Behavioral Food Subsidies*

Andy Brownback† Alex Imas‡ Michael A. Kuhn§

February 16, 2022

We conduct a field experiment with low-income shoppers to study how behavioral interventions can improve the effectiveness of healthy food subsidies. Our unique design enables us to elicit choices and deliver subsidies both *before* and *at* the point of purchase. We examine the effects of two non-restrictive changes to the choice environment: giving shoppers agency over what subsidy they receive and introducing a waiting period before the shopping trip to prompt deliberation about their purchases. Combined with healthy food subsidies, these interventions increase healthy food spending by 61% more than a healthy food subsidy alone, resulting in 199% greater healthy spending than in our control group.

KEYWORDS: nutrition, subsidies, agency, deliberation, waiting periods, field experiment
JEL Classifications: I12, D91, D12, C93

*Thanks to Marianne Bitler, Benjamin Enke, Xavier Gabaix, Tristan Gagnon-Bartsch, Ralph Hertwig, Eric Johnson, David Laibson, Matthew Rabin, Gautam Rao, Gal Zauberman, and seminar participants at Harvard University and Florida State University for valuable feedback. We also sincerely thank the Robert Wood Johnson Foundation for providing funding and valuable feedback. Many thanks to Marilyn Soukup and the team at Field Agent for their incredibly hard work on the program implementation. We also received excellent research assistance from Mohammed Ibrahim. Pre-registration for the project is at AsPredicted: https://aspredicted.org/5f3kc.pdf. University of Arkansas Internal Review Board approval for this project was obtained prior to any research being conducted.

†University of Arkansas, Department of Economics. E-mail: abrownback@walton.uark.edu. Web page: www.andybrownback.com

‡University of Chicago, Booth School of Business, and CESifo. E-mail: aimas@uchicago.edu . Web page: www.aleximas.com

§University of Oregon, Department of Economics, and CEGA. E-mail: mkuhn@uoregon.edu. Web page: www.makuhn.net

1 Introduction

A growing literature has documented extensive socioeconomic inequality both in developing (Alvaredo, Assouad, and Piketty 2019) and developed countries (Piketty and Saez 2014). Inequality in nutrition has received particular attention from both scientists and policy makers because dietary differences are viewed as both a downstream consequence of economic inequality (Drewnowski and Specter 2004), and as a contributor to its persistence (Wolf 2012). Poor diets are now 25% more prevalent amongst low-income Americans than higher-income Americans (Rehm, Peñalvo, Afshin, and Mozaffarian 2016). While government programs such as the Supplemental Nutrition Assistance Program (SNAP) do generate beneficial long-run health consequences (Hoynes, Schanzenbach, and Almond 2016), their focus has largely been on food security rather than nutritional quality. Diets of SNAP enrollees are persistently poor—fruit and vegetable consumption among program participants is approximately half of the recommended intake, while higher-income non-participants consume over two-thirds of the recommendation (Cole and Fox 2008).[1]

We consider how insights from behavioral economics can be leveraged to enhance the effectiveness of healthy food subsides in order to close this nutrition gap. There are two reasons we focus on subsidies. First, recent research has shown that supply-side interventions aiming to increase access to healthy food have not closed the nutrition gap (Allcott, Diamond, Dubé, Handbury, Rahkovksy, and Schnell 2019a), suggesting that differences in demand could be driving dietary inequality. Second, the consumption of unhealthy food features both externalities and internalities. In the context of energy consumption, Allcott, Mullainathan, and Taubinsky (2014) show that optimal policy in these cases relies on both taxes and subsidies to achieve the social-planner's first-best solution. While there is substantial research on the role of taxation of unhealthy foods (see Allcott, Lockwood, and Taubinsky (2019c) for an overview of soda taxes), less is known about the potential role of subsidies for healthy food.

[1]Data are from the 1999-2004 National Health and Nutrition Examination Survey.

