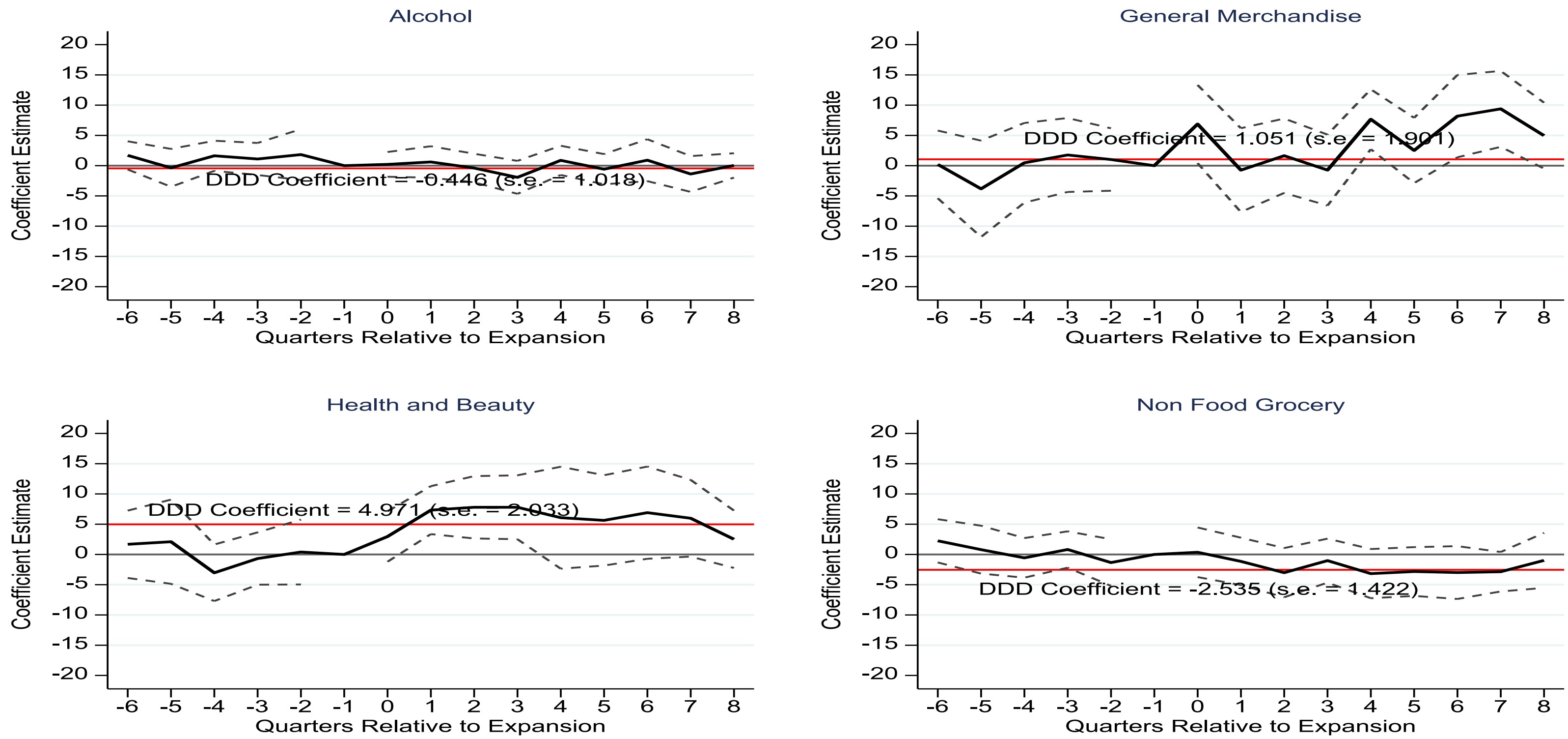


Figure 1.7. Effects of Medicaid expansion on eligible households' quarterly expenditure (2011$) per adult equivalent unit – non-food categories, 2011-2017.

Notes: The red line denotes the triple difference model coefficient. The dark thick solid black line denotes the coefficients of interest from event study. The short, dashed lines are the confidence intervals of the event study coefficient estimates. The solid gray line denotes no change line (y-zero). The omitted expansion quarter is j = -1.

To further analyze what kind of health and beauty products contributed to the increase, I disaggregate the health and beauty products into (1) health products and (2) beauty products and find no significant contribution from beauty products. More than 90% of the significant increase in health and beauty product expenditure comes from the increase in expenditure on health products (Figure 1.8). I further divide the health products subcategory as (2-i) diet aids and vitamins, (2-ii) hygiene and sanitary protection, and (2-iii) over-the-counter medications/remedies/first aid. The results are presented in Figure 1.8. The increase primarily comes from the over-the-counter medications/remedies/first-aid category, which is only partially preventative since many categories of over-the-counter medications are taken to alleviate symptoms.[20] This category includes products such as cough and cold remedies, contraceptives, pregnancy test kits, antacids, insulin syringes, blood pressure kits, pain remedies, etc. This suggests that this low-income sample (i.e., eligible households living in expansion states) made more frequent use of over-the-counter palliative medication, for example Ibuprofen for pain or cold relief following Medicaid expansion. While unlikely to influence mortality, this type of palliative medication substantially improves quality of life when moderately sick or injured.

Though not testable with the data I have, this result is consistent with a pattern in which increased doctor visits and access to preventive care due to Medicaid expansion (Simon, Soni and Cawley 2017), as well as more exposure to health information in general, result in spillovers to these over-the-counter product expenditures. This pattern is also consistent with state level results. Using the Oregon Medicaid lottery experiment, Baicker et al. (2017) found that assignment to Medicaid coverage increased the use of

[20] Preventative care includes certain tests and screenings, vaccinations, contraception, annual checkups.

nonprescription, over-the-counter medications for gastrointestinal conditions such as ulcers. Increased doctor visits allowed for the new diagnoses of gastrointestinal conditions and effective treatment with over-the-counter medications. The category diet aids and vitamins, which is putatively preventative, did not see any significant changes. Hygiene and sanitary protection expenditures were also unaffected.

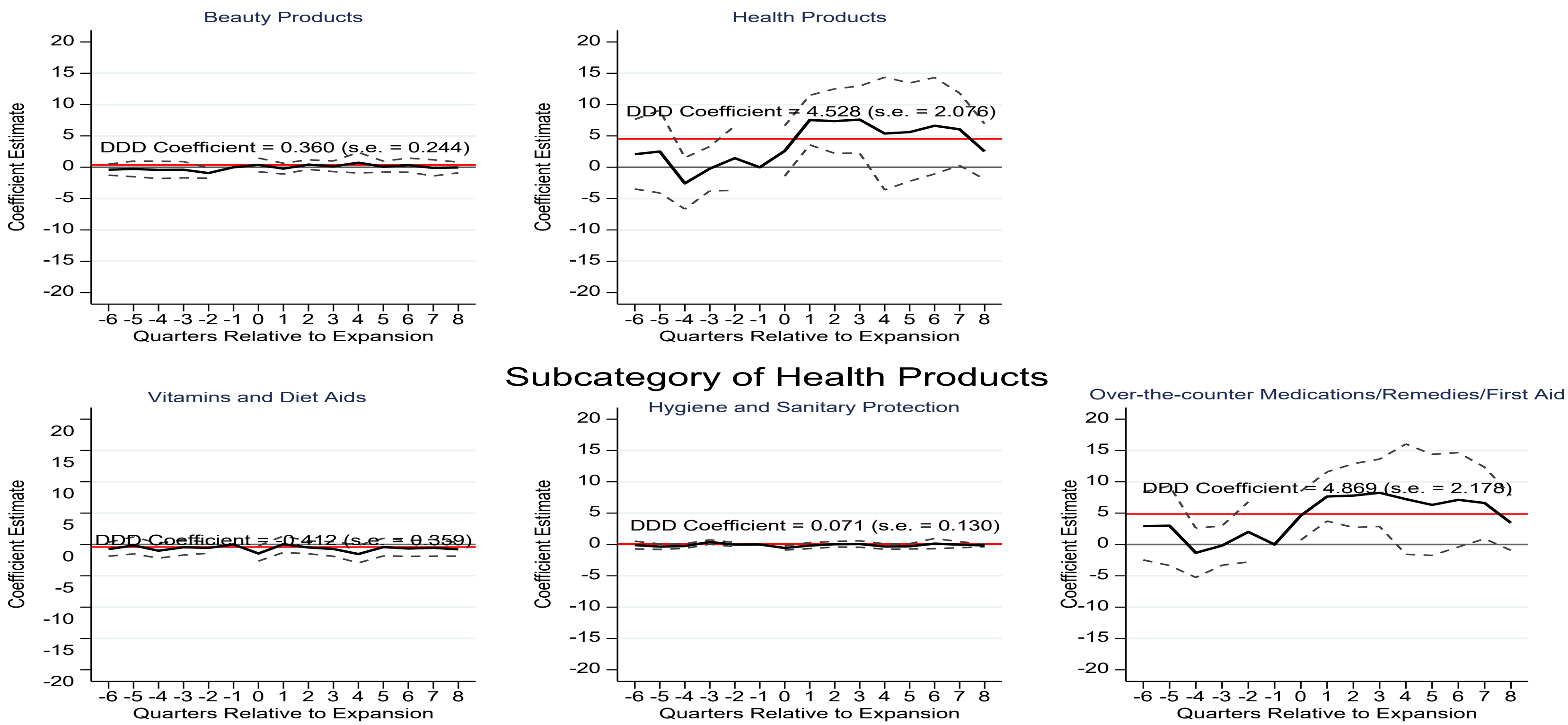


Figure 1.8. Effects of Medicaid expansion on eligible households' quarterly expenditure (2011$) per adult equivalent unit – subcategories of health and beauty products, 2011-2017.

Notes: The red line denotes the triple difference model coefficient. The dark thick solid black line denotes the coefficients of interest from event study. The short, dashed lines are the confidence intervals of the event study coefficient estimates. The solid gray line denotes no change line (y-zero). The omitted expansion quarter is *j = -1*. The bottom three graphs are the subcategories of health products.

1.4.5. Effects on Groups Most Impacted by the Expansion

Low-income childless adults benefitted most from the Medicaid expansion because this group was previously largely ineligible for any public health insurance. I estimate the main models using only households without children under the age of 18 (i.e., childless adults). In the model with only childless households, there is a greater reduction in fresh produce expenditure ($3.73), an insignificant reduction in frozen food expenditure, and a greater increase in health and beauty products expenditure ($7.42) as expected. This suggests more substitution behavior between eating healthy and using more medication among the targeted households.

It is reasonable to assume that another subgroup that benefitted significantly is older people who are not yet eligible for Medicare. So, I conduct the main analysis for a subgroup of households with both heads between ages 55 and 64. The increase in the health and beauty product expenditure is even higher ($11.00) among this group, as expected, even though the decrease in the fresh produce expenditure is smaller ($2) compared to the main sample ($2.88). The results from both groups are presented in tables 1.3, 1.4, and 1.5.

1.4.6. Sensitivity Analysis

I conducted a series of robustness checks by re-estimating the main models presented in tables 1.3, 1.4, and 1.5. First, I estimate the model including the states that had significant prior expansions to test if the results are robust to including or excluding states that have significant prior expansions. Second, I estimate the model without state-specific time trends to verify that the results are not sensitive to state-time trends. Third, I include region-year-quarter fixed effects to test if the results are sensitive to the geographic clustering of the Medicaid expansion because many Southern states did not implement the ACA Medicaid expansion. Fourth, I narrow the age-eligibility by including the households where both heads are between 30 and 60 years old to test if the results are sensitive at the margin of age eligibility requirements. Fifth, I drop the year 2011 since some categories have sharp jumps around the first quarter of 2012. In some cases, researchers use low education to identify eligibility for Medicaid instead of income cutoffs as incomes may respond to changes in Medicaid coverage expansions. As a sensitivity analysis I use both income and education to define eligibility. Households with heads who have less than a high school degree are considered as having low education, and I use this as an alternative proxy for eligibility for Medicaid. I also exclude the late expanders (i.e., states that expanded Medicaid later than 2015) and re-estimate the main models. Next, I estimate a model with all the excluded groups in the main model (excluded states and age groups); this acts as a falsification check as I expect this sub-sample to have no effect. Finally, I exclude the households with incomes above 400 percent of FPL. This narrows the control group to households with incomes between 138 and 400 percentages of FPL.

The results from the sensitivity analysis are similar to the main results with some exceptions (tables 1.3, 1.4, and 1.5). Comparisons made are relative to the main models. When states with significant prior expansions are included, there is an economically insignificant but statistically significant reduction in expenditure on packaged meat (in addition to the significant effects on fresh produce and frozen foods) and the effect on health and beauty products is no longer significant. Results similar to the main results are obtained when the model includes households with income between 100 and 138 percentage of federal poverty level. Results are robust to the exclusion of state specific time trends and to the inclusion of region-year-quarter fixed effects. The effects on health and beauty products are not significant when age eligibility is narrowed. When the year 2011 is dropped, both general merchandise and health and beauty product expenditures significantly increase resulting in an increase in the total non-food expenditure. Results are robust to using both income and education for identifying Medicaid eligibility.[21] Similar results are obtained when late expanders are excluded. Estimations with excluded groups did not yield significant estimates. When a narrowly defined income group is used as control, the effects on health and beauty products are still significant but higher in magnitude while the effects on fresh produce are still significant but lower compared to the main results.

[21] When only education is used to define eligibility, none of the earlier effects are significant even though the health and beauty product expenditure saw greater increase. This might be due to a lack of statistical power and education may not be a good predictor of eligibility in this data.

Table 1.3. Sensitivity Analysis – Total Expenditures

	(1) **All**	(2) **Food**	(3) **Non-food**
1. Include states with significant prior expansions (N=1,079,983)	-5.964 (8.016)	-7.107 (5.400)	1.514 (3.470)
2. No state-specific time trends (N=990,914)	-3.485 (8.820)	-6.548 (6.039)	3.435 (3.657)
3. Add region-year-quarter fixed effects (N=990,914)	-3.976 (8.936)	-6.683 (6.093)	3.100 (3.724)
4. Both heads are between 30 and 60 years old (N=789,659)	-6.683 (11.517)	-9.387 (7.510)	2.371 (5.060)
5. Drop year 2011 (N=843,805)	0.138 (9.298)	-7.540 (5.986)	8.065** (3.956)
6. Education in addition to income cutoffs for eligibility (N=990,914)	-1.832 (8.293)	-5.570 (5.692)	4.452 (3.638)
7. Childless adults (N=679,753)	-2.921 (9.635)	-6.436 (6.881)	4.328 (4.447)
8. Older heads (both heads between ages 55 and 64) (N=331,463)	7.769 (10.911)	-0.454 (7.050)	9.994 (6.484)
9. Exclude late expanders (states expanded after 2015) (N=974,280)	-4.314 (9.054)	-6.473 (6.185)	2.637 (3.712)
10. Excluded states and excluded age group (N=48,717)	-32.596 (25.175)	-28.625 (18.794)	-1.714 (7.051)
11. Upper income threshold is limited at 400% of FPL (N=544,460)	5.864 (9.044)	-1.658 (5.978)	7.275* (3.867)
Estimates from the base model	-3.553 (8.805)	-6.594 (6.034)	3.487 (3.649)

Notes: This table shows the robustness tests from triple difference models for the aggregated total expenditures, food, and non-food expenditures as the outcome variables. The rows show what each robustness test includes, and the columns show the outcome categories. The coefficients shown are the intent-to treat effects of Medicaid expansion on the expenditures. The last highlighted row shows the estimates of the base model. The tests are outlined in the text above. *, **, *** indicate significance at the 10%, 5%, 1% level, respectively. All models include controls for age, marital status, years of education of the household head, number of household heads, hours of employment, and presence of children in the household. Additionally, all models include household, year and quarter fixed-effects, and state-specific time trends. Robust standard errors clustered at the state-level are in parentheses. These estimates are from the triple-difference model and the units are in inflation adjusted dollar expenditures (2011$).

Table 1.4. Sensitivity Analysis – Food Expenditures

	(1) Dairy	(2) Deli	(3) Fresh Produce	(4) Packaged Meat	(5) Frozen Food	(6) Dry Grocery
1. Include states with significant prior expansions (N=1,079,983)	-0.024	-1.133	-3.045***	-0.735*	-1.894*	-0.276
	(0.715)	(0.960)	(0.403)	(0.368)	(1.070)	(3.135)
2. No state-specific time trends (N=990,914)	0.212	-0.959	-2.858***	-0.581	-2.112*	-0.251
	(0.764)	(1.096)	(0.432)	(0.397)	(1.096)	(3.513)
3. Add region-year-quarter fixed effects (N=990,914)	0.231	-0.968	-2.821***	-0.564	-2.159*	-0.403
	(0.762)	(1.095)	(0.444)	(0.386)	(1.124)	(3.492)
4. Both heads are between 30 and 60 years old (N=789,659)	0.119	-0.820	-3.168***	-0.949*	-3.088**	-1.482
	(0.929)	(1.185)	(0.524)	(0.549)	(1.373)	(4.324)
5. Drop year 2011 (N=843,805)	0.169	-1.053	-2.146***	-0.514	-2.212**	-1.784
	(0.750)	(1.130)	(0.467)	(0.403)	(1.088)	(3.364)
6. Education in addition to income cutoffs for eligibility (N=990,914)	0.124	-0.322	-2.646***	-0.629	-2.003*	-0.094
	(0.689)	(1.162)	(0.427)	(0.377)	(1.076)	(3.314)
7. Childless adults (N=679,753)	0.081	-1.047	-3.728***	-0.283	-1.847	0.388
	(0.860)	(1.461)	(0.585)	(0.410)	(1.213)	(4.210)
8. Older heads (both heads between ages 55 and 64) (N=331,463)	0.075	-1.685	-2.011***	0.020	0.142	3.005
	(1.029)	(1.770)	(0.676)	(0.372)	(1.440)	(4.264)
9. Exclude late expanders (states expanded after 2015) (N=974,280)	0.300	-0.998	-2.927***	-0.558	-2.249*	-0.042
	(0.778)	(1.101)	(0.445)	(0.404)	(1.120)	(3.584)

Continued

Table 1.4 continued

	(1) Dairy	(2) Deli	(3) Fresh Produce	(4) Packaged Meat	(5) Frozen Food	(6) Dry Grocery
10. Excluded states and excluded age group (N=48,717)	0.489	-2.362	-1.995	-3.163	-9.175	-12.419
	(0.917)	(2.674)	(1.274)	(1.822)	(7.446)	(7.804)
11. Upper income threshold is limited at 400% of FPL (N=544,460)	0.806	-0.414	-1.782***	-0.276	-0.888	0.896
	(0.715)	(1.135)	(0.425)	(0.387)	(1.158)	(3.612)
Estimates from the base model	0.225	-0.929	-2.875***	-0.601	-2.126*	-0.289
	(0.761)	(1.092)	(0.430)	(0.397)	(1.096)	(3.513)

Notes: This table shows the robustness tests from triple difference models for the categories of food expenditures as the outcome variables. The rows show what each robustness test includes, and the columns show the outcome categories. The coefficients shown are the intent-to treat effects of Medicaid expansion on the expenditures. The last highlighted row shows the estimates of the base model. The tests are outlined in the text above.

*, **, *** indicate significance at the 10%, 5%, 1% level, respectively. All models include controls for age, marital status, years of education of the household head, number of household heads, hours of employment, and presence of children in the household. Additionally, all models include household, year and quarter fixed-effects, and state-specific time trends. Robust standard errors clustered at the state-level are in parentheses. These estimates are from the triple-difference model and the units are in inflation adjusted dollar expenditures (2011$).

Table 1.5. Sensitivity Analysis – Non-Food Expenditures

	(1) Alcohol	**(2) General Merchandise**	**(3) Health and Beauty**	**(4) Non Food Grocery**
1. Include states with significant prior expansions (N=1,079,983)	-0.372	0.833	3.419	-2.738**
	(0.897)	(1.708)	(2.117)	(1.285)
2. No state-specific time trends (N=990,914)	-0.371	1.039	4.919**	-2.523*
	(1.015)	(1.906)	(2.023)	(1.423)
3. Add region-year-quarter fixed effects (N=990,914)	-0.393	0.891	4.773**	-2.564*
	(1.008)	(1.957)	(2.060)	(1.419)
4. Both heads are between 30 and 60 years old (N=789,659)	0.334	-0.242	4.435	-1.822
	(1.121)	(1.997)	(3.041)	(1.542)
5. Drop year 2011 (N=843,805)	-0.386	4.818**	5.290**	-2.043
	(1.067)	(2.100)	(1.981)	(1.363)
6. Education in addition to income cutoffs for eligibility (N=990,914)	-0.714	1.964	4.642**	-2.153
	(0.952)	(1.745)	(1.989)	(1.481)
7. Childless adults (N=679,753)	-0.812	0.546	7.420**	-3.638*
	(1.427)	(2.281)	(2.954)	(1.836)
8. Older heads (both heads between ages 55 and 64) (N=331,463)	-1.771	3.221	10.996*	-4.223
	(2.395)	(3.499)	(5.786)	(2.594)
9. Exclude late expanders (states expanded after 2015) (N=974,280)	-0.477	0.890	4.125**	-2.378
	(1.038)	(1.942)	(1.972)	(1.444)

Continued

Table 1.5. continued

	(1) Alcohol	(2) General Merchandise	(3) Health and Beauty	(4) Non Food Grocery
10. Excluded states and age groups (N=48,717)	-2.257	1.160	0.706	-3.580
	(2.038)	(4.045)	(7.397)	(2.682)
11. Upper income threshold is limited at 400% of FPL (N=544,460)	0.247	2.953	5.844**	-1.522
	(1.034)	(1.991)	(2.222)	(1.312)
Estimates from the base model	-0.446	1.051	4.971**	-2.535*
	(1.018)	(1.901)	(2.033)	(1.422)

Notes: This table shows the robustness tests from triple difference models for the categories of non-food expenditures as the outcome variables. The rows show what each robustness test includes, and the columns show the outcome categories. The coefficients shown are the intent-to treat effects of Medicaid expansion on the expenditures. The last highlighted row shows the estimates of the base model. The tests are outlined in the text above.

*, **, *** indicate significance at the 10%, 5%, 1% level, respectively. All models include controls for age, marital status, years of education of the household head, number of household heads, hours of employment, and presence of children in the household. Additionally, all models include household, year and quarter fixed-effects, and state-specific time trends. Robust standard errors clustered at the state-level are in parentheses. These estimates are from the triple-difference model and the units are in inflation adjusted dollar expenditures (2011$).

1.4.7. Falsification Test

As a falsification test, I estimate the same models for only the households with heads who are below 26 years in age or above 65 years in age. The expectation is that there are no effects because these households' eligibility for health insurance would not change before and after the Medicaid expansion as they are either eligible for Medicare or other provisions. The triple-difference model estimates are presented in Table 1.6. None of the effects except the dry grocery category is significant, as expected. The dry category also had a significant pre-trend so this coefficient might not be well-identified.[22]

[22] I also tested and confirmed that the household income, employment, and work hours are not responsive to Medicaid expansion using an event study design.

Table 1.6. Falsification Test: Triple Difference Model Coefficients for Households Unaffected by Medicaid Expansion (N = 319,160)

Total Spending						
	All	Food	Non-food			
Medicaid Eligible*Expansion	2.778	0.644	1.053			
	(11.738)	(6.471)	(6.999)			
Food Category						
Variable	Dairy	Deli	Fresh Produce	Packaged Meat	Frozen Food	Dry Grocery
Medicaid Eligible*Expansion	-0.614	-2.869	-1.410	-0.609	-2.117	8.263**
	(0.930)	(1.741)	(1.112)	(0.665)	(1.825)	(3.352)
Non Food Category						
	Alcohol	General Merchandise		Health and Beauty		Non Food Grocery
Medicaid Eligible*Expansion	1.081	-1.145		-0.012		2.210
	(0.939)	(3.132)		(4.009)		(1.902)

Notes: This tables shows the estimates of the falsification tests. The estimations are intent-to-treat effects and are expected to be insignificant. The models include only the households with heads less than 26 years old or greater than 65 years old, drop households with income between 100 and 138% of FPL and drop DC and states with prior expansions which are, Delaware, Massachusetts, New York, and Vermont.
*, **, *** indicate significance at the 10%, 5%, 1% level, respectively. All models include controls for age, marital status, years of education of the household head, number of household heads, hours of employment, and presence of children in households. Additionally, all models include household, year and quarter fixed-effects, and state-specific time trends. Robust standard errors clustered at the state-level are in parentheses.

1.4.8. Limitations

The information about household participation in other safety net programs that affect household food consumption such as Supplemental Nutrition Assistance Program (SNAP) is not available hence not controlled for in the models. Research shows that Medicaid enrollment increases SNAP participation and vice versa (Baicker et al. 2014; Yelowitz 1996; Schmidt, Shore-Sheppard and Watson 2019). If a household that newly enrolled in Medicaid also started participating in SNAP, then the mechanism for food consumption changes is via Medicaid-induced SNAP participation and not a direct effect of Medicaid. Data show that more than two-thirds of SNAP participants are households with children and a third are households with elderly or disabled people (Center on Budget and Policy Priorities 2019). The robust results in the childless adult sample suggest that the effects are mainly due to the expansion in Medicaid coverage because there is a higher probability that childless households are not SNAP recipients.

It is not clear whether the seasonality of fresh produce prices affects the purchase behavior of fresh produce. I have controlled for seasonality by adding quarter fixed effects, but this may be insufficient. Due to these limitations, the mechanisms through which fresh produce expenditure is reduced are not fully unraveled. However, given the robustness of the results a plausible explanation is the substitution of health care consumption for preventative non-health care consumption (i.e., less fresh produce consumption).

The results on alcohol purchases are similar to some previous studies (Brook et al. 1983; Cotti, Nesson and Tefft 2019; De Preux 2011), which found that insurance coverage had no effect on the probability of purchasing alcohol, and which suggests that the expansion did not create an ex-ante moral hazard in alcohol consumption. There could be an increase in alcohol consumption away from home which is not captured in this data.[23] Further the eligibility instead of actual participation in Medicaid is used, and so the estimates are likely to be the lower bound of the true effects.

1.5. Conclusions

In this study, I investigate the effects of recent Medicaid expansion on the eligible households' quarterly food and non-food expenditures using state and time variation in Medicaid expansion. Eligible households from expansion states spent less on quarterly fresh produce and frozen foods per adult and more on health and beauty products after Medicaid expansion. Almost all the increase in the health and beauty product expenditure is due to an increase in expenditure on over-the-counter medications and remedies, which are more responsive and palliative in nature.

The robust reduction in fresh produce expenditures and increase in expenditures on over-the-counter medications and remedies suggests that while expanded public health insurance increases formal health care activity, it also decreases informal preventative non-health care expenditures. This does not mean that public health insurance coverage should be limited. The evidence on the benefits of public insurance coverage expansions on health and financial outcomes of the beneficiaries is overwhelmingly positive. Further, pricing incentives such as subsidies aimed at promoting fruits and vegetables purchases for SNAP recipients have been shown to be effective in encouraging fruit and vegetable consumption

[23] Households were instructed to scan purchases from liquor stores.

(An 2013; Durward et al. 2019; Rummo et al. 2019). Among SNAP beneficiaries, financial incentives combined with nutrition education are proven to be effective in improving dietary intake relative to single programs (Verghese, Raber and Sharma 2019). Policy makers can combine health insurance programs with education programs on healthy behaviors and price incentives to help people find a better balance between preventative non-health care consumption and health care consumption.

2. Chapter: Food Choice Behavior of Adolescents under Parent-Student Interaction in the Context of US School Lunch Programs

2.1. Introduction

Participation in the national school lunch program in the United States has been declining on both the extensive and intensive margins. The total number of lunches served reached a 13-year low in 2018 after seven years in a row of decline in the participation rate and eight years in a row decline in total student participants while the number of meals served per student declined by more than 2 percent between 2002 and 2018 (U.S. Department of Agriculture, Food and Nutrition Service 2018). Adding local foods to federally subsidized K-12 school lunches is hypothesized to attract more students to school meals instead of alternative food sources such as food brought from home (National Farm to School Network 2017), which often are less nutritious (Hur, Burgess-Champoux and Reicks 2011). However, adding local foods to school meals creates additional logistical (O'Hara and Benson 2019) and budgetary burdens (Watson, Treadwell and Bucklin 2018). Therefore, understanding student preferences is important to decision makers weighing whether the benefits of increased student participation via inclusion of locally sourced foods in school meals outweigh the additional costs. Deciding whether a student will eat school lunch on any particular day may involve input from both parent and student; hence, understanding how parent and student interact during such decisions is important. In this article, we investigate how the interaction of parent and student preferences for locally sourced items influences school lunch decisions.

This work builds upon and adds to three strands of literature. The first concerns preferences for local foods. Much of this literature concludes that consumers are willing to pay a significant premium for local variants of foods (Adams and Adams 2008; Bean and Sharp 2011; Carpio and Isengildina-Massa 2009; Connolly and Klaiber 2014; Costanigro et al. 2014; Darby et al. 2008; Giraud, Bond and Bond 2005; Hu et al. 2013; Meas et al. 2015; Pelletier et al. 2013; Thilmany, Bond and Bond 2008), but less is known about how inclusion of local foods as part of a meal affects demand.

The second strand of literature includes evaluation of the U.S. Department of Agriculture's Farm-to-School (FTS) programs. FTS was formally established by the Healthy, Hunger-Free Kids Act of 2010 to improve access to local foods in schools. Previous work investigates the benefits of FTS programs (Becot et al. 2017; Joshi, Azuma and Feenstra 2008; Schwartz et al. 2015); examines how local food purchases are affected by local agricultural conditions, e.g., local milk production and direct-to-consumer sales in O'Hara and Benson (2019), total adjacent agricultural production (Botkins and Roe 2018), and the definition of local (Plakias, Klaiber and Roe 2020); explores the impact of local food expenditures on school food service revenues and earnings (Motta 2019); and determines the relationship between the per student local food expenditures and the local food supply chain structure (Christensen et al. 2017). A subset of studies measure student food preferences qualitatively using preference questionnaires and are inconclusive on the relationship between FTS activities and preferences for fruits and vegetables (Prescott et al. 2019).

The third strand of literature involves joint stated-preference elicitation. This literature shows why understanding the preferences of a group, a couple or a household as a whole, matters instead of just the preferences of an individual and mostly explores the preferences of cohabiting couples in environmental, transportation, and public goods settings (Bateman and Munro 2009; Beck and Hess 2016; Beharry-Borg, Hensher and Scarpa 2009; Mariel, Scarpa and Vega-Bayo 2018; Rao and Steckel 1991; Rungie, Scarpa and Thiene 2014; Scarpa, Thiene and Hensher 2012). A few studies focus on household level choices (Dellaert, Prodigalidad and Louviere 1998; Marcucci et al. 2011; Zhang et al. 2009). Papoutsi et al. (2015) use parent-child pairs in a discrete choice experiment to examine how food fiscal policies and child pestering influence parental choice of food for their child. However, only parents made the decisions while their children were allowed to sit next to them during the experiment. A few studies analyze parent-student preference differences in non-recurring choices such as school choice (Giustinelli 2016; Huntington-Klein 2018) or electronics purchases (Aribarg, Arora and Bodur 2002; Aribarg, Arora and Kang 2010).

The purpose of our study is to examine joint and separate food choice behavior by parents and students concerning decisions about school meals with varying local food content. Our specific objectives are to investigate: 1) adolescent preferences for locally sourced elements in school meals and willingness-to-pay; 2) how parent-student influence for locally sourced food varies across socio-economic dimensions; and 3) the situations where assessing the preferences of one party is sufficient for predicting choice as opposed to assessing both parent and student preferences.

We make three contributions to the literature. First, we study a recurring choice situation by modeling both joint and separate local food preferences of parent and students using a school lunch choice experiment which is novel in the literature of stated preference elicitation. We use the results from the choice model to further understand how parent-student influence differs across attributes and with household socio-economic characteristics. Secondly, we contribute to the discrete choice experiment literature by developing novel scenarios that create relevant choices for families of all economic strata by embedding lunch credits within choice scenarios to make the price relevant to students who receive free or reduced-price lunches. We also minimize irrelevant choices by eliciting students' favorite meal items and then populating the choice scenarios with only favorite items to reduce the noise concerning the value of local foods. Finally, we extend the existing literature on FTS programs as well as the literature on preferences for local food.

Modeling both parent and student choices is important for several reasons. For many families, participation in federal school meal programs is a decision that involves input from both parent and student. Pham and Roe (2013) surveyed parents in two suburban school districts and document that more than two-thirds of parents had at least some input into which days of the week their student participated in the National School Lunch Program. Botkins (2017) documents that parents and students often disagree on whether particular school meals are desirable, with only a 62% agreement rate within 90 parent-student dyads who evaluated whether students would prefer a described school lunch over non-school meal alternatives. Further, in the survey we conducted for this study, 39% of the parents said that they sent cash with their student to pay for lunch which suggests regular interaction between parent and student concerning participation. Secondly, the food purchasing behaviors and eating habits of children are influenced by the food environment created by parents (Barlow and Dietz 1998). Parents exert considerable influence on their children's obesity rates especially by communicating healthy behavior, exercising control over children's eating patterns, and providing feedback on children's health choices (Andrews, Silk and Eneli 2010). On the other hand, child pestering strongly affects food purchasing behavior of parents, mainly when parents shop for food with children (Nicholls and Cullen 2004). Parents also choose food products with expected child preferences in mind (Søndergaard and Edelenbos 2007), often choosing less nutritious alternatives due to such pestering (Papoutsi et al. 2015). Finally, by identifying situations where assessing the preferences of only one party is sufficient, we can inform policy makers on whom to target for information campaigns.

We conduct a nationwide U.S. survey that embeds a discrete choice experiment. Our sample includes 1,201 parents with children enrolled in schools with midday meal programs. In a choice situation, respondents were shown three school lunch options and an opt-out option. The lunch options included an entrée, a fruit, and a vegetable that differed in whether none, one, two or all three elements are locally sourced. Parent-student dyads faced choice exercises first *individually* and then *jointly*. Results show that joint choices are influenced by both parties, with the relative importance of each party varying by the meal element. The joint preferences for locally sourced vegetables are more likely to be dominated by parents' preferences while joint preferences for locally sourced fruits are more likely to be dominated by student preferences. Parents' influence is higher in households with lower household income, and in dyads featuring a female parent and female student compared to male parent-male student dyads. Parents' influence is also higher if students eat school lunch more frequently. While our findings are consistent with the limited evidence suggesting that adult and young consumers place higher value on local food, they accentuate why analyzing joint parent-student food choice behavior, rather than individual choices, is vital to understanding decision making in this area.

2.2. Methodology

In the economics literature, there are studies on parent-child interactions, and family decision making although studies regarding the parent-child interactions in food choice behavior is scarce. Following the seminal work of Gary Becker (1974), parent-child interactions were modeled assuming an altruistic parent and purely rational, selfish children. Most of these studies model and analyze family decisions that have a bearing on human capital investments (education and health) and wealth accumulation. Later, non-cooperative game theory was used to explain interactions between adolescent children and parents (Lundberg, Romich and Tsang 2009; Romich, Lundberg and Tsang 2009). Other studies used a principal-agent framework to model parent-child interaction (Weinberg 2001). Since we study the preferences rather than the decision making process itself and use preference parameters to study the influence of parents and students in the food choice behaviors, this essay is closely related to the economics literature on joint stated preference elicitation.[24]

We use a random parameter logit (RPL) model to represent the repeated choices of each parent, student, and dyad (Hole 2007a; Hole 2013; Train 2009). In our choice experiment survey of parent-student dyads, respondents are asked to choose among alternative school lunch options with varying locally sourced content and prices.

The utility of individual i who chooses j lunch alternative at choice occasion t is,

$$U_{ijt} = \beta_i{}' x_{ijt} + \varepsilon_{ijt} \qquad \text{(Eq.2.2)}$$

where, x_{ijt} is the vector of attributes related to the j^{th} alternative, β_i' is the vector of individual-specific parameters, and ε_{ijt} is the error term following an iid type I extreme value distribution. Assuming a utility function that is linear in a vector of random parameters β, with a density function $f(\beta|\theta)$, where θ refers to parameters of the density function (i.e., mean and variance), the probability of a sequence of choices is given by

$$S_n = \int \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\frac{\exp\,(x_{ijt}'\beta_i)}{\sum_{j=1}^{J} \exp\,(x_{ijt}'\beta_i)} \right]^{y_{ijt}} f(\beta|\theta) d\beta \qquad \text{(Eq.2.2)}$$

where $y_{ijt} = 1$ if the choice of individual i is j in choice situation t and 0 otherwise (Hole 2007a). Since this expression does not have a closed-form solution, we use a maximum simulated likelihood estimator that maximizes the following simulated log-likelihood function to estimate parameters of θ:

[24] The related literature is discussed under the introduction section.

$$SLL = \sum_{i=1}^{I} ln\left\{\frac{1}{R}\sum_{r=1}^{R}\prod_{t=1}^{T}\prod_{j=1}^{J}\left[\frac{\exp\,(x'_{ijt}\beta_i^{[r]})}{\sum_{j=1}^{J}\exp\,(x'_{ijt}\beta_i^{[r]})}\right]^{y_{ijt}}\right\} \qquad \text{(Eq.2.3)}$$

where $\beta_i^{[r]}$ is the r^{th} draw for individual i from the distribution of β (Hole 2007b). In this article, all models are estimated using the *mixlogit* package in Stata 15 with 1000 Halton draws for simulations (Hole 2007b). Following Hess and Train (2017), we allow all random parameters to have correlated distributions and estimate the full covariance matrix among random parameters. We assume that random parameters follow a multivariate normal distribution with vector mean *m* and variance-covariance matrix $V = L'L$ where L is the lower triangular (Cholesky) matrix. While V captures all sources of correlation including the correlation that arises from scale heterogeneity, we cannot empirically distinguish the sources of heterogeneity. We estimate the average marginal willingness-to-pay (WTP) values and confidence intervals for local content by bootstrapping. Bootstrapping does not impose a symmetric WTP distribution as in other methods, and does not require the coefficients to be joint-normally distributed (Hole 2007a).

2.2.1. Whose Individual Preferences Dominated the Joint Choices?

We estimate individual-specific parameters following the estimation of parent and student RPL models.[25] The mean parameters for joint decisions are estimated from the joint choices. The superscripts *p*, *s*, and *j* indicate parent, student, and joint parameter estimates, respectively. Subscript *i* indicates the dyad, and *k* indicates the random attribute. Adapting

[25] We estimate 500 bootstrap iterations of the mixed logit model and use each set of bootstrapped estimates to generate individual betas. We report the 95% confidence intervals of the individual specific coefficients to see if there is an overlap between the parent versus student dominated decisions.

the classification scheme from Beharry-Borg, Hensher and Scarpa (2009), we compare the individual specific and joint coefficients to categorize them into two groups representing parent and student dominated outcomes, respectively.

If $\hat{\beta}_{ik}^{p}$, $\hat{\beta}_{ik}^{s}$, $\hat{\beta}_{ik}^{j} > 0$ and $\hat{\beta}_{ik}^{j} > \hat{\beta}_{ik}^{p} > \hat{\beta}_{ik}^{s}$, then the joint choice is consistent with the parent's preferences. This implies that the joint preference parameter is stronger than the parent's preference and closer to the parent's preference than the student's preference for attribute *k*. However, the coefficients could be a mix of positive and negative values within the same dyad and/or the joint coefficients could lie between the individual coefficients. In these cases, the absolute distance between the individual and joint preference coefficient will determine whose preferences dominate the joint decision as shown in Table 2.1. Since the parents and students can have overlapping preference distributions, we further compare the 95% confidence intervals of the individual-specific coefficients to further break each category into groups depending on whether or not the confidence intervals overlap.

Table 2.1. Classification of Parent and Student Dominated Outcomes in Joint Decisions

Scenario 1: Within dyad attribute specific coefficients share the same sign and the joint coefficient has the greatest absolute value	
$\lvert\hat{\beta}^{j}_{ik}\rvert > \lvert\hat{\beta}^{p}_{ik}\rvert > \lvert\hat{\beta}^{s}_{ik}\rvert$	Parent’s preference dominates the joint choice
$\lvert\hat{\beta}^{j}_{ik}\rvert < \lvert\hat{\beta}^{p}_{ik}\rvert < \lvert\hat{\beta}^{s}_{ik}\rvert$	Parent’s preference dominates the joint choice
$\lvert\hat{\beta}^{j}_{ik}\rvert > \lvert\hat{\beta}^{s}_{ik}\rvert > \lvert\hat{\beta}^{p}_{ik}\rvert$	Student’s preference dominates the joint choice
$\lvert\hat{\beta}^{j}_{ik}\rvert < \lvert\hat{\beta}^{s}_{ik}\rvert < \lvert\hat{\beta}^{p}_{ik}\rvert$	Student’s preference dominates the joint choice
Scenario 2: Within dyad attribute specific coefficients do NOT share the same sign and/or the absolute value of the joint coefficient is between the individual specific coefficients	
$\lvert\hat{\beta}^{j}_{ik} - \hat{\beta}^{s}_{ik}\rvert < \lvert\hat{\beta}^{j}_{ik} - \hat{\beta}^{p}_{ik}\rvert$	Student’s preference dominates the joint choice
$\lvert\hat{\beta}^{j}_{ik} - \hat{\beta}^{s}_{ik}\rvert > \lvert\hat{\beta}^{j}_{ik} - \hat{\beta}^{p}_{ik}\rvert$	Parent’s preference dominates the joint choice

Note: $\hat{\beta}^{j}_{ik}$, $\hat{\beta}^{p}_{ik}$, and $\hat{\beta}^{s}_{ik}$ are respectively the parent, student, and joint preference parameters for local attributes obtained from RPL models. Subscript *i* indicates the dyad, and *k* indicates the random attribute.

2.2.2. Explaining Within-dyad Influence

Based on the influence of individual preferences within dyads, we generate a dependent variable $y_i = 1$ for dyad i if the total number of within-dyad outcomes indicate that the parent’s preferences dominate the joint decision and $y_i = 0$, otherwise. We estimate a logistic regression of y_i on socioeconomic variables of interest and school lunch participation related characteristics. Aribarg, Arora and Bodur (2002) used a similar covariate analysis approach to explain the preference revision or concession between parents and students in a marketing study. Beharry-Borg, Hensher and Scarpa (2009) use differences in socio-economic covariates within couples to explain differences in estimates of mean taste parameters between members of the couple and found that several variables explain the difference between spouses' preferences for water quality. Because we explore within-dyad influence and our outcome variable in the model relies on the preference parameters generated from the RPL estimates, we estimate parameters for within-dyad influence as a separate model. Preference parameters generated from the model are assumed to be normally distributed and correlated with each other. We model the binary outcome variable created from the estimated parameters as a linear function of covariates using a logistic regression.

School lunch-related variables include indicators whether or not the student is a picky eater, was exposed to local food or activities related to local food at school or eats school lunch at least once a week. Socioeconomic variables of interest include the absolute age difference between the parent and the student,[26] sex mix of the parent and the student, indicators for household income, responding parent's education, whether the student received free or reduced-price meals, and if more than one adult in the household works full or part time.

2.3. Data and Survey Design

We use data obtained from a choice experiment conducted as part of a 2017 national online survey of 1,201 parents with children enrolled in schools with midday meal programs. In the choice experiment each choice alternative (i.e., lunch option) was defined by four attributes: the local content of the entrée, the vegetable, and the fruit, and the price charged

[26] The absolute difference in ages between parent and student proxies for the age at which the adult became a parent, which has been shown to correlate to parenting style (Kendler, Sham and MacLean 1997), which we hypothesize to affect joint decision making.

for the lunch. Table 2.2 lists the attributes and levels. The first three attributes had two levels (locally produced or not) and the price had four levels. For each choice situation, respondents were shown three alternatives. The first two lunch options varied in terms of which of the three attributes were locally produced or not and prices, while the third option always had no local content and was the lowest price (i.e., $2.80).

Table 2.2. Attributes and Levels

Attributes	**Levels**
Entrée	Local or not (2 levels)
Fruit	Local or not (2 levels)
Vegetable	Local or not (2 levels)
Price per lunch	$2.80, $3.20, $3.60 or $4.00 (4 levels)

Note: This table shows the four attributes and levels of a lunch option.

Respondents could choose any of the three options or select "I would not choose any of the options." This choice serves as the reference or base alternative in the choice model and is referred to as the “no-buy” option. Prices were selected based on the prevailing national average school meal prices. A full factorial design of all possible combinations of all attributes and levels would yield 2x2x2x4=32 possible alternative menu profiles. The choice experiment adopted a full factorial design. However, since each dyad was unable to evaluate all 32 permutations, a randomly selected subset of menu profiles was presented to each dyad. A block design with four blocks of four choices was created using the %mktblock macro available in Statistical Analysis System software (Kuhfeld 2003). While each dyad did not see the whole design, every permutation was included across all the blocks.

The parent and student each separately selected the first two choices of the block. At the end of the survey, dyad members came together and jointly selected these first two choice scenarios and jointly made two additional choices for a total of eight decisions for each dyad. Joint decisions were always made last, so individual choices were made without confounding those choices with the influence of preferences by other members. However, this design is unable to isolate the effects of accumulated learning in later choices since two of the choices are the same between individual and joint decisions.