Healthy food subsidies are restrictive in that they limit discounts to a particular set of food (e.g., fruits and vegetables).[2] Such subsidy restrictions are a natural extension of recent proposals to improve nutrition through outright restrictions on food assistance spending (Schwartz 2017).[3] We are aware of two studies on such *restricted* subsidies, and both have found positive effects (Bartlett, Klerman, Wilde, Olsho, Logan, Blocklin, Beauregard, and Enver 2014; Harnack, Oakes, Elbel, Beatty, Rydell, and French 2016). Of particular note is the Healthy Incentives Pilot (HIP), a large-scale SNAP experiment found that restricted subsidies significantly increased purchases and consumption of fruits and vegetables (Klerman, Bartlett, Wilde, and Olsho 2014; Olsho, Klerman, Wilde, and Bartlett 2016; Wilde, Klerman, Olsho, and Bartlett 2015).[4]

However, research in behavioral science suggests that there are several psychological channels that either blunt the impact of restrictive healthy food subsidies or can be leveraged to increase their effectiveness. For example, Read and van Leeuwen (1998) demonstrate how a failure to account for myopia can prevent even those individuals who want to improve their diets from doing so. This paper demonstrates how non-restrictive changes to the choice environment can substantially increase the effectiveness of healthy food policy. We find that combining subsidies with behavioral interventions leads to a 199% increase ($p < 0.001$) in healthy food purchases relative to an unsubsidized control condition. The impact of these "behavioral food subsidies" is 61% larger ($p = 0.004$) than the impact of healthy food subsidies alone, which represent a 124% increase ($p < 0.001$) in healthy food purchases relative to an unsubsidized control.[5]

[2]Compared to other programs, such restrictions are similar to the spending limitations placed in the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) program, while the SNAP program has favored a less restrictive approach.

[3]This is motivated by the relatively low nutritional quality of SNAP purchases. For example, soda takes up the largest share and approximately 20% of SNAP spending goes towards sweets and sugary drinks. Efforts to place restrictions on SNAP purchases include recent proposals to offer in-kind SNAP benefits in the form of USDA Foods Packages (https://www.npr.org/sections/thesalt/2018/02/12/585130274/trump-administration-wants-to-decide-what-food-snap-recipients-will-get) or to restrict SNAP purchases to nutritious foods (https://www.brookings.edu/testimonies/pros-and-cons-of-restricting-snap-purchases/).

[4]Herman, Harrision, Afifi, and Jenks (2008) finds similar effects of healthy food vouchers among WIC participants.

[5]Estimates are from a random effects regression of fruit and vegetable spending during our study on

We conducted a field experiment with low-income grocery shoppers where our outcome of interest is healthy food spending. We partnered with a smartphone-based market-research platform (Field Agent) to obtain repeated access to shoppers across the United States before, during, and after grocery shopping trips. Through the platform, shoppers were able to conduct their regular shopping trips at the time and location they desired, while we induced experimental variation into their budget sets and grocery shopping experience. This unique setting allowed us to elicit subsidy choices from shoppers through their smartphones. Additionally, we are able to push information to shoppers' smartphones both before shopping trips and at the point of purchase—in the aisle of the grocery store. This allows us to manipulate the timing between subsidy-related information and the actual shopping trip. The Field Agent platform collected pictures of shopping receipts from all shoppers immediately following their shopping trip in order to track actual purchases and calculate subsidy payments.

We first corroborate previous findings showing that restricted subsidies are effective in spurring healthy food spending. Shoppers who were randomly assigned to receive a restricted "healthy" subsidy for 30% off of fruits and vegetables ("FV" hereafter) spent 123% more on FV purchases than an un-subsidized control group.[6] The focus of this paper is on the impact of costless, behavioral interventions that aim to enhance these subsidies by overcoming psychological factors which may be blunting their effectiveness.

Our first behavioral intervention manipulates the subsidy choice environment by endowing shoppers with greater agency over their subsidies.[7] In our experiment, shop-

the complete set of treatment-group indicator variables, controlling for an individual's pre-study fruit and vegetable spending, and including recruitment wave fixed effects. Standard errors are clustered at the individual level. We take a variety of alternative approaches to analyzing the data in Section 4.

[6] It is possible that the subsidies induced shoppers to concentrate their purchases into fewer trips to maximize their subsidy payments. It is also possible that the novelty of the subsidies amplified their effectiveness. Both mechanisms would inflate the estimated levels of the restricted subsidy effect alone. However, this will not affect inference about the relative effectiveness of the behavioral interventions—which is the focus of the paper—because all interventions feature subsidies. Specifically, the incentive to concentrate shopping trips will be equally present across interventions.