One challenge posed by a school lunch choice experiment is that many students receive free or reduced-price school meals, potentially making prices attached to different meals unimportant or unfamiliar. To overcome this challenge, we developed a choice scenario in which the student has a particular amount of lunch credit available and where any unspent balance generated by choosing less expensive options or by choosing a non-school lunch option can be carried over to future lunches or used to buy healthy snacks from the school cafeteria. Referencing future lunches and healthy snack alternatives creates a more uniform expectation among respondents concerning the opportunity cost of spending less on the lunch options in the choice set. Even though the reimbursement level is unaffected by whether the food is sourced locally, these current federal reimbursement policies could be altered to include changes in reimbursement levels that reflect the WTP of the students. Further, schools often use other non-federal funds to fund school lunch programs and information on the WTP for local items could be essential for directing review of federal funding rules and for guiding resource on whether to subsidize schools' purchases of local foods, and philanthropic efforts to support local food acquisition by districts.

Nationally, the decline in school lunch participation rate is largely from a decline in participation of students who pay full price. Our work informs the potential for ‘recovering’ that declining segment which could be available to cross-subsidize the district Food Service Authorities. In schools that are mainly or exclusively free/reduced, it is still important to understand the lost welfare from not offering local foods; even if the students cannot pay the increased meal prices in those schools, it does not imply that there is no improvement in social welfare from providing local foods so long as the WTP of students participating in free/reduced priced meals is enhanced by the offering of local foods. In particular, if inclusion of local foods would increase the amount of the meal consumed by students having free/reduced priced meals (a topic we do not investigate), it could yield improved overall nutrition for these students and yield important public health gains.

Another challenge posed by school lunch choice experiments is that some students may have strong pre-existing preferences for certain foods (e.g., love school lunch pizza but dislike school lunch hamburgers). Choosing foods for the choice experiment that individuals may never consider purchasing due to pre-existing preferences would render the incremental value of “local” to be meaningless (i.e., if a student hates beets, making them local is unlikely to make them more desirable). To overcome this challenge, earlier in the survey respondents indicated their first and second choices of entree, fruit, and vegetable from a fixed list which is presented in Appendix Table A.5. Each element of the lunch options in the choice experiment was then populated with these favorite elements with the only varying element being whether the element is locally sourced or not. This approach allows us to capture the preference for local without being confounded by the preference for the food.[27] While food types vary between choice sets, each choice set had three variations of the same entree, fruit, and vegetable with the only distinction being which elements were local and the total price.

During the choice experiment respondents were presented with the following scenario and were given the above-mentioned choice situations: *"Now I'd like you to imagine (you have/your student has) a credit of ($ credit) that can be used at your school for school meals and snacks. If (you do/your student does) not spend the entire credit on lunch, the unspent credit can be carried over to later days or spent on healthy snacks stocked by the school, for example: 0.84 oz. Quaker Chewy Granola Bars for $0.50, 1.5 oz. Cheez-it Baked Snack Crackers for $0.50, or similar snacks. (Your/Your student's) school wants to add a new meal to their menu and is considering the three meals in the table below. Note: "Local" items are those whose ingredients are grown within 250 miles of the school. Locally produced foods are often fresher than items grown and processed further away, can help support local farmers, food processors and manufacturers, can reduce the carbon footprint of school lunch by reducing the total distance that food travels before being eaten, can improve learning by stimulating student questions about how food is grown and the origin of the food we eat."* Figure 2.1 provides a sample choice set that parent and student face in a choice occasion.

[27] The current design eliminates the tradeoff students have to make between their favorite and a less favorite local food. However, the limitation is that we will not know if the students would like certain foods, e.g., locally sourced beans (as they are fresh), which they did not like before.

Meal Choice 1

	Meal A	Meal B	Meal C	None
Entree	Local Hamburger	Hamburger	Hamburger	
Vegetable	Steamed broccoli	Local Steamed broccoli	Steamed broccoli	
Fruit	Fresh grapes	Fresh grapes	Fresh grapes	
Price	$4.00	$3.20	$2.80	
Which of these would you [your student] choose?	Meal A ☐	Meal B ☐	Meal C ☐	None ☐

Figure 2.1. A sample choice situation

A key question in choice experiments is whether decisions made in the proposed hypothetical choice experiments can track real purchases (Chang, Lusk and Norwood 2009; Penn and Hu 2018). However, a pilot study showed that the percent of hypothetical school meals purchased in the choice experiment is monotonic in respondents' revealed purchase frequency, and price sensitivity is associated with household income (Pham and Roe, 2013). The pilot study was conducted in 2013 in two suburban Ohio school districts that had not implemented any FTS activities by the time of the pilot study. By analyzing the pilot data researchers found that the percentage of hypothetical meals purchased increased with the number of days per week a student eats school lunch. Additionally, the probability of purchasing a meal decreased when the meal price was increased in 25 cents increments, and this occurred at a decreasing rate among households with higher income.

These two key results showed evidence that choice experiments could provide responses that follow predictions of economic theory in this context.

A total of 3,988 parent-student dyads were recruited to participate in the survey in 2017. Qualtrics was used to program the survey and recruit participants. The completed survey provided 1,201 usable responses for the choice model. In order to be deemed a usable response, the participating parent had to make decisions about the student's school day lunch more often than any other parent or guardian, the student had to be between age 13 to 18 and had to attend a middle school or high school, the student had to attend a public or private/charter school that serves lunch, and the survey must have been completed successfully (including correct answers to questions designed to measure attentiveness).

The recruits who are not included feature 340 who started but did not complete the survey, 202 who completed the survey but did not pass questions designed to ensure respondents were paying attention, 1,147 where either the parent or student did not consent or assent to participate, 601 where the student was either too young or too old for the study (ineligible), and 497 where the student did not attend a U.S. school serving a midday meal (ineligible). This translates to 41.6% completion rate among eligible recruits. The survey also elicited socioeconomic characteristics of the respondents, characteristics of the school attended by the student, student's recent school lunch participation, awareness about local food served in school meals, student's and parent's perceptions and expectations of school lunch, and general food preferences of students. A comparison of key demographic averages of the final sample (N = 1,201) against the groups who dropped out of the final sample for reasons other than not meeting the criteria is provided in Appendix Table A.6.

Descriptive statistics are shown in Appendix Table A.7. More than 50% of the students purchased school lunch five days a week. Taste and the way food looked were most likely to stimulate student lunch participation. About 58% of responding students said serving more locally grown items would make them more likely to participate in school lunches. Students paying full price for lunch reported paying an average of $3.39 per lunch.[28] The average including free and reduced-price meals is $2.40 per lunch. About 25% of sample households reported receiving free school meals, and 10% received reduced-price school meals. Nationally, about 68% (6%) of all school lunches are free (reduced price) (School Nutrition Association 2019).

2.4. Results

The results are based on 38,432 choices made by the 1,201 parent- student dyads and report estimates from the RPL model. We specify the price and no-buy option coefficients as fixed to aid in calculation of willingness to pay, while all the local attribute coefficients are specified as normally distributed with full covariance among random coefficients (Colombo, Hanley and Louviere 2009). This specification allows average marginal WTP estimates to have the same distribution as the coefficient of the attribute. We calculate the average marginal WTP for an attribute as the negative of the ratio between the attribute and price coefficient.

[28] For comparison, the 2017-18 federal reimbursement rate for school lunch was $3.23 (Federal Register 2017) while the average price reported by 1,550 member-schools of the School Nutrition Association was $2.74 for high schools in the 2016-17 school year (School Nutrition Association 2018).

Separate models are estimated to model the food choice behavior of parent-only, student-only, and joint parent-student choices (Table 2.3). The survey contained four different choice sets, denoted A-D. Choice sets A and B were each presented multiple times (parent only, student only, and joint), while choice sets C and D were presented only in a joint decision. Prior to finalizing the three separate models, we also estimated an RPL model by pooling the choice responses of parents and students. The pooled model controls for potential differences in scale and imposes the null hypothesis of identical preference parameters for parent and student. Using the likelihood ratio test between pooled and separate models, we conclude that the preferences are different across parents and students with a chi-square test statistic of 27.08 (df=11, $p<0.01$). Similarly, we tested the null hypothesis of equality of preferences across the choice responses of parents, students, and parent-student dyads by comparing the sum of likelihood estimates from each model to the pooled model. The null is rejected with a chi-square test statistic of 141.80 (df=22, $p<0.01$). Thus, we use separate models for parents and students throughout the chapter as opposed to one composite model and also a separate model for the joint responses.[29]

[29] We also estimate the mixed logit models with full covariance and repeat the likelihood ratio tests and conclude that preferences are different across parents and students with a chi-square test statistic of 139.93 (df=20, $p<0.01$). This suggests that scale heterogeneity between parents and students do not confound the results.

Table 2.3. Design and Models

Order of the choice experiment in the survey	Choice sets used	Who is choosing	Parent-only model	Student-only model	Pooled model	Joint model
1^{st} & 2^{nd}	A & B	Parent only	X		X	
3^{rd} & 4^{th}	A & B	Student only		X	X	
5^{th} & 6^{th}	A & B	Joint decision				X
7^{th} & 8^{th}	C & D	Joint decision				X

Note: This table shows the four RPL models estimated (i.e., parent-only, student-only, pooled, joint) and what each model includes in terms of choice experiment.

Table 2.4 presents the regression results of RPL models with standard errors clustered at the individual level or dyad level in parentheses. Based on the estimated coefficient means, the presence of a locally-sourced element in a school lunch increases the associated utility level of students compared to the same non-locally sourced element. Higher prices of school lunches decrease utility. The coefficients of the no-buy option were negative and highly significant suggesting that the respondents preferred to spend their lunch credit on a meal described in the scenario instead of carrying over the whole credit to later days or spending the whole credit on healthy snacks. The significant standard deviations on the fully correlated local attribute coefficients indicate the presence of substantial heterogeneity in preferences and that some respondents may prefer lower levels of certain attributes. The significant off-diagonal elements of the same Cholesky matrices in Appendix Table A.8 illustrate the presence of significant pair-wise correlations across the three attributes and justifies the specification of correlated coefficients. The correlation matrix and the coefficient covariant matrix are reported in Appendix Tables A.9 and A.10.[30]

The highest increase in utility is associated with locally produced entrées across all choice situations. There are some differences in preferences for other local attributes between parents and students as well as between individual and joint choices. On average parents perceive that students would get higher incremental utility from locally produced vegetables than local fruit. However, the opposite is true with the student sample. Joint parent-student parameter estimates for locally sourced fruits show that they are more extreme than either the estimate from just the parents or just the students.

To better interpret the magnitudes of RPL coefficients, we estimated the average marginal willingness-to-pay (WTP) for each attribute (Table 2.5. Average Marginal Willingness-to-Pay Estimates). Based on student preferences, holding all other attributes constant, a locally produced entrée[31] would add an estimated value of 41¢ on average to a school lunch compared to the same entrée that is not locally produced. Similarly, on average, a locally produced fruit would add a premium of 23¢ to a school lunch, and a locally produced vegetable would add a premium of 18¢ compared to the corresponding non-local elements. The higher WTP for entrées than for the side items is consistent with the fact the entrée is generally the costliest element in a school meal (Matts 2009) and

---

[30] The correlations between the estimated coefficients in parent-only and student-only models have the same signs but differ in magnitude. The correlations from the joint model have changes in both magnitude and signs. Allowing correlations between the random preference parameters slightly decreases mean WTP estimates and increases their confidence intervals.

[31] The survey instructions do not articulate whether all major ingredients within an entrée must be produced locally to qualify as local, which means that different participants may have different interpretations of the term local for entrees.

mostly consumed without plate waste (Cohen et al. 2013). Parents have higher marginal WTP for locally produced entrées (45¢) and vegetables (33¢) than students. The joint marginal WTP estimates are higher than either the individual student's or parent's estimates and this is stable across elements. The estimated marginal WTP based on the pooled parent and student are in between the individual parent and student estimates.

The results show that incorporating locally sourced items in school meals may enhance utility so long as the lunch price does not increase too much. Among the three items, switching to a local entrée is likely to be more popular among the students followed by switching to local fruits and local vegetables. To put things into perspective, an average school meal price was $2.71/meal in middle and high schools in 2016-17 school year based on the State of School Nutrition Survey 2018 (School Nutrition Association 2018). Thus, if we consider the highest and the lowest mean WTP estimates for each element from Table 2.5, keeping everything else constant, respondents would be willing to pay a premium of 15-20% for a school lunch with a locally produced entrée, compared with 7-13% and 6-15% for locally produced fruit and vegetables elements, respectively. However, school food authorities (SFAs) need to consider the additional cost for adding local elements to conclude whether these numbers imply adequate cost savings or revenue generation.

Table 2.4. Model Results from RPL Models with Correlated Coefficients

	(1) Parent-only		(2) Student-only		(3) Joint		(4) Pooled Parent and Student	
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Entrée: locally produced	0.931*** (0.148)	1.719*** (0.334)	0.733*** (0.132)	1.985*** (0.317)	0.735*** (0.072)	1.365*** (0.093)	0.830*** (0.097)	1.829*** (0.225)
Fruit: locally produced	0.412** (0.181)	2.432*** (0.280)	0.396** (0.164)	2.090*** (0.259)	0.475*** (0.072)	0.971*** (0.111)	0.413*** (0.120)	2.265*** (0.192)
Vegetable: locally produced	0.654*** (0.137)	1.474*** (0.259)	0.322*** (0.322)	1.502*** (0.268)	0.550*** (0.061)	0.708*** (0.113)	0.484*** (0.088)	1.489*** (0.187)
Price	-2.030*** (0.229)		-1.781*** (0.205)		-1.319*** (0.096)		-1.900*** (0.153)	
No-buy option	-9.070*** (0.688)		-7.701*** (0.605)		-6.577*** (0.335)		-8.309*** (0.453)	
Log likelihood	-2346.995		-2524.262		-4703.056		-4884.798	

Notes: The table provides estimates (i.e., means and standard deviations for the random parameters), means for the fixed parameters of the RPL models for parent-only (column 1), student-only (column 2), joint (column 3), and pooled parent and student (column 4) sub-samples. Standard errors, clustered at the individual level in models (1), (2), and (3) and at the dyad-level in model (4) are shown in parantheses. Number of respondents/dyads in all models include 1,201 individuals or dyads, and 9608 observations in model (1) and (2), and 19216 observations in models (3) and (4).

Table 2.5. Average Marginal Willingness-to-Pay Estimates

	Parent (¢)	Student (¢)	Joint (¢)	Pooled Parent and Student (¢)
Entrée: locally produced	45***	41***	55***	44***
	(0.067)	(0.081)	(0.049)	(0.048)
	[32,58]	[25, 57]	[46, 65]	[34, 53]
Fruit: locally produced	20**	23***	36***	22***
	(0.091)	(0.088)	(0.057)	(0.062)
	[3, 38]	[6, 40]	[25, 47]	[10, 35]
Vegetable: locally produced	33***	18***	42***	25***
	(0.060)	(0.063)	(0.044)	(0.040)
	[21, 45]	[6, 31]	[33, 51]	[17, 33]

Notes: WTP estimates for the local attributes derived from the bootstrapping estimations in table 2.4. *p < 0.1, **p < 0.05, ***p < 0.01, the standard errors are shown in parantheses, the confidence interval for WTP estimates are shown in brackets.

We conduct pairwise between-meal element comparisons of WTP estimates within each respondent group by bootstrapping (i.e., between rows within each column of Table 2.5). The marginal WTP estimates for local entrée and fruit are significantly different from each other among parents. Among students, the significantly different pair involves the local entrée and the vegetable. Both these significant differences occur in the joint responses. The marginal WTPs between the local vegetable and fruit are not significantly different from each other in any group, hence the relative attractiveness of adding a local entrée versus local fruit/vegetable depends on the decision-making group.

2.4.1. Parent and Student Dominated Outcomes in Joint Decisions

Table 2.6 reports whether the parent or the student preferences dominate the joint decision. The numbers in each cell show what percentage of individual parent or individual student preferences dominate joint decisions. The first (second) inset column for each dyad member features the percentage of joint decisions in which confidence intervals do not (do) overlap with the other dyad member's confidence interval, revealing an indication of the degree of dominance. When it comes to entrées and vegetables, parents' preferences (the shaded columns) appear to dominate as the most frequent outcome was a parent attempting to dominate the student's preference for local entrées and vegetables. The opposite holds for locally sourced fruits with the students' preferences (shaded columns) dominating. The values in Table 2.6 are descriptive, therefore comparisons between any two values need to be made with caution. For example, for entrées parents dominate 20.07% of the joint preferences with no overlapping CI with student preferences and dominate 49.79% of the joint preferences with overlapping CIs.[32]

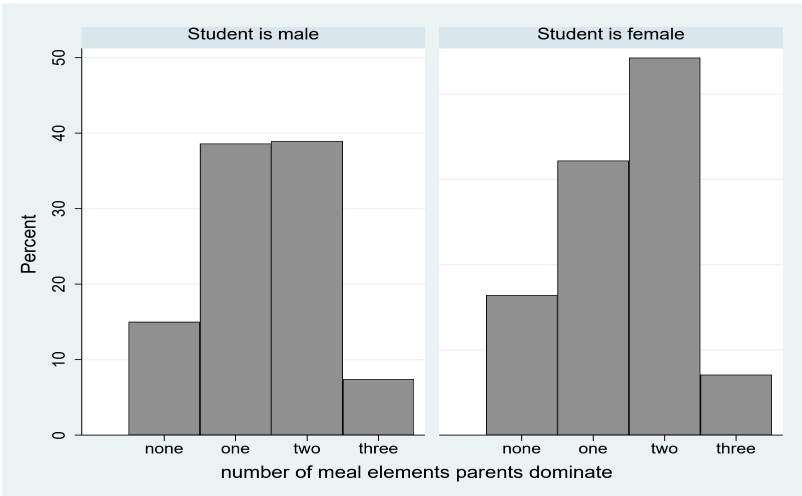
[32] In other words, 71.28% of the parent-dominated joint decisions for entrées have overlapping CI compared to only 54.97% of the student-dominated joint decisions for entrées (i.e., 49.79/(20.07+49.79) versus 16.57/(13.57+16.57)).

Table 2.6. Whose Preferences Dominate the Joint Decision?

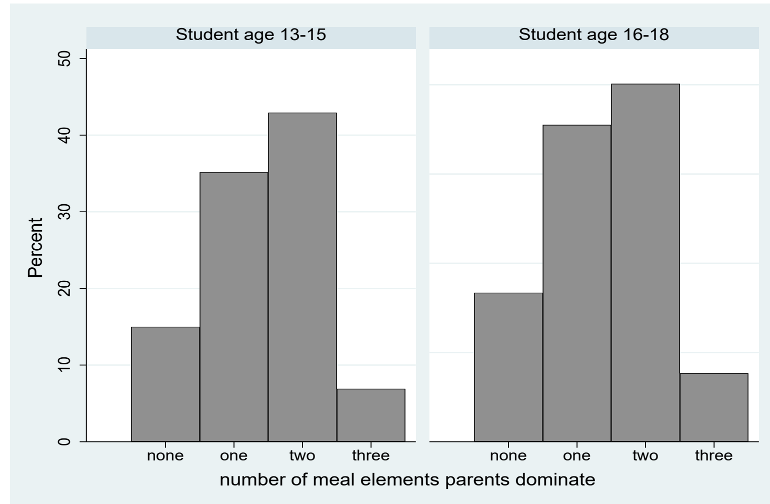
Parameters	Percentage of parent versus student dominated joint decisions with and without overlapping Confidence Intervals (CI's) of individual specific coefficients (%)			
	Parent Dominates		Student Dominates	
	CI's Do not overlap	CI's Do overlap	CI's Do not overlap	CI's Do overlap
Entrée: locally produced	20.07	49.79	13.57	16.57
Fruit: locally produced	16.15	24.90	21.98	36.97
Vegetable: locally produced	17.15	40.47	15.82	26.56

Note: This table reports whether the parent or the student preferences dominate the joint decision. The numbers in each cell show what percentage of individual parent or individual student preferences dominate joint decisions and whether the confidence intervals overlap with the other dyad member's confidence interval. See text for derivation of percentages. Shaded cells denote for each row/item, which party dominates the majority of choices.

Figure 2.2 shows which person's preferences dominate the joint decisions by student sex (left panel) and student age (right panel) using the individual specific coefficients. We plot histograms with the percentage of parent-dominated joint outcomes against the total number of meal elements (entrée, fruit, vegetable) for which parents dominate students to the parent's preferences in the joint decisions. Parents dominate slightly more successfully when the student's sex is female since there is an increase in the number of meal elements that parents dominate successfully when the student's sex is female than when the student's sex is male. However, complete parent domination (i.e., parent-dominated outcomes across all three elements) or complete parent non-domination (no parent-dominated outcomes across the three elements) are not much different between the two student sexes. Based on the right panel in Figure 2.2, there is no specific pattern by student age in terms of parent-dominated outcomes. Interestingly, both age groups (i.e., 13-15 years and 16-18 years) show similar patterns across the number of meal elements. We also split the data to see if patterns emerge across different meal elements by sex or age and we did not see significant differences by sex or age. We group the data to observe within dyad consistency in parent-dominated joint decisions across meal elements. Figure 2.3 presents the results. We do not see any strong evidence for either the parent or the student within the dyad dominating preferences for all the meal elements in the joint decisions. Most dyads fall into the category of one member dominating the preferences for one or two of the three attributes.



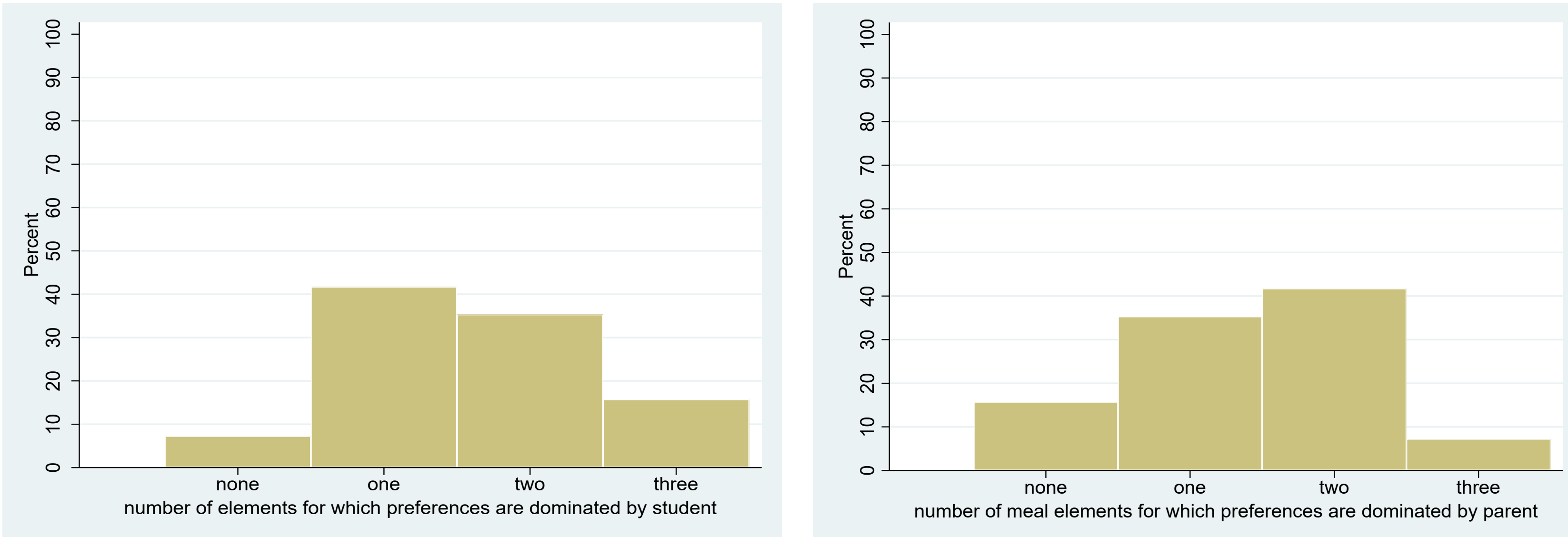
Parent-dominated outcomes by student sex



Parent-dominated outcomes by student age

Figure 2.2. Parent- versus student-dominated outcomes by age and sex

Notes: This figure shows which person's preferences dominate the joint decisions by student sex (left panel) and student age (right panel) using the individual specific coefficients. Each bar shows the percentage of interactions in which parent dominates the preferences for certain number of meal elements. The number of meal elements vary from zero to three.



Within dyad consistency in student-dominated outcomes across elements

Within dyad consistency in parent-dominated outcomes across elements

Figure 2.3. Within dyad consistency in parent- versus student-dominated outcomes across elements
Notes: This figure shows the within dyad consistency in parent-dominated or student-dominated joint decisions across meal elements. Each bar indicates the percentage of times within a dyad in which 1) the student dominates each number of joint preferences (left panel), and 2) the parent dominates each number of joint preferences (right panel).

2.4.2. Influence Regressions

Finally, we regress an indicator variable for higher overall influence by parents on the covariates of interest. The purpose of this analysis is to identify situations in which assessing the preferences of only one party is enough to yield policy relevant insights. The dependent variable is equal to one if the overall preferences across elements are influenced by parents (i.e., based on the classification in Table 2.1) and zero otherwise. This means that when a parent and student make choices jointly, at least two of the three attribute specific coefficients were dominated by the parent's preferences for local sourcing. Results are reported in Table 2.7 where column 2 reports the logistic regression coefficients and column 3 reports the average marginal effects.

Parent influence is lower when the parent-student age difference is greatest, which could reflect past results that parents who were older when children were first born are less likely to have authoritarian parenting styles (Kendler, Sham and MacLean 1997). Parent influence is higher when the student consumes a school lunch at least weekly. This may imply that in situations where the student eats school lunch on a regular basis, parents have more input in lunch-related decision making. On average, a student who consumes school lunch on a weekly basis is 16 percentage points more likely to be influenced by the parent in the majority of the meal attributes than a comparable student who doesn't consume school lunch on a weekly basis. Among the sex mixes of parent and student, female parent-female student dyads are 9 percentage points more likely to have more input from the parent compared to a male parent-male student dyad. The student is 8 percentage points less likely to be influenced by the parent if the yearly household income is greater than

$50,000. We do not see any evidence for significant associations between the following variables and overall influence by the parent via dominance of their individual preferences for different attributes: parent's education, student's age, whether more than one adult in the household works full or part time, whether or not the student is a picky eater, whether or not the student reports being exposed to local foods in school, and whether or not the student received free or reduced-price meals. The small, adjusted R-squared value implies that there is considerable variation not explained by the covariates.

Table 2.7. Regression Estimates of the Within-Dyad Influence

Dependent variable: y =1 (if parent's influence > student's influence | dyad)	Coefficients	Average Marginal Effects
Difference between parent and student age	-0.0237***	-0.0057***
	(0.009)	(0.0022)
Student age	-0.0181	-0.0044
	(0.0451)	(0.0109)
Male parent - Female student	0.212	0.0514
	(0.219)	(0.0531)
Female parent - Male student	0.446	0.1082
	(0.329)	(0.0793)
Female parent - Female student	0.370**	0.0899**
	(0.188)	(0.0453)
Male parent - Male student	(omitted)	(omitted)
Student is a picky eater	0.0789	0.0192
	(0.230)	(0.0558)
Received free or reduced-price meals	-0.253	-0.0613
	(0.172)	(0.0414)
Have school lunch at least once per week	0.700***	0.1698***
	(0.204)	(0.0482)
Student has awareness of local food in school cafeteria	0.0534	0.0130
	(0.191)	(0.0463)
Yearly household income is higher than $50k	-0.341*	-0.0827*
	(0.196)	(0.0471)
Parent has a Bachelor's degree	-0.0431	-0.0105
	(0.168)	(0.0407)
More than one adult work full or part time	-0.0765	-0.0186
	(0.139)	(0.0336)
Constant	0.451	
	(0.858)	
Observations	862	
R-squared	0.0212	

Notes: The table shows the outcomes of the logistic regression of an indicator variable for higher overall influence by parents against the covariates of interest. The column on the far right shows the average marginal effects. *, **, *** denotes statistical significance at the 10, 5 and 1 percent levels. The 862 observations is less than the full sample count of 1,201 due to mis-reporting or missing ages for parents.

2.5. Discussion

Overall, the presence of locally sourced elements compared to non-local elements in school lunches increases the associated utility level of students and parents regardless of whether the decision is made individually or jointly. Locally sourced entrées are the most preferred by both parents and students. Based on WTP estimates, entrée vs. fruits and entrée vs. vegetables show significant between-element differences in preferences for local sourcing. There is no significant between-element difference for fruits vs. vegetables across all the models. This might be an important result to SFAs in deciding which elements to source locally as there does not seem to be significant tradeoffs in terms of marginal WTP between locally sourced fruits and vegetables regardless of whether parents or students made the choices. If we assume that all three locally sourced elements are equally costly, and there are no budgetary or personnel constraints, then locally sourced entrées would generate more surplus followed by vegetables and fruits. Even if school food authorities are supposed to maintain a non-profit status, they would still aim to at least break even. Hence, when we relax the assumptions on costs and constraints and consider whether the schools can generate adequate revenue or achieve cost savings by featuring locally sourced items depends on factors such as seasonality and availability of the products, cost, capacity of the school kitchen, the flexibility if the school has to charge higher prices to full price students, and the percentage of students who pay full lunch price.

The differences between actual student preferences for local food and the parents' perceived student preferences have useful implications for schools and researchers depending on their goals. However, one needs to bear in mind that in this study we force parents and students to jointly decide in part of the choice experiment whereas the frequency of a parent and student jointly make lunch decisions varies across households. Since parents successfully dominate students' preferences for vegetables when they make the choices together with their students, if an SFA specifically aims to serve more locally sourced vegetables, then targeting parents for a promotional campaign would be more sensible. For researchers, the findings imply that when predicting lunch choice by only interviewing one member of the family, at least asking that member about how the school lunch-related decisions are made in their households is important. Ignoring the parent-student interactions in school meal-related decisions may lead to inaccurate targeting of campaigns to increase demand for locally sourced elements in school meals.

Results from the influence regression hint that campaigns targeting parents may work better with female students and female parents and students who regularly participate in school lunches. Students from high income households with older parents are less likely to be influenced by parents, thus targeting such parents to promote local food in those circumstances could be less effective. Given that joint decisions are more favorable towards locally sourced foods, if SFAs have adequate resources, then targeting both parents and students would yield decisions in which meals with locally sourced elements are chosen more often. However, we re-emphasize that the extremeness of the joint estimates could be due to factors such as accumulated learning or fatigue by the participants.

There are several limitations to our study. Firstly, we specify the price and no-buy option coefficients as fixed, and all the local attribute coefficients to be normally distributed with full covariance among random coefficients. Fixed price coefficients make the distribution of derived WTP the same as the distribution of the associated attribute parameter, yielding defined moments (Ortega et al. 2012; Veronesi et al. 2014; Ward et al. 2014). However, the tradeoff is that we assume that all individuals have the same preferences for price. It is reasonable to assume that the local attribute coefficients follow a normal distribution since parents and students can exhibit either positive or negative preferences for locally sourced lunches. Specifying all coefficients as nonrandom is observed to yield problems in model identification (Hensher, Shore and Train 2005; Ruud 1996). Secondly, with the available information, we could not completely explain why joint estimates of preferences for local meal elements are more extreme than the individual estimates. Factors such as palatability, and other food related preferences (e.g., preference for freshness) could matter. It is possible that parents and students could consider entirely different factors when they make the choice for a meal option. What is important to parents could be student's nutrition, and they might be willing to spend some extra money for their students to meet the caloric and nutritional requirements. This is also evident from the higher values that students and parents place on caloric entrées relative to fruits and vegetables. For students, factors such as palatability may matter more, and they might not value the "local" attribute as much as their parents would. However, recall all choice scenarios feature foods that respondents deemed to be among their favorite items and the choice scenarios merely reflect a different number of meal elements with local content. Even if the students, once at school, will not always choose the same options, school district food service managers will benefit from knowing whether or not promoting local items increases either the parent or the student's interest in participating in the NSLP.

The survey was administered online, thus one could question if the respective parties actually completed their sections as prompted. In other words, we must rule out that the parent just answered all the questions to receive the compensation for completing the survey. If one party answered all questions on a regular basis, we would expect no difference between parent and student preferences for attributes. The extremeness of joint preferences and WTP estimates suggest that two distinct parties answered the different parts of the survey individually and then jointly as prompted. We also acknowledge that participants in an online survey may systematically differ from the entirety of households engaged with the National School Lunch Program. Finally, we acknowledge that the data are responses to hypothetical scenarios rather than revealed preference data. Given our objectives of understanding differences between parent and student preferences and the potential role of parent-student interaction during decision making, a hypothetical choice experiment permits the cleanest approach to meeting these objectives and including responses from a national sample.

2.6. Conclusions

Whether questioned alone or jointly, we find students who obtain lunch through the National School Lunch Program and their parents would prefer that locally produced items be added to school lunch menus, and that the willingness to pay for such offerings range from \$0.18 to \$0.55 per meal (6% to 20% of meal price). The highest willingness to pay is associated with the addition of a main meal entrée that is locally sourced, while the addition of a locally sourced fruit or vegetable evokes a smaller willingness to pay.

While parent and student preferences align on some aspects of locally sourced meal elements, their preferences are not identical, and preferences identified from joint parent-student decisions differ significantly from those identified from decisions cast individually.

For example, examining only parent responses would result in more emphasis on locally sourced vegetables while examining only student responses would result in more focus on locally produced fruit. We discuss the usefulness and implications of such preference differences. When parent and student make choices together, the marginal WTP for local content estimates are 25% to 133% greater than either the estimates from just the parents or just the students. We find parents are more likely to be the driver of the additional value placed on locally produced vegetables, with majority having overlapping preferences with students in these joint decisions. In contrast, students are more likely to be the driver in the cases of heightened willingness to pay for fruits in joint decisions. Parents are more likely to influence the joint decision when students report frequently relying upon school lunch compared to students who only occasionally eat a school lunch. We caution that joint choices were always made last in our experimental sequence; therefore, the altered magnitude may be due to the joint decision process as well as due to accumulated individual learning or fatigue by both parent and student respondents.

Local food items are currently included in thousands of schools. Our work suggests that both parents and students value the local sourcing of school meals and local sourcing is not imposed on students by parents; thus, adds value to the motivation behind the Farm-to-School programs. Our findings are consistent with the limited evidence suggesting that adult and young consumers place higher value on local foods. We also provide new evidence as to why analyzing both parent and student food choice behavior, rather than individual choices, is vital in this context. Our findings may hold implications for efforts to promote locally sourced food elements in school lunches and the role of parent engagement in that process.

3. Chapter: Who Buys Portion-Controlled Sizes of Full Calorie Soda? Evidence from Scanner Data

3.1. Introduction

Curbing the rising rates of adult and childhood obesity has been a focus of recent public health policy discussions. The prevalence of obesity in the United States has rapidly increased in the past three decades. Rising obesity rates increase the risk of type 2 diabetes, cardiovascular diseases, hypertension, some types of cancer, and premature deaths in obese individuals, posing a public health threat (Bogers et al. 2007; Dengo et al. 2010; Greenberg 2006; Guh et al. 2009; McGee and Diverse Populations Collaboration 2005). Conventional obesity interventions such as special diets and intense physical exercise have not proven to be successful in the long term for physiological and behavioral reasons; most people who lose weight through dietary and lifestyle changes tend to regain the weight over time (Anderson et al. 2001; Jeffery et al. 2000; Wadden et al. 1989; Wadden, Butryn and Byrne 2004). An alternative approach to induce weight loss aims for smaller behavioral changes. Reducing unit sizes of consumer-packaged foods and requiring restaurants and food outlets to display the number of calories contained in standard items are examples of changes to the prevailing food environment that have been suggested to induce behavioral changes that reduce the intake of unhealthy foods. In this essay, I focus on the household level purchases of portion-controlled sizes of carbonated beverages which I define as full calorie (i.e., regular) carbonated beverages sold in less than 12 oz sized containers. 'Portion' refers to the overall volume of a single unit of carbonated beverage purchased for consumption. Non-conventionally sized packages (less than 12 oz) were launched to attract health conscious consumers who want to restrict calories consumed from full calorie beverages. Specifically, I investigate what characteristics of households, if any, predict purchase of portion-controlled sizes of full calorie carbonated beverages and whether this behavior is associated with other healthy dietary habits.

The link between portion sizes, food intake, and body weight has gained increased focus in the recent obesity literature (Ledikwe, Ello-Martin and Rolls 2005; Steenhuis and Poelman 2017; Rolls et al. 2004; Rolls 2014; Young and Nestle 2002). Systematic evidence from randomized control trials conducted in laboratory or field settings (Diliberti et al. 2004; Hollands et al. 2015; Rolls, Roe and Meengs 2006) shows that increased portion sizes lead to a substantial increase in calories consumed, with the concern that this has contributed to excess weight gain (Ello-Martin, Ledikwe and Rolls 2005; Hieke et al. 2016; Young and Nestle 2002). However, there is a lack of observational studies proving causality between increased portion sizes and weight gain. A meta-analytic review of laboratory based studies concludes that for a doubling of portion size, consumption increases by 35% on average (Zlatevska, Dubelaar and Holden 2014). Research also shows that consumers in general underestimate the calories in food, and thus miscalculate the calories consumed (Burton et al. 2006). The miscalculation of calories is exacerbated when marketplace portion sizes of ready-to-eat prepared foods exceed federal standard serving sizes (Young and Nestle 2003). Making portion-controlled packages that are predominantly single servings available in the marketplace could help consumers to limit portions.

Full calorie carbonated beverages contain large amounts of added sugars and few or no nutrients, so higher consumption of these sugary drinks is associated with weight gain and increased risk of type 2 diabetes (Schulze et al. 2004). Moreover, in the past there was an increase in the size of the carbonated beverage packages sold as single servings. For example, one of the prominent brands, Coca-Cola, was originally sold in 6.5 oz. (192 ml) bottles with larger bottles introduced beginning in 1955. Coca-Cola currently sells single-serve bottles in 16 oz. (500 ml) and 20 oz. (600 ml) and single-serve cans in 12 oz. (Zlatevska et al. 2014). Since the rise of public health concerns, followed by dramatic reduction in carbonated full calorie beverage consumption, Coca-Cola introduced mini-cans (7.5 oz and 90 calories per can) in 2009 as one of the many ways to attract health conscious consumers to buy more regular carbonated beverages (Young and Nestle 2012).

Even though carbonated beverage consumption has declined over the past decade, the average daily consumption of added sugars from carbonated beverages still exceeds recommended limits (Welsh, Lundeen and Stein 2013). Chen et al. (2009) show that a 100 kcal/day reduction in consumption of regular carbonated beverages over a six-month period is associated with a weight loss of 0.25 kg in adults. While taxing based on the sugar content of carbonated beverages yields some positive results in some cities, it does not provide the expected results in all contexts (Fletcher, Frisvold and Tefft 2010a; Fletcher, Frisvold and Tefft 2010b; Sturm et al. 2010). Further, taxes on food and beverages are shown to be regressive (Lin et al. 2011) and beverage tax rates differ across different localities. Given this background, I choose full calorie (i.e., regular) carbonated beverage purchases as the focus of this research. Regular carbonated beverages exclude diet and low-calorie carbonated beverages, though these enter the model as controls.

In the existing literature, little is known about the effects of the availability of portion-controlled packages of food and drinks on calories consumed. Specifically, it is not clear if, outside of clinically-monitored settings, consumers benefit from purchasing portion-controlled regular carbonated beverage packages by experiencing significant reduction in calories consumed relative to the conventional packages (sold in greater than 12 oz per container). While past studies have examined the effect of portion size on consumption, most portion size studies rely upon laboratory and field experiments (Rolls et al. 2004; van Kleef, Kavvouris and van Trijp 2014; Vermeer, Bruins and Steenhuis 2010). Field experiments may not fully represent free-living conditions where participants have complete control over their food choices (Raynor and Wing 2007). Some of the drawbacks of these studies are a lack of repeated observations over an extended period, lack of sufficient statistical power, and a focus on the consumption decisions where consumers have limited autonomy. Further, almost all studies rely on individual decisions and do not observe household-level food and beverage consumption or purchases. Most of the food and beverages consumed at-home are purchased through household-level decisions, which indicates the need for a household-level analysis.

Two studies are most closely related to this research. Wilson, Stolarz-Fantino and Fantino (2013), examined whether a maximum limit on per-unit volume of sugary drinks sold in fast food restaurants would still be effective if businesses including restaurants convert a larger-sized drink into bundles of smaller-sized drinks. They found that participants bought significantly more sugary drinks with bundled drink options (multi pack) as opposed to when varying sized single drinks were available without bundling. However, Wilson, Stolarz-Fantino and Fantino's (2013) study relied on self-reported hypothetical purchase choices rather than actual purchases and their focus was to analyze the consequences of a specific policy which was later repealed in New York City that restricted the serving size of sugary drinks to be a maximum size of 16 oz at restaurants and other food outlets.

John, Donnelly and Roberto (2017) overcome the drawbacks in Wilson, Stolarz-Fantino and Fantino (2013). However, they found the opposite result, in which bundling caused people to buy fewer sugary drinks. They also focus on the same policy as Wilson, Stolarz-Fantino and Fantino (2013), however, John, Donnelly and Roberto (2017) use laboratory experiments where the participants actually made purchases and consumed the drinks they bought in either simulated waiter-service style or self-service style restaurants. John, Donnelly and Roberto (2017) argue that the results from Wilson, Stolarz-Fantino and Fantino (2013) could be measurement error caused by the way the orders were elicited. However, neither study explores the relationship between the purchases of specific bundles and other healthy dietary habits.

This essay will overcome some of the limitations in the previous research by observing households' regular carbonated beverage purchases and associated characteristics under free-living conditions using consumer panel data. Observing purchasing decisions is essential for understanding portion size effects because purchasing is an antecedent to the self-control of consumption behavior. For example, consumers who want to control their consumption of caloric foods that are likely to be consumed on impulse can voluntarily and strategically ration their purchased quantities (Wertenbroch 1998). This kind of control is difficult to exercise after purchases are made, though purchasing smaller units provides another dimension for exercising such control. Given this background, my specific objectives are as follows:

1) Identify the characteristics of households, if any, that predict purchases of relatively more portion-controlled sizes (i.e., less than 12 oz) of full calorie carbonated beverages by volume
2) Assess if households who purchase relatively more portion-controlled sizes of full calorie carbonated beverages by volume also engage in other dietary behaviors that signal healthful intentions

3.2. Data

I use NielsenIQ Homescan Consumer Panel (NHCP) data on purchases made by consumers from 2012 to 2017. The NHCP consists of a panel of households who scan their purchases using at-home scanner technology after all grocery and other shopping trips from stores they usually visit. Besides the price and quantity of the carbonated sugary drinks bought, the data contain information on volume per unit, whether the product bought is a multipack or not, and several product attributes such as flavor, type of container (e.g., plastic, glass, can), and type (e.g., caffeinated, sugar-free). Each product is uniquely identified using a

Universal Product Code (UPC). Various household characteristics as documented from 2012 to 2017 are also available.

The NielsenIQ dataset comprises a representative panel of 40,000-60,000 active panelist households in each panel year with a retention rate of 80% from one year to another. The sampling of panelists follows a proportionate random sampling approach in which the key demographic characteristics of panelists are matched to the demographics of the continental US population and regular checks are made to ensure the representativeness. NielsenIQ samples all states except Alaska and Hawaii.

For the purpose of this essay, I use a balanced panel of 10,050 households who purchase both regular and diet or low-calorie carbonated beverages at least once in every three months from 2012-2017, for a span of six years. I aggregate the purchases into years. I exclude households whose total spending on food is in the bottom 5%. I also exclude households who only purchase diet carbonated beverages. I use indicator variables to identify the 15 major brands of carbonated beverages which account for 78% of market by sales volume and include the rest in 'Other Brands' category. There are 60,300 household-year observations in total.

3.3. Methods

I use machine learning methods, specifically random forest methods, to predict households who purchase relatively more portion-controlled sizes of regular (i.e., full calorie) carbonated beverages by volume and to identify which demographic, socio-economic, and dietary characteristics can predict household behavior towards purchasing portion-controlled sizes of regular carbonated beverages.

The random forest method is based on a tree-based machine-learning algorithm (Breiman 2001). As a non-parametric method, it does not require an explicit functional form for the relationship between the outcome and predictors. Further, compared to other machine learning methods, random forest allows non-linear relationships and interactions among predictor variables, and nonlinear relationships between predictors and outcomes (Hut and Oster 2018; Sage 2018). The trees grown are called regression trees and they group households that, for the purposes of this study, are similar in their volume shares of less than 12 oz beverages. The upper limit of a single unit's volume is limited at two liters in this study. The trees are grown by splitting bootstrapped samples of the training data using a set of features (i.e., independent variables or predictors) to generate predictions of the outcome of interest. The training data is a fraction of the data used to train the random forest. The fit is evaluated based on the out-of-bootstrapped sample's prediction performance (known as out-of-bag mean-squared-error). I use a variant called historical random forest ('*htree:hrf*' package in R), which is suitable for panel data (Sexton 2018).[33]

Historical regression trees produce a non-parametric estimate of how the response variable depends on all of its prior realizations and that of any time-varying features (Sexton 2018). Historical random forest estimates a model for the outcome variable y_{ij} for household i ($i = 1, \ldots, n$) at the j^{th} ($j = 1, \ldots, n_i$) observation time t_{ij} using a vector of

[33] The random forest algorithms for panel data are still limited and the standard random forest packages available are not tested on panel data.

predictors x_{ij}. Data is assumed to be of form: $(y_{ij}, x_{ij},\ t_{ij})$. The subscripts i and j in time t_{ij} is allowed to vary across households though this is not needed for this study because the sample is a balanced panel of households. The vector of predictors are input as two groups—historic and concurrent predictors. The estimation uses both $(t_{ij},\ x_{ij})$ and all preceding observations of the i^{th} household leading up to but not including time t_{ij}. For concurrent predictors, nodes are split using the approach for standard regression trees, i.e., looking for the best concurrent predictor and best cut point to split the data to minimize the residual sums of squares within each of the resulting regions. For historical predictors, the nodes are split to minimize:

$$argmin \sum_{(ij)\in Node} (y_{ij} - \mu_L I(s(\eta; \bar{z}_{ijk}) < c) - (y_{ij} - \mu_R I(s(\eta; \bar{z}_{ijk}) \geq c))^2 \quad \text{(Eq.3.3)}$$

where the minimization is over the vector (k, μ, c, η); k is the predictor, μ is the region mean, c is the cut point, and η is the argument vector of the summary function. In the above equation, μ_L and μ_R are the mean outcome of the observations in left and right split regions, respectively. The preceding values of a historical predictor are transformed into a summary function and is denoted by $s(\eta; \bar{z}_{ijk})$ where $\bar{z}_{ijk}$ indicates the set of historical values of the k^{th} predictor and η is the argument vector of the summary function. Various summary functions are available and the one used in this study is "mean0" which is shown below.

$$s(\eta; \bar{z}_{ijk}) = \sum_{h: t_{ij} - \eta_1 \leq t_{ih} < t_{ij}} \frac{z_{ihk}}{\eta_{ij}(\eta)} \quad \text{(Eq.3.2)}$$

where $\eta_{ij}(\eta)$ is the number of observations of household i in the time window $[t_{ij} - \eta_1, t_{ij})$.