[7] We do not mean to imply that shoppers in our setting have no agency *per se*. Rather, we argue that relative agency constraints imposed by policy, such as in the case of aid that is restricted to particular subsets of food, can be perceived by recipients as restricting their individual agency. Our intervention aims to examine the behavioral effect of relaxing this constraint.

pers were given the choice between the healthy FV subsidy and an "unhealthy" subsidy for 30% off of baked goods ("BG" hereafter). We elicited binding subsidy choices from shoppers through their smartphones at different points in time—either before shopping or during the shopping trip. We then compare their shopping behavior to subjects restricted to the healthy subsidy. In the standard economic framework, this expansion of subsidy choice should increase the average prices of FV relative to BG and result in *decreased* FV spending. However, work in psychology and behavioral economics documents how an active choice can generate a preference for consistency (Bénabou and Tirole 2004; Falk and Zimmermann 2017, 2018). Specifically, a prior choice, such as selecting the healthy subsidy, can prompt congruent behavior, such as purchasing healthy food. This effect is a common feature of several behavioral frameworks that build on the self-perception model of Bem (1967). These include the information-processing model of Falk and Zimmermann (2018) and the self-signaling model of Bénabou and Tirole (2004). In our setting, 83% of shoppers report a desire to increase FV consumption and the majority of the sample (78%) chooses the FV subsidy. As a result, a preference for consistency predicts that prompting an active subsidy choice can lead to an increase in FV spending.[8] Additionally, recent research has shown that endowing people with greater agency mitigates myopia and encourages more goal-oriented behavior, which should further increase demand for healthy food (Gneezy, Imas, and Jaroszewicz 2020).

In our context, these behavioral mechanisms predict that being able to *choose* a healthy subsidy—rather than being forced into one—will prompt more congruent, patient decisions. The results reveal that the positive behavioral effects of agency are strong enough to counteract the negative price effect of expanding subsidy choice.

[8]Importantly, an active subsidy choice allows shoppers to select a healthy subsidy as a means of sending themselves a *costly* signal of their intended shopping behavior—using their own choices to motivate subsequent behavior. In Section 4.3, we present evidence that shoppers pursue this strategy: 28% of shoppers leave money on the table by choosing the healthy subsidy. We show that this behavior cannot be explained by random errors or confusion; rather, it is consistent with self-signaling (Bénabou and Tirole 2004). If the majority chose the unhealthy subsidy, then these predictions would not hold.

Nearly 78% of shoppers choose the healthy subsidy, meaning that the expanded choice set increased the relative price of FV compared to the restricted case. However, the overall effect of agency is a 21% *increase* in FV spending. While the marginal effect is not statistically significant on its own, we can reject any meaningfully-sized negative effects of expanding subsidy choice on FV spending.

Poor nutritional choices are frequently thought of as a specific manifestation of myopic or present-focused preferences (Laibson 1997; O'Donoghue and Rabin 1999), potentially resulting from Pavlovian responses to unhealthy foods (Bushong, King, Camerer, and Rangel 2010). Previous research into food choices has focused on pre-commitment as a means of overcoming failures of self-control and the resulting time inconsistency (Allcott, Lockwood, and Taubinsky 2019b; Read and van Leeuwen 1998; Sadoff, Samek, and Sprenger 2020).[9] Our focus on agency is conceptually distinct from this line of work. In our setting, the most "committed" group are those in the restricted healthy subsidy treatment. The agency condition thus tests the impact of allowing shoppers the opportunity to *relax* this commitment by choosing an unhealthy subsidy. Those who select the healthy subsidy face identical prices as under the restricted subsidy, but have made an *active choice*. By *expanding* the choice set, our agency intervention circumvents the welfare tradeoffs between short-run and long-run selves that arise from limitations on choice in standard commitment contracts.[10]

Our second behavioral intervention introduces a waiting period into the subsidy choice environment. In our experiment, we randomly assigned shoppers to complete a pre-shopping task 4-48 hours prior to their next shopping trip. The unique smart-

[9] Laibson (2015) and Carrera, Royer, Stehr, Sydnor, and Taubinsky (2019) argue that evidence from Read and van Leeuwen (1998) and Sadoff et al. (2020) may be examples of uncertainty about future tastes rather than time-inconsistency due to failures of self-control. Allcott et al. (2019b) attempt to avoid this issue by showing that survey measures of self control predict spending on sugar-sweetened beverages.