The performance of the random forest is based on the mean-squared error (MSE). Since trees are fitted on a bootstrapped subset of observations, the remaining observations that are not used to fit a given tree—out-of-bag (OOB) observations—are used to calculate the OOB MSE. Tree parameters can be tuned to enhance the prediction and performance of the trees and are discussed in section 3.4.

3.3.1. Response Variable, Historical Predictors, and Concurrent Predictors

The response variable is the volume share of portion-controlled (less than 12 oz) full calorie carbonated beverages of household *i* bought in year *j*. The size bin is chosen based on the pre-dominant market category for portion-controlled sizes. Rather than focusing on individual sizes, I group the sizes as less than 12 oz given that the frequency of purchases in this category is modest (i.e., there are not sufficient observations in each size category for group analysis). Table 3.1 summarizes the share of purchases by portion size, container, and formula types. Only about 2.7 percent of all yearly household purchases are of less than 12 oz beverages. About 3.9 percent of full calorie and 1.2 percent of diet purchases include less than 12 oz beverages. About 16.2 percent of the beverages sold in glass containers are of less than 12 oz.

Table 3.1. Share of household carbonated beverage purchases by size, container type, and formula type (2012-2017)

Characteristics	Percentage
<u>Size category</u>	
Purchases of portion control sizes (less than 12 oz)	2.67
Purchases of conventional sizes (12 oz or greater)	97.33
<u>Container type</u>	
Share of purchases of cans (e.g., aluminum) that are less than 12 oz	6.30
Share of purchases of cans (e.g., aluminum) that are more than 12 oz	93.70
Share of purchases of plastic bottles that are less than 12 oz	0.13
Share of purchases of plastic bottles that are more than 12 oz	99.87
Share of purchases of glass bottles that are less than 12 oz	16.19
Share of purchases of glass bottles that are more than 12 oz	83.81
<u>Formula type</u>	
Share of purchases of full calorie beverages that are less than 12 oz	3.89
Share of purchases of full calorie beverages that are more than 12 oz	96.11
Share of purchases of diet or low calorie beverages that are less than 12 oz	1.16
Share of purchases of diet or low calorie beverages that are more than 12 oz	98.84

Source: Author's calculations from NHCP.

I use three sets of features indexed by household-year (*ij*). The first set includes household demographics (age, education, income, household size, marital status, employment status, race, presence of elderly and children, whether the household ever participated or currently participates in Women, Infants, and Children (WIC) program, and type of housing, including single-family house, apartment/condo, and trailer/mobile home). The type of house is a proxy for storage conditions, which influences bulk carbonated beverage purchases (Wang 2015). The second set of features includes household dietary characteristics as proxies for relatively healthy versus unhealthy food and beverage purchases (food expenditure shares on fresh produce, snacks, and volume share of diet carbonated beverages). The final set includes factors related to carbonated beverage purchases (average price per ounce, whether the purchase includes only metal cans, only plastic bottles or only glass containers, brand indicators, volume share of multipacks, volume share of deals and promotions).

I implement the main random forest using the '*hrf*' package in R-Studio which uses historical regression trees. I categorize the diet-related variables and factors related to carbonated beverage purchases as the historical predictors. All variables also act as concurrent predictors. Since the price of the beverages can be endogenous, I include the average price from nearby markets as an instrument for price. The households are categorized into 56 geographically-defined markets based on the scan-track market codes used by NielsenIQ.

3.4. Results and Discussion

Table 3.2 shows the summary statistics of the demographic features for the NielsenIQ sample used in this study. Even though the NielsenIQ data is designed to be representative of US demographics, the sample developed for this study is particularly skewed towards higher income households and racial composition is mostly white (Einav, Leibtag and Nevo 2008; Zhen et al. 2009).

Table 3.2. Summary Statistics

Variable	Mean	Standard Deviation	US average (2017)
HH size	2.44	1.16	2.65
HH head years of education	14.27	1.93	13.7
HH head age	59.06	10.37	51.9
White (0/1)	0.88	0.33	0.76
Children (0/1)	0.18	0.39	0.31
Married (0/1)	0.75	0.43	0.48
Employed (proportion)			
Unemployed	0.29		0.28
Fully employed	0.46		0.53
Partly employed	0.25		0.18
HH Income	$67,997	$40,097	$61,372^{+}
Presence of adults older than 64 years old (0/1)	0.14	0.34	0.40
Participation in WIC (0/1)	0.09	0.29	
House type (proportion)			
One-unit structures	0.89		0.69
Two- or more unit structures	0.06		0.26
Mobile homes/ Trailers	0.05		0.06

Notes: $^{+}$ indicates the median. US averages are obtained from US Census Bureau and Statistics – American Community Survey 2017.

Table 3.3 shows the summary statistics of sampled households' carbonated beverage purchase behavior and dietary characteristics. The average purchased volume share of full calorie carbonated beverages of less than 12 oz is 1.4 percent annually. The yearly average volume share of diet beverages purchased is higher at 50 percent among the households who buy both diet and regular carbonated beverages. About 3 and 9 percent of the households only buy beverages sold in metal cans and plastic bottles, respectively. Less than 1 percent of the households only purchase beverages sold in glass containers. The yearly volume share of deals and promotions does not exceed half of all purchases. The average annual food expenditure share of fresh produce and snacks are about 7 and 5 percent, respectively.

Table 3.3. Carbonated Beverage Purchases and Diet Characteristics

Variable	Mean	Standard Deviation
Volume share of carbonated beverages of less than 12 oz (out of all carbonated beverages)	0.014	0.067
Volume share of diet carbonated beverages (out of all carbonated beverages)	0.503	0.394
Whether the carbonated beverage purchases include only metal cans (0/1)	0.031	0.173
Whether the carbonated beverage purchases include only plastic bottles (0/1)	0.091	0.288
Whether the carbonated beverage purchases include only glass (0/1)	0.004	0.021
Volume share of multipacks (out of all carbonated beverages)	0.624	0.348
Volume share of deals and promotions (out of all carbonated beverages)	0.443	0.372
Average price per ounce	\$0.027	\$0.001
Food expenditure share of fresh produce	0.067	0.049
Food expenditure share of snacks	0.050	0.032

Source: Author's calculations from NHCP. *Note:* The volume shares are calculated as a proportion of all purchased carbonated beverage ounces for each household and averaged across all households and years. Based on the zip codes of households' residents, there are 240 households with corresponding zip codes that do not have any observations with carbonated beverage purchases of less than 12 oz containers; however, it is possible that these households have access to stores located at different zip codes that carry portion-controlled sizes of carbonated beverages. The store zip codes are restricted to only the first three digits in the Nielsen data, so this information cannot be used to identify the zip codes corresponding to stores carrying the portion-controlled sizes. Assuming these 240 households have access to portion-controlled sizes but choose not to purchase them, they are included in the analysis.

The following tree parameters are used to fit the historic random forest based on the tree performance and default rules for regression trees (Table 3.4). The number of trees are chosen based on the existing literature using the NHCP data set for studying consumers' diet-related behaviors (Oster 2018). The recommended number of predictors sampled at each split is $p/3$ where p is the number of predictors used; in this study p = 39 (Probst, Wright and Boulesteix 2019). The default number of bootstrap samples with replacement is 100. Due to the computational burden this choice can impose, it is set at 200 rather than a higher number. The minimum number of training observations in a terminal node is 5 by default, but in this study, it is set to 50. This was chosen based on the tree performance and stability after growing trees with varying node sizes from 25 to 100. The fraction of the sample used to train each tree was chosen based on the existing literature.

Table 3.4. Tree Parameters

Parameters	**Size**
Number of trees in ensemble (*ntrees*)	300
Historical summary method (*method*)	mean0
Number of predictors	39
Number of predictors sampled at each split ($mtry = p/3$)	13
Number of bootstrap samples (B)	200
Minimum number of training observations in a terminal node (*nodesize*)	50
Fraction of data sample to train each tree (*sample_fraction*)	0.75

Figure 3.1 plots the OOB-MSE against the number of trees. This is used to observe the tree performance as the number of trees is increased. The OOB-MSE stabilizes close to 0.0042 around 75 trees. In the main estimation I grow 300 trees. Figure 3.2 shows three OOB-MSE curves when different numbers of trees and node sizes are used. The resulting graphs curves are close to the original one.

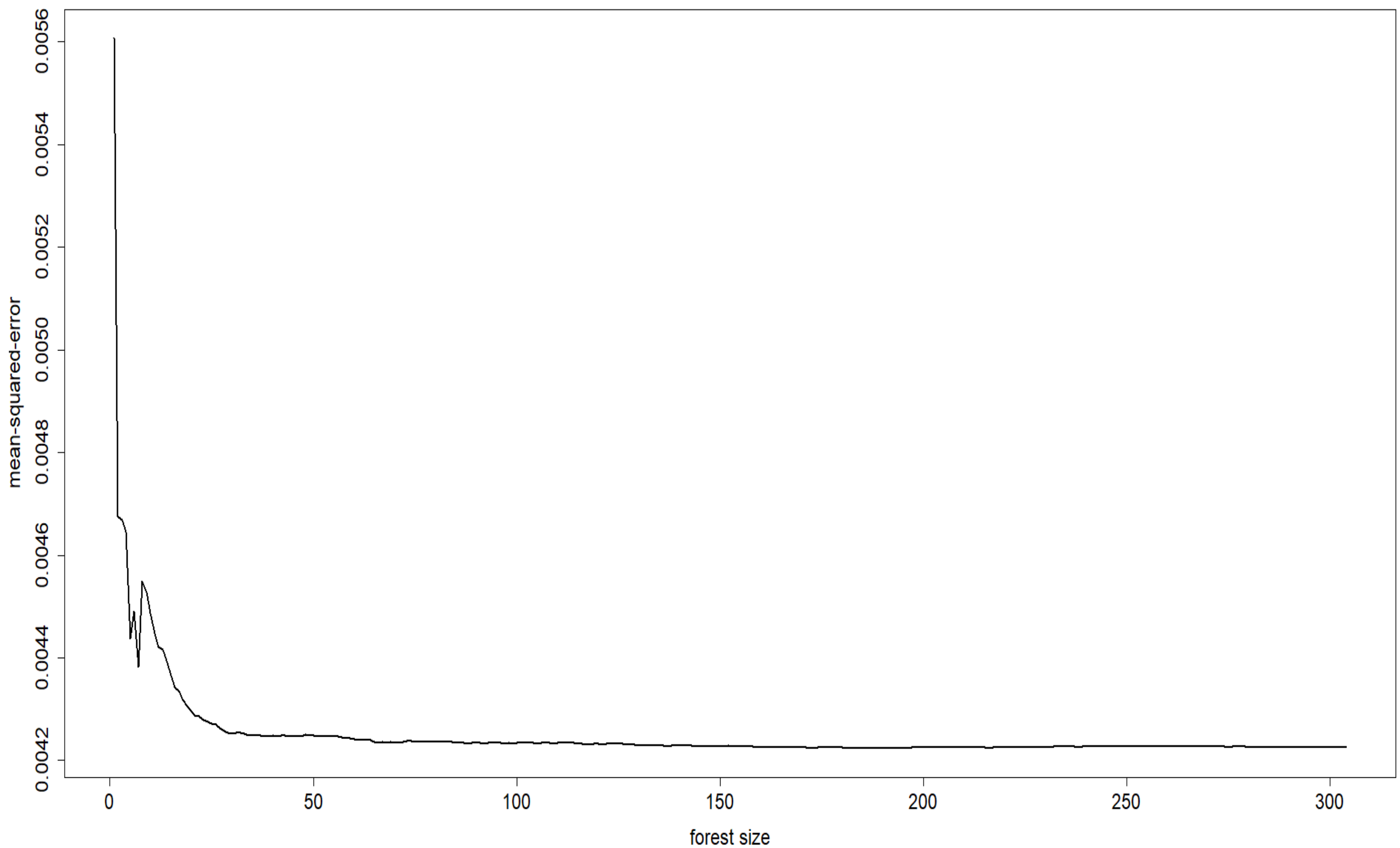


Figure 3.1. Out-of-bag mean squared error of the final tree

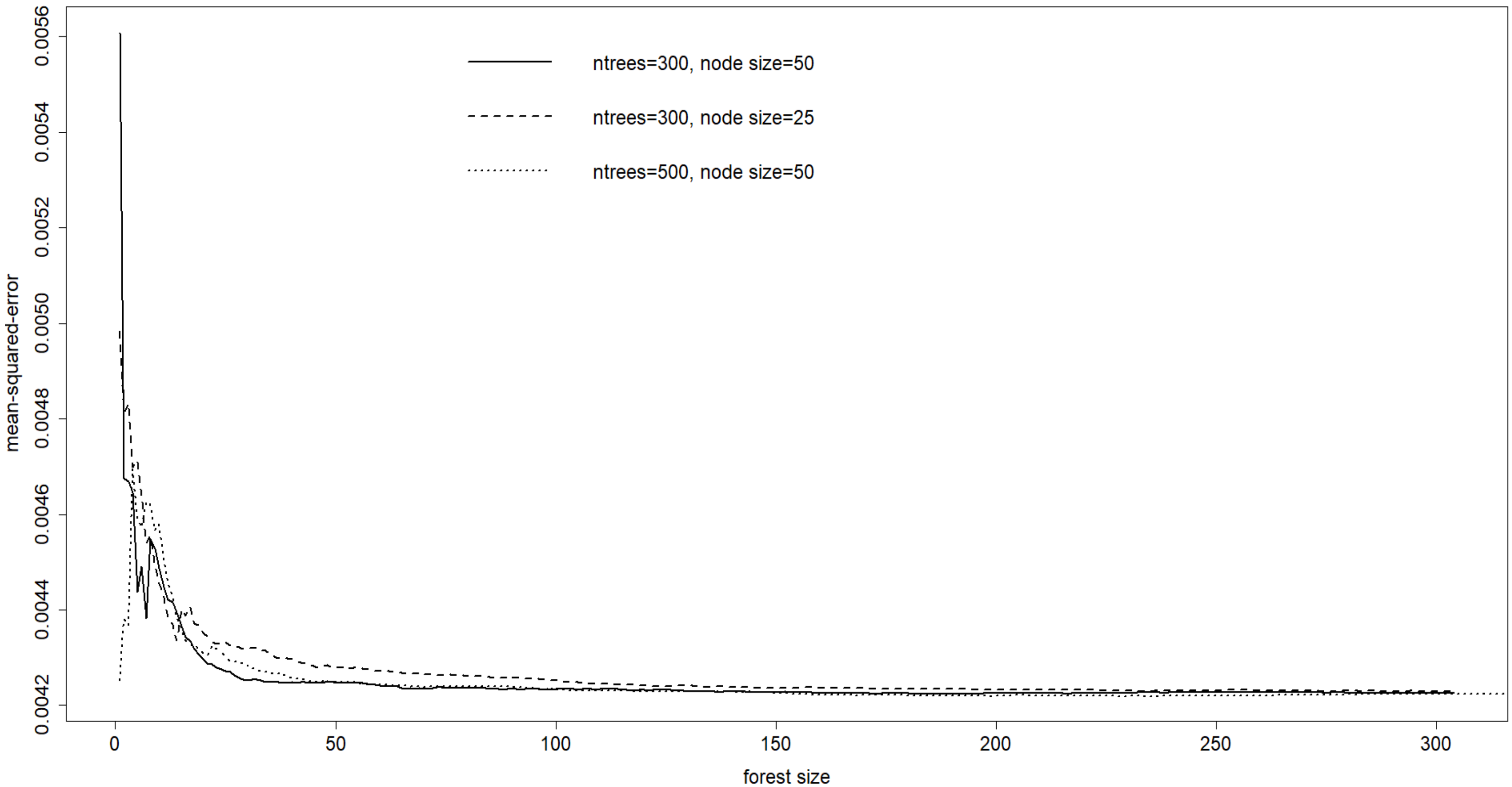


Figure 3.2. OOB-MSE with different number of trees and node sizes.
Note: The final OOB-MSE with parameters in Table 3.4 is shown in solid line. The dashed line indicates OOB-MSE when node size is changed to 25, and the dotted line indicates the OOB-MSE when the number of trees is changed to 500.

3.4.1. Variable Importance

The random forest method allows me to determine which features of the forest are the most predictive of the outcome variable. The standard method finds the important features via the variable importance permutation (VIMP). By default, VIMP is calculated for the training data and measured by comparing the estimation prediction errors with the prediction errors after integrating the predictor out of the model (Sexton 2018). The integration is done by averaging over multiple predictions from the estimated model, each obtained using a random permutation of the observed values of a feature. A corresponding Z-score from a paired test for the equality of the prediction errors is also estimated for each feature. Larger Z-scores mean the feature is important, however sometimes this also can be due to smaller standard errors so, comparing estimated predictor error and marginalized error is more accurate. Table 3.5 shows the VIMP measures of the top features ranked by their importance. Based on the VIMP, some purchase-related and dietary variables are important in predicting the volume share of portion-controlled size purchases. Two of the three dietary variables that are proxies for healthy purchases—the expenditure share of fresh produce and the volume share of diet carbonated beverages—are among the top important features. However, being richer or more educated does not predict the same behavior.

Table 3.5. Variable Importance Measures

(1) Predictor	(2) Marginalized error	(3) Model error	(4) Relative change	(5) Z-value
Volume share of diet carbonated beverages	0.0044	0.0042	0.045	38.294
Volume share of multipack purchases	0.0043	0.0042	0.008	12.057
Volume share of deals and promotions	0.0043	0.0042	0.013	4.157
Expenditure share of fresh produce	0.0043	0.0042	0.006	20.285

Notes: This table shows the variable importance permutation measures of the top four features ranked by their importance. The first column lists the most predictive features of the outcome variable. Column 3 shows the estimation prediction errors of the estimated model. Column 2 lists the prediction errors obtained after integrating the predictors in column 1 out. The relative change between column 2 and column 3 is given in column 4 and the corresponding Z-score from a paired test for the equality of columns 2 and 3 is listed in column 5.

Another key finding is the structure of the relationship. Partial dependence plots illustrate the marginal effects of the features (Figures 3.3 through 3.6). There is a linear relationship between volume share of diet beverage purchases and volume share of portion-controlled sizes of full calorie carbonated beverages (Figure 3.3). Purchases of portion-controlled sizes are more common in households that purchase proportionately more full calorie carbonated beverages. This suggests those households' tendencies toward limiting calories consumed in a single serving from full calorie carbonated beverages, as those who have a larger share of full calorie beverages lean towards buying the portion-controlled sizes.

Volume share of multipack purchases has a somewhat non-linear relationship with the volume share of portion-controlled sizes; the volume share of multipack is positively related to the volume share of portion-controlled sizes up to certain extent and flattens afterwards. This is plausible because as a household's share of multipack purchases increases, the non-portion-controlled sizes are more likely to be purchased to extract the price advantage of multipacks sold in greater than 12 oz containers (Figure 3.4). The volume share of less than 12 oz full calorie beverages is lower as households buy more beverages in deals and promotions (Figure 3.5). This is expected given that the less than 12 oz beverages have a higher price per ounce than the larger ones. The food expenditure share of fresh produce is associated positively with the volume share of less than 12 oz full calorie beverages (Figure 3.6). This also suggests that healthy dietary behavior is associated with portion-controlled sizes of beverage purchases. Given that the mean and standard deviation of food expenditure share of fresh produce are 0.067 and 0.049, respectively, the curve beyond the 0.2 mark on the *x*-axis is mostly flat. Almost all other features have flat partial dependence curves suggesting that those features do not associate with the portion-controlled sizes of full calorie beverage purchases (figures are excluded).

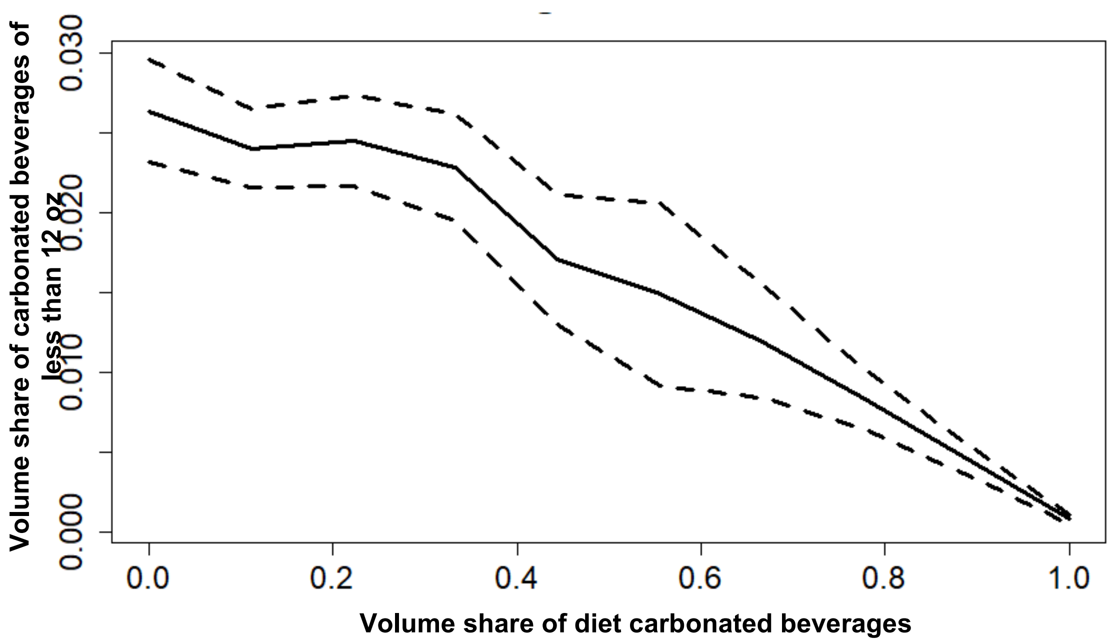


Figure 3.3. Marginal effects of volume share of diet carbonated beverages.
Note: The dashed lines show the 90% confidence intervals.

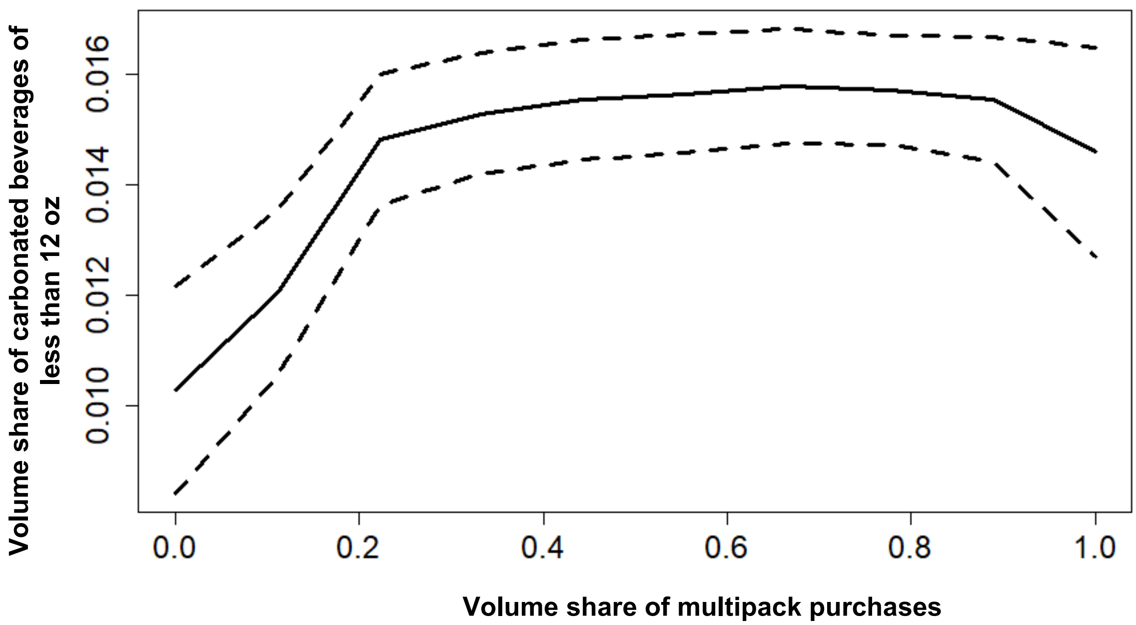


Figure 3.4. Marginal effects of volume share of multipack purchases.
Note: The dashed lines show the 90% confidence intervals.

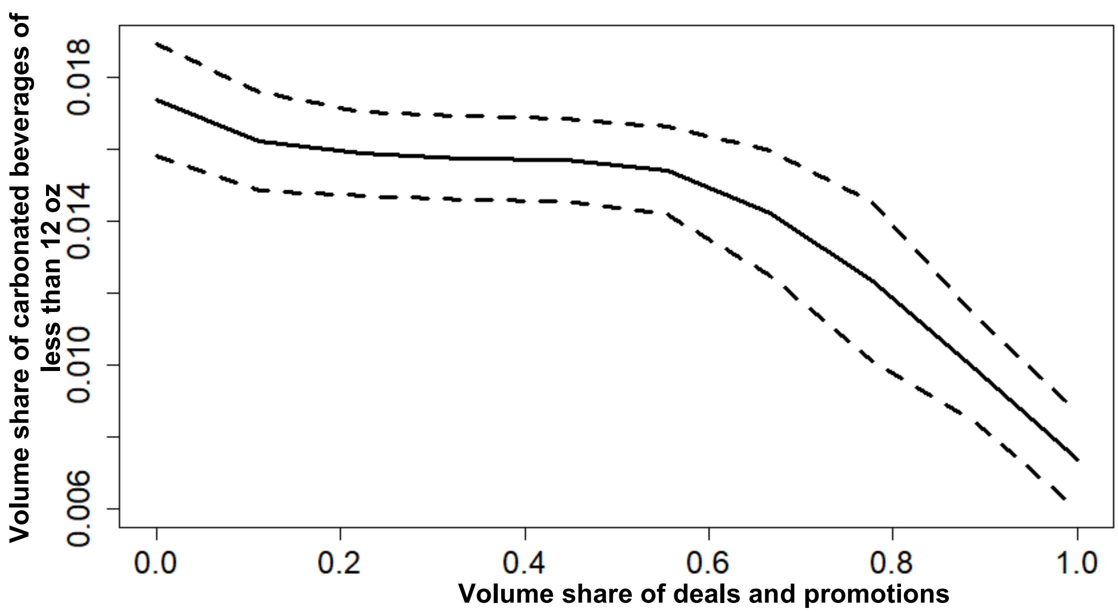


Figure 3.5. Marginal effects of volume share of deals and promotions

Note: The dashed lines show the 90% confidence intervals.

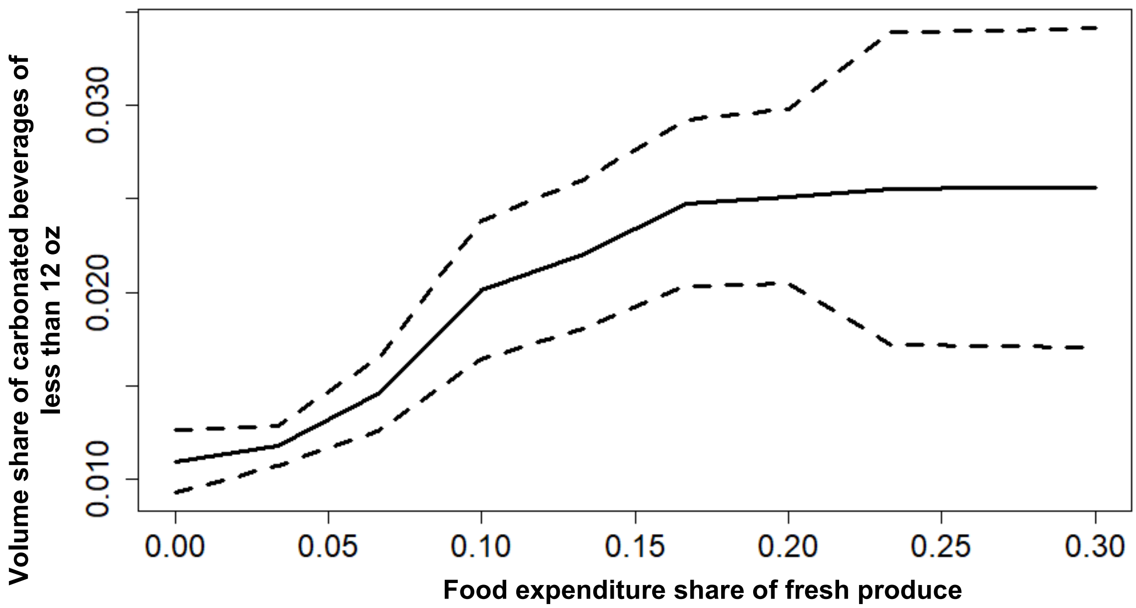


Figure 3.6. Marginal effects of expenditure share of fresh produce
Note: The dashed lines show the 90% confidence intervals.

Based on the findings, among the dietary variables, the volume share of diet carbonated beverages and the food expenditure share of fresh produce predict purchases of portion-controlled sizes of full calorie carbonated beverages. Income or education do not predict this purchasing behavior. Neither do the presence of elderly people or children. The volume share of less than 12 oz full calorie beverages is lower as households buy more beverages in deals and promotions. Multipacks are positively associated with purchasing, but the size of this effect is small. Overall, there is some suggestive evidence that is associated with portion control behavior in carbonated beverage consumption in real life settings.

Coca-Cola introduced mini-cans in 2009 as one of the many ways to attract health conscious consumers to buy more regular carbonated beverages. Since these were rolled out in some markets before others, a difference-in-differences research design would be ideal for this study. However, the NielsenIQ data do not have sufficient observations for a difference-in-differences design and the timeline of the mini-can rollout is not publicly available. This could be an avenue for future research.

There are several limitations to this work. First, I only observe households' food-at-home consumption of carbonated beverages. The NHCP does not include food-away-from-home consumption. Therefore, it is impossible to know the differences or similarities between the two kinds of purchasing and consumption behaviors. Second, even though the household storage condition is proxied by house type, it may not accurately capture the actual storage condition which influences bulk purchase decisions, and no proxies are included to capture the sharing of beverages between household members or the convenience aspect such as occasional purchases for a one-time consumption. Further, it is presumed that reduced purchase from portion-controlled sizes would not merely lead to perfectly correlated reductions in waste, but it correlates with consumption. Third, there are portion-controlled food products, such as 100-calorie snack packages, which can be used to test the external validity of the approach used in this essay for other types of products. This is another avenue for future research.

3.5. Conclusions

In this essay, I investigate what characteristics of households, if any, predict purchases of portion-controlled sizes of full calorie carbonated beverages and whether this behavior is associated with other healthy dietary habits. I use NielsenIQ Homescan Consumer Panel data on purchases made by consumers from 2012 to 2017 and use a machine learning method called historical random forest to answer the research questions. I find that household demographics including income and education are not associated with purchasing behavior of less than 12 oz beverages. Neither are the presence of children or the elderly. However, the behavior is predicted by volume share of diet carbonated beverages and food expenditure share of fresh produce, which are two of the three proxies used to capture healthy dietary habits. These results suggest that there is an association between purchases of smaller packages (less than 12 oz) of regular carbonated beverages and portion control behavior.

Bibliography

Adams, D.C., and A.E. Adams. 2008. “Availability, Attitudes and Willingness to Pay for Local Foods: Results of a Preliminary Survey.”

An, R. 2013. “Effectiveness of subsidies in promoting healthy food purchases and consumption: a review of field experiments.” *Public Health Nutrition* 16(7):1215–1228.

Anderson, J.W., E.C. Konz, R.C. Frederich, and C.L. Wood. 2001. “Long-term weight-loss maintenance: a meta-analysis of US studies.” *The American Journal of Clinical Nutrition* 74(5):579–584.

Andrews, K.R., K.S. Silk, and I.U. Eneli. 2010. “Parents as Health Promoters: A Theory of Planned Behavior Perspective on the Prevention of Childhood Obesity.” *Journal of Health Communication* 15(1):95–107.

Angrist, J.D., and J.-S. Pischke. 2014. *Mastering ’Metrics: The Path from Cause to Effect*. Princeton University Press.

Aribarg, A., N. Arora, and H.O. Bodur. 2002. “Understanding the Role of Preference Revision and Concession in Group Decisions.” *Journal of Marketing Research* 39(3):336–349.

Aribarg, A., N. Arora, and M.Y. Kang. 2010. “Predicting Joint Choice Using Individual Data.” *Marketing Science* 29(1):139–157.

Baicker, K., H.L. Allen, B.J. Wright, and A.N. Finkelstein. 2017. “The Effect of Medicaid On Medication Use Among Poor Adults: Evidence from Oregon.” *Health Affairs* 36(12):2110–2114.

Baicker, K., A. Finkelstein, J. Song, and S. Taubman. 2014. “The Impact of Medicaid on Labor Market Activity and Program Participation: Evidence from the Oregon Health Insurance Experiment.” *American Economic Review* 104(5):322–328.

Baicker, K., S.L. Taubman, H.L. Allen, M. Bernstein, J.H. Gruber, J.P. Newhouse, E.C. Schneider, B.J. Wright, A.M. Zaslavsky, and A.N. Finkelstein. 2013. “The Oregon experiment—effects of Medicaid on clinical outcomes.” *New England Journal of Medicine* 368(18):1713–1722.

Barbaresco, S., C.J. Courtemanche, and Y. Qi. 2015. "Impacts of the Affordable Care Act dependent coverage provision on health-related outcomes of young adults." *Journal of Health Economics* 40:54–68.

Barlow, S.E., and W.H. Dietz. 1998. "Obesity Evaluation and Treatment: Expert Committee Recommendations." *Pediatrics* 102(3):e29–e29.

Bateman, I.J., and A. Munro. 2009. "Household Versus Individual Valuation: What's the Difference?" *Environmental and Resource Economics* 43(1):119–135.

Bean, M., and J.S. Sharp. 2011. "Profiling Alternative Food System Supporters: The Personal and Social Basis of Local and Organic Food Support." *Renewable Agriculture and Food Systems* 26(03):243–254.

Beck, M.J., and S. Hess. 2016. "Willingness to accept longer commutes for better salaries: Understanding the differences within and between couples." *Transportation Research Part A: Policy and Practice* 91:1–16.

Becker, G.S. 1974. "A Theory of Social Interactions", *Journal of Political Economy* 82(6): 1095–117.

Becot, F., J.M. Kolodinsky, E. Roche, A.E. Zipparo, L. Berlin, E. Buckwalter, and J. McLaughlin. 2017. "Do Farm-to-School Programs Create Local Economic Impacts?" *Choices* 32(1).

Beharry-Borg, N., D.A. Hensher, and R. Scarpa. 2009. "An Analytical Framework for Joint vs Separate Decisions by Couples in Choice Experiments: The Case of Coastal Water Quality in Tobago." *Environmental and Resource Economics* 43(1):95–117.

Binkley, J.K., J. Eales, and M. Jekanowski. 2000. "The relation between dietary change and rising US obesity." *International Journal of Obesity* 24(8):1032–1039.

Bogers, R.P., W.J.E. Bemelmans, R.T. Hoogenveen, H.C. Boshuizen, M. Woodward, P. Knekt, R.M. van Dam, F.B. Hu, T.L.S. Visscher, A. Menotti, R.J. Thorpe, K. Jamrozik, S. Calling, B.H. Strand, M.J. Shipley, and BMI-CHD Collaboration Investigators. 2007. "Association of overweight with increased risk of coronary heart disease partly independent of blood pressure and cholesterol levels: a meta-analysis of 21 cohort studies including more than 300 000 persons." *Archives of Internal Medicine* 167(16):1720–1728.

Botkins, E.R. 2017. *Three Essays on the Economics of Food and Health Behavior*. The Ohio State University. Available at: https://etd.ohiolink.edu/pg_10?0::NO:10:P10_ACCESSION_NUM:osu149208205990797 [Accessed August 26, 2019].

Botkins, E.R., and B.E. Roe. 2018. “Understanding Participation in Farm to School Programs: Results Integrating School and Supply-Side Factors.” *Food Policy* 74: 126–137.

Breiman, L., 2001. “Random forests.” *Machine Learning* 45:5–32.

Brook, R.H., J.E. Ware, W.H. Rogers, E.B. Keeler, A.R. Davies, C.A. Donald, G.A. Goldberg, K.N. Lohr, P.C. Masthay, and J.P. Newhouse. 1983. “Does Free Care Improve Adults’ Health?” *New England Journal of Medicine* 309(23):1426–1434.

Burton, S., E.H. Creyer, J. Kees, and K. Huggins. 2006. “Attacking the Obesity Epidemic: The Potential Health Benefits of Providing Nutrition Information in Restaurants.” *American Journal of Public Health* 96(9):1669–1675.

Carpio, C.E., and O. Isengildina-Massa. 2009. “Consumer Willingness to Pay for Locally Grown Products: The Case of South Carolina.” *Agribusiness* 25(3):412–426.

Center on Budget and Policy Priorities. 2019. “Policy Basics: The Supplemental Nutrition Assistance Program (SNAP).” *Center on Budget and Policy Priorities*. Available at: https://www.cbpp.org/research/food-assistance/policy-basics-the-supplemental-nutrition-assistance-program-snap [Accessed September 10, 2020].

Chang, J.B., J.L. Lusk, and F.B. Norwood. 2009. “How Closely Do Hypothetical Surveys and Laboratory Experiments Predict Field Behavior?” *American Journal of Agricultural Economics* 91(2):518–534.

Chen, L., L.J. Appel, C. Loria, P.-H. Lin, C.M. Champagne, P.J. Elmer, J.D. Ard, D. Mitchell, B.C. Batch, L.P. Svetkey, and B. Caballero. 2009. “Reduction in consumption of sugar-sweetened beverages is associated with weight loss: the PREMIER trial123.” *The American Journal of Clinical Nutrition* 89(5):1299–1306.

Christensen, L.O., B. Jablonski B.R.R., L. Stephens, and A. Joshi. 2017. “Economic Impacts of Farm to School: Case Studies and Assessment Tools.” Available at: http://www.farmtoschool.org/resources-main/economic-impacts-of-farm-to-school [Accessed April 1, 2018].

Cohen, J.F.W., S. Richardson, S.B. Austin, C.D. Economos, and E.B. Rimm. 2013. “School Lunch Waste Among Middle School Students: Nutrients Consumed and Costs.” *American Journal of Preventive Medicine* 44(2):114–121.

Colombo, S., N. Hanley, and J. Louviere. 2009. “Modeling Preference Heterogeneity in Stated Choice Data: An Analysis for Public Goods Generated by Agriculture.” *Agricultural Economics* 40(3):307–322.

Connolly, C., and H. Allen Klaiber. 2014. "Does Organic Command a Premium when the Food is Already Local?" *American Journal of Agricultural Economics*, 96(4): 1102-1116.

Costanigro, M., S. Kroll, D. Thilmany, and M. Bunning. 2014. "Is It Love for Local/Organic or Hate for Conventional? Asymmetric Effects of Information and Taste on Label Preferences in an Experimental Auction." *Food Quality and Preference* 31:94–105.

Cotti, C., E. Nesson, and N. Tefft. 2019. "Impacts of the ACA Medicaid expansion on health behaviors: Evidence from household panel data." *Health Economics* 28(2):219–244.

Courtemanche, C., J. Marton, B. Ukert, A. Yelowitz, and D. Zapata. 2018. "Effects of the Affordable Care Act on Health Care Access and Self-Assessed Health After 3 Years." *INQUIRY: The Journal of Health Care Organization, Provision, and Financing* 55:004695801879636.

Courtemanche, C.J., and D. Zapata. 2014. "Does Universal Coverage Improve Health? The Massachusetts Experience." *Journal of Policy Analysis and Management* 33(1):36–69.

Darby, K., M.T. Batte, S. Ernst, and B. Roe. 2008. "Decomposing Local: A Conjoint Analysis of Locally Produced Foods." *American Journal of Agricultural Economics* 90(2):476–486.

Dave, D., and R. Kaestner. 2009. "Health insurance and ex ante moral hazard: evidence from Medicare." *International Journal of Health Care Finance and Economics* 9(4):367–390.

De Preux, L.B. 2011. "Anticipatory ex ante moral hazard and the effect of medicare on prevention." *Health Economics* 20(9):1056–1072.

Dellaert, B.G.C., M. Prodigalidad, and J.J. Louviere. 1998. "Family Members' Projections of Each Other's Preference and Influence: A Two-Stage Conjoint Approach." *Marketing Letters* 9(2):135–145.

Dengo, A.L., E.A. Dennis, J.S. Orr, E.L. Marinik, E. Ehrlich, B.M. Davy, and K.P. Davy. 2010. "Arterial destiffening with weight loss in overweight and obese middle-aged and older adults." *Hypertension (Dallas, Tex.: 1979)* 55(4):855–861.

Diliberti, N., P.L. Bordi, M.T. Conklin, L.S. Roe, and B.J. Rolls. 2004. "Increased portion size leads to increased energy intake in a restaurant meal." *Obesity Research* 12(3):562–568.

Dillender, M. 2017. “Medicaid, family spending, and the financial implications of crowd-out.” *Journal of Health Economics* 53:1–16.

Duflo, E., R. Glennerster, and M. Kremer. 2007. “Chapter 61 Using Randomization in Development Economics Research: A Toolkit.” In T. P. Schultz and J. A. Strauss, eds. *Handbook of Development Economics*. Elsevier, pp. 3895–3962. Available at: https://www.sciencedirect.com/science/article/pii/S1573447107040612 [Accessed February 24, 2021].

Durward, C.M., M. Savoie-Roskos, A. Atoloye, P. Isabella, M.D. Jewkes, B. Ralls, K. Riggs, and H. LeBlanc. 2019. “Double Up Food Bucks Participation is Associated with Increased Fruit and Vegetable Consumption and Food Security Among Low-Income Adults.” *Journal of Nutrition Education and Behavior* 51(3):342–347.

Einav, L., and A. Finkelstein. 2018. “Moral Hazard in Health Insurance: What We Know and How We Know It.” *Journal of the European Economic Association* 16(4):957–982.

Einav, L., E.S. Leibtag, and A. Nevo. 2008. “On the Accuracy of Nielsen Homescan Data.” *AgEcon Search*. Available at: https://ageconsearch.umn.edu/record/56490 [Accessed June 1, 2020].

Ello-Martin, J.A., J.H. Ledikwe, and B.J. Rolls. 2005. “The influence of food portion size and energy density on energy intake: implications for weight management.” *The American Journal of Clinical Nutrition* 82(1 Suppl):236S-241S.

Federal Register. 2017. National School Lunch Payments. 82(144): 35176.

Ferrier, P., and C. Zhen, 2014. “Explaining the Shift from Preserved to Fresh Vegetable Consumption.” 2014 Annual Meeting, Agricultural and Applied Economics Association, Minneapolis.

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J.P. Newhouse, H. Allen, K. Baicker, and O.H.S. Group. 2012. “The Oregon health insurance experiment: evidence from the first year.” *The Quarterly journal of economics* 127(3):1057–1106.

Fletcher, J.M., D. Frisvold, and N. Tefft. 2010a. “Can Soft Drink Taxes Reduce Population Weight?” *Contemporary economic policy* 28(1):23–35.

Fletcher, J.M., D. Frisvold, and N. Tefft. 2010b. “Taxing Soft Drinks and Restricting Access to Vending Machines to Curb Child Obesity.” *Health Affairs* 29(5):1059–1066.

Giraud, K.L., C.A. Bond, and J.J. Bond. 2005. “Consumer Preferences for Locally Made Specialty Food Products Across Northern New England.” *Agricultural and Resource Economics Review* 34(02):204–216.

Giustinelli, P. 2016. “Group Decision Making with Uncertain Outcomes: Unpacking Child-Parent Choice of the High School Track: Group Decision Making.” *International Economic Review* 57(2):573–602.

Greenberg, J.A. 2006. “Correcting biases in estimates of mortality attributable to obesity.” *Obesity (Silver Spring, Md.)* 14(11):2071–2079.

Gruber, J., and A. Yelowitz. 1999. “Public health insurance and private savings.” *Journal of Political Economy* 107(6):1249–1274.

Guh, D.P., W. Zhang, N. Bansback, Z. Amarsi, C.L. Birmingham, and A.H. Anis. 2009. “The incidence of co-morbidities related to obesity and overweight: A systematic review and meta-analysis.” *BMC Public Health* 9(1):88.

He, X., R.A. Lopez, and R. Boehm. 2020. “Medicaid expansion and non-alcoholic beverage choices by low-income households.” *Health Economics*: hec.4133.

Hensher, D., N. Shore, and K. Train. 2005. “Households’ Willingness to Pay for Water Service Attributes.” *Environmental and Resource Economics* 32(4): 509–531.

Hess, S., and K. Train. 2017. “Correlation and scale in mixed logit models.” *Journal of Choice Modelling* 23:1–8.

Hieke, S., A. Palascha, C. Jola, J. Wills, and M.M. Raats. 2016. “The pack size effect: Influence on consumer perceptions of portion sizes.” *Appetite* 96:225–238.

Himmelstein, G. 2019. “Effect of the Affordable Care Act’s Medicaid Expansions on Food Security, 2010–2016.” *American Journal of Public Health* 109(9):1243–1248.

Hole, A.R. 2007a. “A Comparison of Approaches to Estimating Confidence Intervals for Willingness to Pay Measures.” *Health Economics* 16(8):827–840.

Hole, A.R. 2007b. “Fitting Mixed Logit Models by Using Maximum Simulated Likelihood.” *The Stata Journal* 7(3):388–401.

Hole, A.R. 2013. “Mixed logit modeling in Stata--an overview.” United Kingdom Stata Users’ Group Meetings 2013 No. 23, Stata Users Group. Available at: https://ideas.repec.org/p/boc/usug13/23.html [Accessed August 26, 2019].

Hollands, G.J., I. Shemilt, T.M. Marteau, S.A. Jebb, H.B. Lewis, Y. Wei, J.P.T. Higgins, and D. Ogilvie. 2015. “Portion, package or tableware size for changing selection and consumption of food, alcohol and tobacco.” *The Cochrane Database of Systematic Reviews* 2015(9). Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4579823/ [Accessed February 26, 2021].

Hu, W., P. Qing, M. Batte, T. Woods, and S. Ernst. 2013. “What is Local and for What Foods Does It Matter?” *Agricultural Economics (Zemědělská ekonomika)* 59(No. 10):454–466.

Huntington-Klein, N. 2018. “College Choice as a Collective Decision.” *Economic Inquiry* 56(2):1202–1219.

Hur, I., T. Burgess-Champoux, and M. Reicks. 2011. “Higher Quality Intake from School Lunch Meals Compared With Bagged Lunches.” *ICAN: Infant, Child, & Adolescent Nutrition* 3(2):70–75.

Hut, S., and E. Oster. 2018. “Changes in Household Diet: Determinants and Predictability.” No. w24892, National Bureau of Economic Research. Available at: https://www.nber.org/papers/w24892 [Accessed April 13, 2021].

Jeffery, R.W., A. Drewnowski, L.H. Epstein, A.J. Stunkard, G.T. Wilson, R.R. Wing, and D.R. Hill. 2000. “Long-term maintenance of weight loss: current status.” *Health Psychology: Official Journal of the Division of Health Psychology, American Psychological Association* 19(1S):5–16.

John, L.K., G.E. Donnelly, and C.A. Roberto. 2017. “Psychologically Informed Implementations of Sugary-Drink Portion Limits.” *Psychological Science* 28(5):620–629.

Joshi, A., A.M. Azuma, and G. Feenstra. 2008. “Do Farm-to-School Programs Make a Difference? Findings and Future Research Needs.” *Journal of Hunger & Environmental Nutrition* 3(2–3):229–246.

Kaiser Family Foundation. 2019. “Distribution of Eligibility for ACA Health Coverage Among Those Remaining Uninsured as of 2019.” Available at: https://www.kff.org/health-reform/state-indicator/distribution-of-eligibility-for-aca-coverage-among-the-remaining-uninsured/.

Kaiser Family Foundation. 2019. “Status of State Action on the Medicaid Expansion Decision.” Available at: https://www.kff.org/health-reform/state-indicator/state-activity-around-expanding-medicaid-under-the-affordable-care-act/ [Accessed June 1, 2020].

Kaiser Family Foundation. 2013. “Summary of the Affordable Care Act.” Available at: https://www.kff.org/health-reform/fact-sheet/summary-of-the-affordable-care-act/ [Accessed June 1, 2020].

Kendler, K.S., P. C. Sham, and C. J. MacLean. 1997. “The Determinants of Parenting: An Epidemiological, Multi-informant, Retrospective Study.” *Psychological Medicine*, 27(3): 549-563.

van Kleef, E., C. Kavvouris, and H.C.M. van Trijp. 2014. “The unit size effect of indulgent food: How eating smaller sized items signals impulsivity and makes consumers eat less.” *Psychology & Health* 29(9):1081–1103.

Kuhfeld, W.F. 2003. Marketing Research Methods in SAS. SAS Institute Incorporated.

Ledikwe, J.H., J.A. Ello-Martin, and B.J. Rolls. 2005. “Portion Sizes and the Obesity Epidemic.” *The Journal of Nutrition* 135(4):905–909.

Leininger, L., H. Levy, and D. Schanzenbach. 2010. “Consequences of SCHIP Expansions for Household Well-Being.” *Forum for Health Economics & Policy* 13(1). Available at: https://www.degruyter.com/view/journals/fhep/13/1/article-fhep.2010.13.1.1201.xml.xml [Accessed June 1, 2020].

Levy, H., T. Buchmueller, and S. Nikpay. 2019. “The Impact of Medicaid Expansion on Household Consumption.” *Eastern Economic Journal* 45(1):34–57.

Lin, B.-H., T.A. Smith, J.-Y. Lee, and K.D. Hall. 2011. “Measuring weight outcomes for obesity intervention strategies: The case of a sugar-sweetened beverage tax.” *Economics & Human Biology* 9(4):329–341.

Lindhjem, H., and S. Navrud. 2009. “Asking for Individual or Household Willingness to Pay for Environmental Goods?” *Environmental and Resource Economics* 43(1):11–29.

Lundberg, S., J. L. Romich, and K. P. Tsang. 2009. “Decision-Making by Children,” *Review of Economics of the Household* 7(1): 1–30.

Marcucci, E., A. Stathopoulos, L. Rotaris, and R. Danielis. 2011. “Comparing Single and Joint Preferences: A Choice Experiment on Residential Location in Three-Member Households.” *Environment and Planning A: Economy and Space* 43(5):1209–1225.

Mariel, P., R. Scarpa, and A. Vega-Bayo. 2018. “Joint parental school choice: Exploring the influence of individual preferences of husbands and wives.” *Regional Science and Urban Economics* 68:23–35.

Matts, C. 2009. School Food 101: The Cost of School Lunch. Available online at: https://www.canr.msu.edu/resources/cost-of-school-lunch (accessed 3 September 2019).

Mazumder, B., and S. Miller. 2016. "The effects of the Massachusetts health reform on household financial distress." *American Economic Journal: Economic Policy* 8(3):284–313.

McGee, D.L. and Diverse Populations Collaboration. 2005. "Body mass index and mortality: a meta-analysis based on person-level data from twenty-six observational studies." *Annals of Epidemiology* 15(2):87–97.

Meas, T., W. Hu, M.T. Batte, T.A. Woods, and S. Ernst. 2015. "Substitutes or Complements? Consumer Preference for Local and Organic Food Attributes." *American Journal of Agricultural Economics* 97(4):1044–1071.

Millen, B.E., S. Abrams, L. Adams-Campbell, C. A. Anderson, J. T. Brenna, W. W. Campbell, S. Clinton, F. Hu, M. Nelson, M. L. Neuhouser, and R. Perez-Escamilla. 2016. "The 2015 dietary guidelines advisory committee scientific report: development and major conclusions." *Advances in nutrition*, 7(3):438-444.

Miller, S., N. Johnson, and L.R. Wherry. 2019. "Medicaid and Mortality: New Evidence from Linked Survey and Administrative Data." Working Paper Series No. 26081, National Bureau of Economic Research. Available at: http://www.nber.org/papers/w26081 [Accessed October 9, 2020].

Motta, V. 2019. "The Impact of Local Food Expenditure on School Foodservice Revenues." *Journal of School Health* 89(9):722–729.

National Farm to School Network. 2017. Benefits of Farm to School. Available online at: http://www.farmtoschool.org/resources-main/the-benefits-of-farm-to-school (accessed 1 September 2019).

Nguyen, B.T., X. Han, A. Jemal, and J. Drope. 2016. "Diet quality, risk factors and access to care among low-income uninsured American adults in states expanding Medicaid vs. states not expanding under the affordable care act." *Preventive Medicine* 91:169–171.

Nicholls, A.J., and P. Cullen. 2004. "The child–parent purchase relationship: 'pester power', human rights and retail ethics." *Journal of Retailing and Consumer Services* 11(2):75–86.

Nicklas, T.A., T. Baranowski, K.W. Cullen, and G. Berenson. 2001. "Eating Patterns, Dietary Quality and Obesity." *Journal of the American College of Nutrition* 20(6):599–608.

O'Hara, J.K., and M.C. Benson. 2019. "The impact of local agricultural production on farm to school expenditures." *Renewable Agriculture and Food Systems* 34(03):216–225.

Oster, E. 2018. "Diabetes and Diet: Purchasing Behavior Change in Response to Health Information." *American Economic Journal: Applied Economics* 10(4):308–348.

Ortega, D.L., H.H. Wang, N.J. Olynk, L. Wu, and J. Bai. 2012. "Chinese Consumers' Demand for Food Safety Attributes: A Push for Government and Industry Regulations." *American Journal of Agricultural Economics* 94(2): 489–495.

Papoutsi, G.S., R.M. Nayga, P. Lazaridis, and A.C. Drichoutis. 2015. "Fat tax, subsidy or both? The role of information and children's pester power in food choice." *Journal of Economic Behavior & Organization* 117:196–208.

Pelletier, J.E., M.N. Laska, D. Neumark-Sztainer, and M. Story. 2013. "Positive Attitudes toward Organic, Local, and Sustainable Foods Are Associated with Higher Dietary Quality among Young Adults." *Journal of the Academy of Nutrition and Dietetics* 113(1): 127–132.

Penn, J.M., and W. Hu. 2018. "Understanding Hypothetical Bias: An Enhanced Meta-Analysis." *American Journal of Agricultural Economics* 100(4):1186–1206.

Pham, M.V., and B.E. Roe. 2013. "Will Reducing the Calorie Content of School Lunches Affect Participation? Evidence from a Choice Experiment with Suburban Parents." *AgEcon Search*. Available at: https://ageconsearch.umn.edu/record/149816 [Accessed August 26, 2019].

Plakias, Zoë T., H. Allen Klaiber and B. E. Roe. 2020. "Tradeoffs in Farm to School Implementation: Larger Foodsheds Drive Greater Local Food Expenditures." *Journal of Agricultural and Resource Economics*.45(2):232-243.

Prescott, M.P., R. Cleary, A. Bonanno, M. Costanigro, B.B.R. Jablonski, and A.B. Long. 2019. "Farm to School Activities and Student Outcomes: A Systematic Review." *Advances in Nutrition*: nmz094.

Probst, P., M.N. Wright, and A.-L. Boulesteix. 2019. "Hyperparameters and tuning strategies for random forest." *WIREs Data Mining and Knowledge Discovery* 9(3):e1301.

Rao, V.R., and J.H. Steckel. 1991. "A Polarization Model for Describing Group Preferences." *Journal of Consumer Research* 18(1): 108–118.

Raynor, H.A., and R.R. Wing. 2007. "Package unit size and amount of food: do both influence intake?" *Obesity (Silver Spring, Md.)* 15(9):2311–2319.

Rolls, B.J. 2014. “What is the role of portion control in weight management?” *International Journal of Obesity* 38(S1):S1–S8.

Rolls, B.J., J.A. Ello-Martin, and B.C. Tohill. 2004. “What Can Intervention Studies Tell Us about the Relationship between Fruit and Vegetable Consumption and Weight Management?” *Nutrition Reviews* 62(1):1–17.

Rolls, B.J., L.S. Roe, T.V.E. Kral, J.S. Meengs, and D.E. Wall. 2004. “Increasing the portion size of a packaged snack increases energy intake in men and women.” *Appetite* 42(1):63–69.

Rolls, B.J., L.S. Roe, and J.S. Meengs. 2006. “Larger portion sizes lead to a sustained increase in energy intake over 2 days.” *Journal of the American Dietetic Association* 106(4):543–549.

Romich, J. L., S. Lundberg, and K. P. Tsang. 2009. “Independence Giving or Autonomy Taking? Childhood Predictors of Decision-Sharing Patterns Between Young Adolescents and Parents.” *Journal of Research on Adolescence* 19(4):587–600.

Ruud, P., 1996, “Approximation and Simulation of the Multinomial Probit Model: An Analysis of Covariance Matrix Estimation.” Working paper, Department of Economics, University of California, Berkeley. Available online at: https://pdfs.semanticscholar.org/be7c/4d2170c60452497c4a21789b518167a0a8d2.pdf

Rummo, P.E., D. Noriega, A. Parret, M. Harding, O. Hesterman, and B.E. Elbel. 2019. “Evaluating A USDA Program That Gives SNAP Participants Financial Incentives to Buy Fresh Produce in Supermarkets.” *Health Affairs* 38(11):1816–1823.

Rungie, C., R. Scarpa, and M. Thiene. 2014. “The influence of individuals in forming collective household preferences for water quality.” *Journal of Environmental Economics and Management* 68(1):161–174.

Sage, A. 2018. *Random forest robustness, variable importance, and tree aggregation*. Doctor of Philosophy. Ames: Iowa State University, Digital Repository. Available at: https://lib.dr.iastate.edu/etd/16453/ [Accessed April 14, 2021].

Scarpa, R., M. Thiene, and D.A. Hensher. 2012. “Preferences for tap water attributes within couples: An exploration of alternative mixed logit parameterizations: WTP Differences for Tap Water Within Couples.” *Water Resources Research* 48(1). Available at: http://doi.wiley.com/10.1029/2010WR010148 [Accessed December 16, 2019].

Schmidt, L., L. Shore-Sheppard, and T. Watson. 2019. "The Impact of Expanding Public Health Insurance on Safety Net Program Participation: Evidence from the ACA Medicaid Expansion." No. w26504, National Bureau of Economic Research. Available at: http://www.nber.org/papers/w26504.pdf [Accessed October 19, 2020].

School Nutrition Association 2018. School Nutrition Operations Report: The State of School Nutrition 2018. Arlington, VA.

School Nutrition Association 2019. School Meal Trends & Stats. Available online at: https://schoolnutrition.org/aboutschoolmeals/schoolmealtrendsstats/ (accessed 3 September 2019).

Schulze, M.B., J.E. Manson, D.S. Ludwig, G.A. Colditz, M.J. Stampfer, W.C. Willett, and F.B. Hu. 2004. "Sugar-Sweetened Beverages, Weight Gain, and Incidence of Type 2 Diabetes in Young and Middle-Aged Women." *JAMA* 292(8):927–934.

Schwartz, M.B., K.E. Henderson, M. Read, N. Danna, and J.R. Ickovics. 2015. "New School Meal Regulations Increase Fruit Consumption and Do Not Increase Total Plate Waste." *Childhood Obesity* 11(3):242–247.

Sexton, J., 2018. "Historical tree ensembles for longitudinal data." Available at: https://CRAN. R-project.org/package=htree

Simon, K., A. Soni, and J. Cawley. 2017. "The impact of health insurance on preventive care and health behaviors: evidence from the first two years of the ACA Medicaid expansions." *Journal of Policy Analysis and Management* 36(2):390–417.

Sommers, B.D., and S. Rosenbaum. 2011. "Issues in health reform: how changes in eligibility may move millions back and forth between Medicaid and insurance exchanges." *Health affairs* 30(2):228–236.

Søndergaard, H.A., and M. Edelenbos. 2007. "What parents prefer and children like – Investigating choice of vegetable-based food for children." *Food Quality and Preference* 18(7):949–962.

Steenhuis, I., and M. Poelman. 2017. "Portion Size: Latest Developments and Interventions." *Current Obesity Reports* 6(1):10–17.

Sturm, R., L.M. Powell, J.F. Chriqui, and F.J. Chaloupka. 2010. "Soda taxes, soft drink consumption, and children's body mass index." *Health Affairs (Project Hope)* 29(5):1052–1058.

Taubman, S.L., H.L. Allen, B.J. Wright, K. Baicker, and A.N. Finkelstein. 2014. “Medicaid Increases Emergency-Department Use: Evidence from Oregon’s Health Insurance Experiment.” *Science (New York, N.Y.)* 343(6168):263–268.

The Patient Protection and Affordable Care Act. *Public law* 111(48):759–762.

Thilmany, D., C.A. Bond, and J.K. Bond. 2008. “Going Local: Exploring Consumer Behavior and Motivations for Direct Food Purchases.” *American Journal of Agricultural Economics* 90(5,). Available at: http://www.jstor.org/stable/20492389.

Todd, J.E., L. Mancino, and B.-H. Lin. 2010. “The Impact of Food Away from Home on Adult Diet Quality.” No. ID 1557129, Social Science Research Network. Available at: https://papers.ssrn.com/abstract=1557129 [Accessed September 10, 2020].

Train, K.E. 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press.

U.S. Census Bureau. 2018. Health Insurance Coverage in the Unites States. Retrieved from: https://www.census.gov/library/publications/2019/demo/p60-267.html

U.S. Department of Agriculture, Food and Nutrition Service. 2018. Child Nutrition Tables. Available online at https://www.fns.usda.gov/pd/child-nutrition-tables

Verghese, A., M. Raber, and S. Sharma. 2019. “Interventions targeting diet quality of Supplemental Nutrition Assistance Program (SNAP) participants: A scoping review.” *Preventive Medicine* 119:77–86.

Vermeer, W.M., B. Bruins, and I.H.M. Steenhuis. 2010. “Two pack king size chocolate bars. Can we manage our consumption?” *Appetite* 54(2):414–417.

Veronesi, M., F. Chawla, M. Maurer, and J. Lienert. 2014. “Climate change and the willingness to pay to reduce ecological and health risks from wastewater flooding in urban centers and the environment.” *Ecological Economics* 98: 1–10.

Volpe, R., and A. Okrent. 2012. “Assessing the Healthfulness of Consumers’ Grocery Purchases”, EIB-102, U.S. Department of Agriculture, Economic Research Service, November 2012.

Wadden, T.A., M.L. Butryn, and K.J. Byrne. 2004. “Efficacy of lifestyle modification for long-term weight control.” *Obesity Research* 12 Suppl:151S–62S.

Wadden, T.A., J.A. Sternberg, K.A. Letizia, A.J. Stunkard, and G.D. Foster. 1989. “Treatment of obesity by very low calorie diet, behavior therapy, and their combination: a five-year perspective." *International Journal of Obesity* 13 Suppl 2:39–46.

Ward, P.S., D.L. Ortega, D.J. Spielman, and V. Singh. 2014. "Heterogeneous Demand for Drought-Tolerant Rice: Evidence from Bihar, India." *World Development* 64: 125–139.

Wang, E.Y. 2015. "The impact of soda taxes on consumer welfare: implications of storability and taste heterogeneity." *The RAND Journal of Economics* 46(2):409–441.

Watson, J., D. Treadwell, and R. Bucklin. 2018. "Economic Analysis of Local Food Procurement in Southwest Florida's Farm to School Programs." *Journal of Agriculture, Food Systems, and Community Development*:61–84.

Weinberg, B. A. 2001. "An Incentive Model of the Effect of Parental Income on Children." *Journal of Political Economy* 109(2):266-280.

Welsh, J.A., E.A. Lundeen, and A.D. Stein. 2013. "The sugar-sweetened beverage wars: public health and the role of the beverage industry." *Current opinion in endocrinology, diabetes, and obesity* 20(5):401–406.

Wertenbroch, K. 1998. "Consumption Self-Control by Rationing Purchase Quantities of Virtue and Vice." *Marketing Science* 17(4):317–337.

Wilson, B.M., S. Stolarz-Fantino, and E. Fantino. 2013. "Regulating the Way to Obesity: Unintended Consequences of Limiting Sugary Drink Sizes" A. Bruce, ed. *PLoS ONE* 8(4):e61081.

Yelowitz, A.S. 1996. "Did recent medicaid reforms cause the caseload explosion in the food stamp program?" Institute for Research on Poverty Discussion Papers No. 1109–96, University of Wisconsin Institute for Research on Poverty. Available at: https://ideas.repec.org/p/wop/wispod/1109-96.html [Accessed October 9, 2020].

Young, L.R., and M. Nestle. 2003. "Expanding portion sizes in the US marketplace: Implications for nutrition counseling." *Journal of the American Dietetic Association* 103(2):231–240.

Young, L.R., and M. Nestle. 2012. "Reducing Portion Sizes to Prevent Obesity: A Call to Action." *American Journal of Preventive Medicine* 43(5):565–568.

Young, L.R., and M. Nestle. 2002. "The Contribution of Expanding Portion Sizes to the US Obesity Epidemic." *American Journal of Public Health* 92(2):246–249.

Zhang, J., M. Kuwano, B. Lee, and A. Fujiwara. 2009. “Modeling household discrete choice behavior incorporating heterogeneous group decision-making mechanisms.” *Transportation Research Part B: Methodological* 43(2):230–250.

Zhen, C., J.L. Taylor, M.K. Muth, and E. Leibtag. 2009. “Understanding differences in self-reported expenditures between household scanner data and diary survey data: a comparison of Homescan and consumer expenditure survey.” *Review of Agricultural Economics* 31(3):470–492.

Zlatevska, N., C. Dubelaar, and S.S. Holden. 2014. “Sizing Up the Effect of Portion Size on Consumption: A Meta-Analytic Review.” *Journal of Marketing* 78(3):140–154.

Appendix A. Additional Tables and Figures

Table A.1. State Medicaid Expansion and Dates

State	**Expansion**
Alaska	9/2015
Arizona	1/2014
Arkansas	1/2014
California	1/2014
Colorado	1/2014
Connecticut	1/2014
Delaware	1/2014
District of Columbia	1/2014
Hawaii	1/2014
Illinois	1/2014
Indiana	2/2015
Iowa	1/2014
Kentucky	1/2014
Louisiana	7/2016
Maryland	1/2014
Massachusetts	1/2014
Michigan	4/2014
Minnesota	1/2014
Montana	1/2016
Nevada	1/2014
New Hampshire	8/2014
New Jersey	1/2014
New Mexico	1/2014
New York	1/2014
North Dakota	1/2014
Ohio	1/2014
Oregon	1/2014
Pennsylvania	1/2015
Rhode Island	1/2014
Vermont	1/2014
Washington	1/2014
West Virginia	1/2014
Wisconsin	1/2014

Note: Data from Simon, Soni, and Cawley (2017), and Kaiser Family Foundation, 2019.

Table A.2. Effects of Medicaid Expansion on Quarterly Expenditure ($) Per Adult Equivalent Unit: Event-Study Estimates (N = 990,914)

Variables	(1) All Categories	(2) Food	(3) Non-Food
Eligible*Quarters to Expansion = 7	2.906	-0.304	1.663
	(10.830)	(7.928)	(6.285)
Eligible*Quarters to Expansion = 6	7.842	1.990	4.150
	(10.051)	(7.407)	(4.733)
Eligible*Quarters to Expansion = 5	-7.771	-6.490	-0.917
	(11.411)	(6.877)	(7.231)
Eligible*Quarters to Expansion = 4	1.898	3.378	-3.115
	(10.263)	(6.367)	(5.691)
Eligible*Quarters to Expansion = 3	10.931	7.909	1.918
	(9.763)	(6.500)	(4.941)
Eligible*Quarters to Expansion = 2	8.175	6.256	0.085
	(10.913)	(6.779)	(5.291)
Eligible*Quarters to Expansion = 1		Omitted	
Eligible*Quarters after Expansion = 0	6.559	-3.863	10.208*
	(10.616)	(6.245)	(5.285)
Eligible*Quarters after Expansion = 1	12.701	6.637	5.450
	(8.096)	(5.379)	(4.671)
Eligible*Quarters after Expansion = 2	8.005	1.955	6.441
	(9.696)	(5.339)	(5.822)

Continued

Table A.2. continued

Variables	(1) All Categories	(2) Food	(3) Non-Food
Eligible*Quarters after Expansion = 3	2.054	-2.069	6.056
	(9.068)	(5.912)	(4.625)
Eligible*Quarters after Expansion = 4	13.852*	2.395	10.580**
	(7.649)	(5.525)	(4.794)
Eligible*Quarters after Expansion = 5	5.250	0.505	5.329
	(11.226)	(6.860)	(5.253)
Eligible*Quarters after Expansion = 6	20.585	7.564	12.114*
	(13.089)	(7.253)	(6.075)
Eligible*Quarters after Expansion = 7	14.350	3.198	12.525***
	(9.540)	(6.604)	(4.238)
p-value of the F-test that coefficients on quarters to expansion are jointly equal to zero	0.2904	0.1715	0.4684

Notes: *, **, *** indicate significance at the 10%, 5%, 1% level, respectively. All models include controls for age, marital status, years of education of the household head, number of household heads, hours of employment, and presence of children in the household. Additionally, all models include household, year and quarter fixed-effects, and state-specific time trends. Robust standard errors clustered at the state-level are in parentheses. The main model drops households with income between 100 and 138% of FPL and drops the DC and states with prior expansions which are, Delaware, Massachusetts, New York, and Vermont, and households with heads less than or equal to 25 years old or greater than or equal to 65 years old.

Table A.3. Effects of Medicaid Expansion on Quarterly Expenditure ($) Per Adult Equivalent Unit of Food Categories: Event-Study Estimates
(N = 990,914)

Variables	(1) Dairy	(2) Deli	(3) Fresh Produce	(4) Packaged Meat	(5) Frozen Food	(6) Dry Grocery
Eligible*Quarters to Expansion = 7	0.191	1.467	-1.153*	-0.513	0.605	-0.901
	(0.947)	(1.364)	(0.672)	(0.731)	(1.590)	(4.443)
Eligible*Quarters to Expansion = 6	0.100	2.537*	-1.107	-0.398	0.367	0.490
	(1.013)	(1.292)	(0.923)	(0.883)	(1.440)	(4.014)
Eligible*Quarters to Expansion = 5	-1.082	1.979*	-0.405	-0.142	-0.591	-6.250
	(0.958)	(1.168)	(0.812)	(0.678)	(1.195)	(3.910)
Eligible*Quarters to Expansion = 4	0.081	0.394	-1.103	0.205	-0.088	3.889
	(0.871)	(0.995)	(0.827)	(0.501)	(1.588)	(3.448)
Eligible*Quarters to Expansion = 3	1.024	1.107	-1.604*	-0.083	2.436	5.028
	(0.822)	(1.117)	(0.861)	(0.619)	(1.598)	(3.599)
Eligible*Quarters to Expansion = 2	0.378	1.812*	-0.805	0.081	1.048	3.742
	(0.878)	(1.076)	(0.792)	(0.536)	(1.592)	(3.466)
Eligible*Quarters to Expansion = 1				Omitted		
Eligible*Quarters to Expansion = 0	-0.559	2.959	-3.422***	-0.465	-2.165	-0.212
	(0.980)	(1.942)	(0.719)	(0.612)	(1.471)	(3.226)
Eligible*Quarters after Expansion = 1	1.312	2.194	-3.092***	-0.638	0.457	6.405**
	(0.852)	(1.819)	(0.653)	(0.560)	(1.441)	(2.805)

Continued

Table A.3. continued

Variables	(1) Dairy	(2) Deli	(3) Fresh Produce	(4) Packaged Meat	(5) Frozen Food	(6) Dry Grocery
Eligible*Quarters after Expansion = 2	0.853	2.010	-2.489***	-0.270	-0.353	2.204
	(0.801)	(1.579)	(0.692)	(0.522)	(1.416)	(3.010)
Eligible*Quarters after Expansion = 3	0.509	0.728	-0.918	-0.467	0.222	-2.142
	(0.867)	(2.112)	(0.596)	(0.591)	(1.618)	(2.456)
Eligible*Quarters after Expansion = 4	0.322	1.999	-2.747***	0.540	-0.517	2.797
	(0.818)	(1.617)	(0.643)	(0.539)	(1.693)	(4.089)
Eligible*Quarters after Expansion = 5	1.180	0.262	-4.012***	0.174	-1.073	3.974
	(0.931)	(1.413)	(0.726)	(0.494)	(1.845)	(4.478)
Eligible*Quarters after Expansion = 6	2.194***	1.828	-2.187***	0.800	-0.152	5.081
	(0.814)	(1.692)	(0.588)	(0.513)	(1.711)	(5.105)
Eligible*Quarters after Expansion = 7	0.224	2.517	-1.058	0.571	-0.282	1.226
	(0.959)	(1.599)	(0.668)	(0.486)	(1.726)	(4.016)
p-value of the F-test that coefficients on quarters to expansion are jointly equal to zero	0.0860	0.2058	0.0687	0.9388	0.3855	0.0222

Notes: *, **, *** indicate significance at the 10%, 5%, 1% level, respectively. All models include controls for age, marital status, years of education of the household head, number of household heads, hours of employment, and presence of children in the household. Additionally, all models include household, year and quarter fixed-effects, and state-specific time trends. Robust standard errors clustered at the state-level are in parentheses. The main model drops households with income between 100 and 138% of FPL and drops the DC and states with prior expansions which are, Delaware, Massachusetts, New York, and Vermont, and households with heads less than or equal to 25 years old or greater than or equal to 65 years old.

Table A.4. Effects of Medicaid Expansion on Quarterly Expenditure ($) Per Adult Equivalent Unit of Non-Food Categories: Event-Study Estimates
(N = 990,914)

Variables	(1) Alcohol	(2) General Merchandise	(3) Health and Beauty	(4) Non Food Grocery
Eligible*Quarters to Expansion = 7	1.546	2.434	-1.059	0.288
	(1.564)	(3.271)	(3.642)	(2.056)
Eligible*Quarters to Expansion = 6	1.702	0.212	1.679	2.259
	(1.190)	(2.854)	(2.841)	(1.822)
Eligible*Quarters to Expansion = 5	-0.364	-3.820	2.106	0.797
	(1.597)	(4.060)	(3.550)	(2.018)
Eligible*Quarters to Expansion = 4	1.635	0.476	-3.031	-0.560
	(1.268)	(3.360)	(2.384)	(1.670)
Eligible*Quarters to Expansion = 3	1.104	1.769	-0.673	0.822
	(1.349)	(3.117)	(2.207)	(1.532)
Eligible*Quarters to Expansion = 2	1.834	1.016	0.406	-1.337
	(2.144)	(2.631)	(2.735)	(1.966)
Eligible*Quarters to Expansion = 1		Omitted		
Eligible*Quarters to Expansion = 0	0.215	6.880**	2.965	0.363
	(1.044)	(3.294)	(2.126)	(2.090)
Eligible*Quarters after Expansion = 1	0.614	-0.733	7.324***	-1.140
	(1.326)	(3.535)	(2.022)	(1.998)

Continued

Table A.4. continued

Variables	(1) Alcohol	(2) General Merchandise	(3) Health and Beauty	(4) Non Food Grocery
Eligible*Quarters after Expansion = 2	-0.391	1.647	7.799***	-3.006
	(1.204)	(3.135)	(2.630)	(2.073)
Eligible*Quarters after Expansion = 3	-1.933	-0.726	7.798***	-1.016
	(1.390)	(2.990)	(2.694)	(1.854)
Eligible*Quarters after Expansion = 4	0.878	7.670***	6.076	-3.167
	(1.231)	(2.527)	(4.298)	(2.073)
Eligible*Quarters after Expansion = 5	-0.584	2.517	5.630	-2.818
	(1.274)	(2.750)	(3.810)	(2.058)
Eligible*Quarters after Expansion = 6	0.908	8.193**	6.907*	-2.986
	(1.764)	(3.473)	(3.888)	(2.226)
Eligible*Quarters after Expansion = 7	-1.372	9.391***	5.986*	-2.853*
	(1.509)	(3.198)	(3.234)	(1.665)
p-value of the F-test that coefficients on quarters to expansion are jointly equal to zero	0.0006	0.2899	0.2349	0.2944

Notes: *, **, *** indicate significance at the 10%, 5%, 1% level, respectively. All models include controls for age, marital status, years of education of the household head, number of household heads, hours of employment, and presence of children in the household. Additionally, all models include household, year and quarter fixed-effects, and state-specific time trends. Robust standard errors clustered at the state-level are in parentheses. The main model drops households with income between 100 and 138% of FPL and drops the DC and states with prior expansions which are, Delaware, Massachusetts, New York, and Vermont, and households with heads less than or equal to 25 years old or greater than or equal to 65 years old.

Table A.5. List Used to Generate Favorite Entrée, Vegetable, and Fruit

Favorite lunch food	**Percent of Students (%)**	**Favorite vegetable**	**Percent of Students (%)**	**Favorite fruits**	**Percent of Students (%)**
Chicken patty sandwich	12.66	Cooked sweet corn	32.06	Apple slices	20.48
Hamburger	15.99	Steamed broccoli	24.65	Canned pear slices	3.41
Cheese pizza	33.72	Steamed green beans	7.83	Fresh grapes	23.81
Macaroni and Cheese	9.08	Fresh baby carrots	13.24	Canned plums	0.33
Beef tacos	14.49	Cooked peas	2.58	Fresh red raspberries	4.91
Spaghetti with meat sauce	6.58	Fresh red peppers	2.50	Fresh strawberries	34.72
Cheese and bean enchilada	2.33	Fresh cucumber slices	15.90	Canned peach slices	6.00
Meatball sub sandwich	5.16	Cooked beets	1.25	Fresh cherries	6.33

Table A.6. Comparison of Key Demographic Averages of Final Sample (N = 1,201) against the Groups who Dropped Out of the Final Sample due to Reasons Other than not Meeting the Criteria

Variable	**Category**	**Final sample (%) N = 1,201**	**Attention Fail Test 1 or 2 (%) N = 202**	**Parent does not make decisions regarding school lunches (%) N= 169**	**Did not complete the survey (%) N = 340**
Parent's sex	Male	26.7	No Observations	No Observations	No Observations
	Female	73.2			
	Prefer not to answer	0.1			
Student's sex	Male	48.3	56.9	40.2	22.7
	Female	51.7	42.6	59.2	26.5
	Prefer not to answer	0.0	0.5	0.6	0.0
	Missing	N/A	N/A	N/A	50.9
Parent's race	African American	12.5	15.9	12.3	5.8
	Asian/Pacific Islander	3.4	4.7	6.2	2.3
	Hispanic	12.1	15.0	15.6	7.5
	Multi-Racial	1.3	1.4	0.6	1.4
	Native American	0.9	1.9	2.2	0.9
	White	69.5	61.2	61.5	28.5
	Other	0.2	0.0	1.9	0.0
	Missing				53.6
School type	Public	87.7	81.2	85.2	27.4
	Private	7.9	15.8	11.8	2.6
	Charter	4.0	3.0	2.4	1.8

Continued

Table A.6. continued

Variable	Category	Final sample (%) N = 1,201	Attention Fail Test 1 or 2 (%) N = 202	Parent does not make decisions regarding school lunches (%) N= 169	Did not complete the survey (%) N = 340
	Other	0.4	0.0	0.6	0.3
	Online/Home-schooled	N/A	N/A	N/A	0.6
Student's age (years)	13	16.3	18.3	13.0	5.6
	14	20.3	20.8	23.1	5.6
	15	20.0	16.3	20.7	9.1
	16	21.7	18.3	20.1	7.9
	17	19.2	21.8	21.3	9.1
	18	2.4	4.5	1.8	1.2
	Missing	N/A	N/A	N/A	61.5
Student's race	White	64.6	No Observations	No Observations	No Observations
	African American	13.4			
	Native American	1.2			
	Asian/Pacific Islander	3.5			
	Other	0.5			
	Multi-Racial	4.0			
	Hispanic	12.9			
School level	Primary/Elementary School	0.5	3.0	0.6	0.3
	Middle School or Junior High School	19.2	20.8	15.4	4.7
	High School	79.5	75.7	83.4	27.4

Continued

Table A.6. continued

Variable	**Category**	**Final sample (%) N = 1,201**	**Attention Fail Test 1 or 2 (%) N = 202**	**Parent does not make decisions regarding school lunches (%) N= 169**	**Did not complete the survey (%) N = 340**
	Other	0.8	0.5	0.6	0.6
	Missing	N/A	N/A	N/A	67.1
Parent's education (National average represents 25 years and over)	Some High School, no diploma	1.9	No Observations	No Observations	No Observations
	High School degree or equivalent	16.7			
	Some College, no degree	26.5			
	Associate's degree	15.5			
	Bachelor's degree	26.6			
	Graduate or Professional degree	12.8			
Household income	Less than $25,000	17.6	21.3	27.2	9.7
	$25,000 to $49,999	22.6	19.3	19.5	10.0
	$50,000 to $74,999	19	19.8	15.4	7.4
	$75,000 to $99,999	13.6	15.4	13.6	6.8
	$100,000 to $149,999	15.2	15.4	13.6	5.9
	$150,000 to 199,999	6	3.5	5.3	2.1
	$200,000 or greater	6.2	5.5	5.3	0.9
	Missing	N/A	N/A	N/A	57.4

Table A.7. Descriptive Statistics (% Unless Specified Otherwise)

Variable	Description	Sample (N= 1,201)	Population/Census
Parent's sex	Percentage of female	73.2	51.4
Student's sex	Percentage of female	51.7	
Parent's age	Years (mean and s.d.)	42.4 (8.1)	
Student's age	Years (mean and s.d.)	15.2 (1.4)	
Parent's race	African American	12.3	12.3
	Asian/Pacific Islander	3.4	5.5
	Hispanic	9.8	17.8
	Multi-Racial	1.1	2.4
	Native American	0.6	0.7
	White	72.5	61.1
Student's race	African American	12.4	14.0
	Asian/Pacific Islander	3.3	5.0
	Hispanic	9.8	24.9[a]
	Multi-Racial	3.9	6.5
	Native American	0.7	1.0
	White	69.5	66.8
Level of school	Elementary School	0.5	
	Middle School	19.2	
	High School	79.5	
	Other	0.8	

Continued

Table A.7. continued

Variable	Description	Sample (N= 1,201)	Population /Census
Household income	Less than $25,000	17.6	22.3
	$25,000 to $49,999	22.6	23.1
	$50,000 to $74,999	19	17.8
	$75,000 to $99,999	13.6	12.2
	$100,000 to $149,999	15.2	13.5
	$150,000 to 199,999	6	5.4
	$200,000 or greater	6.2	5.7
Parent's education	Some High School, no diploma	1.9	7.2
	High School degree or equivalent	16.7	27.2
	Some College, no degree	26.5	20.6
	Associate's degree	15.5	8.4
	Bachelor's degree	26.6	19.3
	Graduate or Professional degree	12.8	11.9
Lived in	Rural	24.9	
	Urban	50.1	
	Sub-urban	25	
Frequency of eating school lunch	None	8.6	
	1-3 days a week	11.7	
	3-4 days a week	20.2	
	5 days a week	58.1	
Average cost per lunch	All meals	$2.41	
	Excluding free meals	$3.39	
	Excluding free and reduced-price meals	$3.61	
Receive free or reduced-price meals	Received free school meals	25.2	

Continued

Table A.7. continued

Variable	Description	Sample (N= 1,201)	Population /Census
	Received reduced price school meals	9.7	
Changes to school lunch that would increase participation much more likely	Make items from scratch	40.1	
	Improve taste	55	

[a]Hispanic origin is asked separate from other race questions

Table A.8. Cholesky Matrix from RPL Estimates

parent-only			
	Entrée: locally produced	*Fruit: locally produced*	*Vegetable: locally produced*
Entrée: locally produced	1.7188***		
Fruit: locally produced	-1.0259**	2.2055***	
Vegetable: locally produced	-0.0825	0.1478	1.4645***
student-only			
	Entrée: locally produced	*Fruit: locally produced*	*Vegetable: locally produced*
Entrée: locally produced	1.9850***		
Fruit: locally produced	-0.3803	2.0553***	
Vegetable: locally produced	-0.5853	0.3644	1.3344***
joint			
	Entrée: locally produced	*Fruit: locally produced*	*Vegetable: locally produced*
Entrée: locally produced	1.3650***		
Fruit: locally produced	0.5288***	0.8148***	
Vegetable: locally produced	0.6745***	-0.1391	-0.1646

Table A.9. Coefficient Covariant Matrix from RPL Estimates

parent-only			
	Entrée: locally produced	*Fruit: locally produced*	*Vegetable: locally produced*
Entrée: locally produced	2.9544***		
Fruit: locally produced	-1.7634*	5.9166***	
Vegetable: locally produced	-0.1418	0.4107	2.1734***
student-only			
	Entrée: locally produced	*Fruit: locally produced*	*Vegetable: locally produced*
Entrée: locally produced	3.9404***		
Fruit: locally produced	-0.7549	4.3690***	
Vegetable: locally produced	-1.1618	0.9716	2.2560***
joint			
	Entrée: locally produced	*Fruit: locally produced*	*Vegetable: locally produced*
Entrée: locally produced	1.8631***		
Fruit: locally produced	0.7218***	0.9435***	
Vegetable: locally produced	0.9207***	0.2433*	0.5014***

Table A.10. Correlation Matrix from RPL Estimates

parent-only			
	Entrée: locally produced	*Fruit: locally produced*	*Vegetable: locally produced*
Entrée: locally produced	1	-0.4218	-0.0560
Fruit: locally produced	-0.4218	1	0.1145
Vegetable: locally produced	-0.0560	0.1145	1
student-only			
	Entrée: locally produced	*Fruit: locally produced*	*Vegetable: locally produced*
Entrée: locally produced	1	-0.1819	-0.3896
Fruit: locally produced	-0.1819	1	0.3095
Vegetable: locally produced	-0.3896	0.3095	1
joint			
	Entrée: locally produced	*Fruit: locally produced*	*Vegetable: locally produced*
Entrée: locally produced	1	0.5443	0.9526
Fruit: locally produced	0.5443	1	0.3538
Vegetable: locally produced	0.9526	0.3538	1

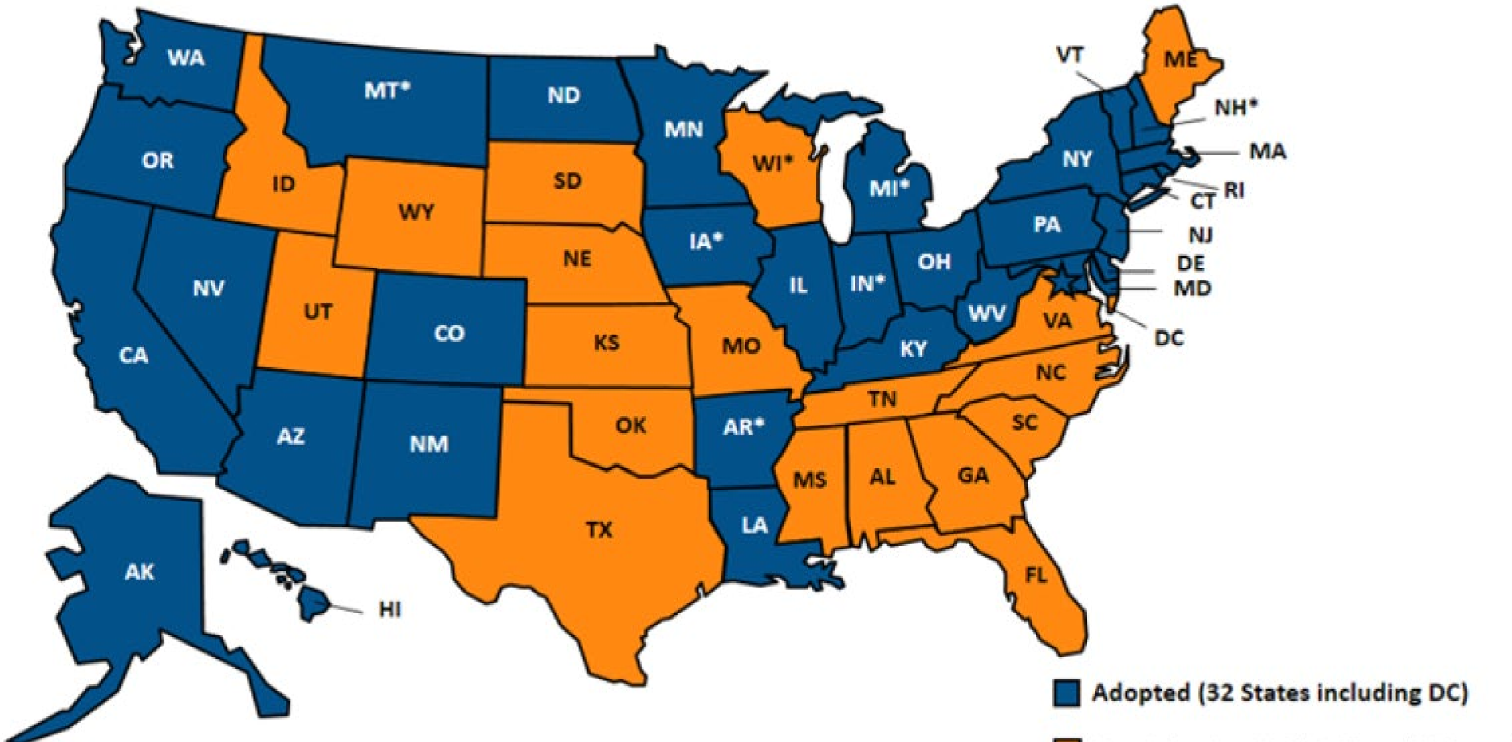


Figure A.1. State Medicaid expansion decisions as of end of 2017