[10] A rich literature has examined commitment opportunities where consumers select a seemingly dominated option as a means of restricting their choice set. See, for example, Ariely and Wertenbroch (2002); DellaVigna and Malmendier (2006); Kaur, Kremer, and Mullainathan (2010, 2015); Sadoff et al. (2020); Schilbach (2019); Schwartz, Mochon, Wyper, Maroba, Patel, and Ariely (2014). Take-up of such contracts, however, is not always straightforward evidence for time-inconsistency as illustrated in Carrera et al. (2019).

phone platform allowed this pre-shopping task to enforce a delay between the arrival of information about the subsidy and the shopping trip on which the subsidy would be used.[11] Laboratory experiments have demonstrated that waiting periods prompt more deliberative processing, leading to less myopic and more goal-oriented decision-making (Dai and Fishbach 2013; DeJarnette 2018; Imas et al. 2018).[12] This intervention is distinct from traditional healthy-eating interventions in the nutrition and behavioral literatures for two main reasons. First, it provides no new information that could alter the costs and benefits of food purchases. Second, it does nothing to separate grocery choices from consumption; rather, it maintains the timing of decisions. Nonetheless, we find that the marginal effect of a waiting period is a 29% increase in healthy food purchases relative to a subsidy without a waiting period.

While the effects of waiting periods are not predicted by the workhorse β-δ model of present bias (O'Donoghue and Rabin 1999), they are consistent with the Gabaix and Laibson (2017) framework that models myopia as being generated by imperfect forecasts of future utility. In their setting, people are uncertain about the utility derived from potential future outcomes. By prompting deliberation, waiting periods reduce noise when forecasting utility. This process leads to choices that are less myopic and more in-line with a person's underlying goals such as healthier eating.[13]

We further explored the role of deliberation by varying when shoppers made the subsidy choice: at the start of the waiting period ("Early Choice") or at the end of the waiting period at the point of purchase ("Delayed Choice"). The framework of Gabaix and Laibson (2017) predicts that Early Choice can further increase the efficacy of

[11] As in Imas, Kuhn, and Mironova (2018), we define a waiting period as a delay between information about a prospective choice and the choice itself. This information need not be new—as it is at the outset of the study—as deliberation can still facilitate the use of old information.

[12] Outside of the lab, there is related non-experimental research. Koenig and Schindler (2018) and Luca, Malhotra, and Poliquin (2017) use observational data to explore the effects of waiting periods for gun ownership on violence, and Burlando, Kuhn, and Prina (2020) uses administrative lender data to estimate the effect of loan delivery speed on credit repayment. While focused on anticipation rather than deliberation effects, Thakral and To (2020a,b) find that households receiving cash transfers save more and spend less on temptation goods when the delivery delay is longer.

[13] Section 3 considers how other models, such as goal-setting, adaptive reference points, and preference adaptation may relate to waiting periods.

waiting periods by narrowing the set of potential future outcomes over which one must deliberate. This leads to more focused deliberation and further decreases myopia.[14] Indeed, we find that despite no differences in the healthy subsidy choice rates across all treatments, shoppers in the Early Choice treatment have the highest FV spending—over 60% more compared to the restricted healthy subsidy and 20% more than those in the Delayed Choice condition.

Together, our interventions increase healthy food spending by more than 60% relative to the restricted healthy subsidy. This increase comes at no additional marginal cost to the policy maker apart from the subsidy itself. Overall, we find that coupling food subsidies with these behavioral interventions *triples* observed purchases of healthy food, increasing FV spending by 199% relative to a control group.

To assess the welfare consequences of our interventions, we follow the behavioral welfare framework of Bernheim and Taubinsky (2018) to assess whether we have over-corrected for any biases that led to under-consumption of healthy food in the first place.[15] In Section 5.2 we compare observed healthy food consumption to official dietary recommendations to show that our interventions are not sufficiently large to raise concerns that the marginal cost of our subsidies exceed the marginal benefits to our shoppers. This suggests that the observed increases in FV consumption are welfare-improving.

Our findings contribute to a large literature that uses insights from psychology and behavioral economics to develop interventions in policy-relevant domains. Ex-

[14]Note that the Early Choice treatment can also change purchase behavior by affecting the subsidy choice. While most grocery shopping is for future consumption, myopia and present-bias may tempt shoppers to select the unhealthy subsidy if it immediately precedes the purchase decision. Indeed Houser, Liu, Reiley, and Urbancic (2021) show that the grocery store can be a tempting environment for unhealthy food choice. Early Choice may allow shoppers to overcome this by separating the purchase decision from the subsidy choice, permitting them to "pre-commit" to the healthy subsidy. As discussed in Section 4, we find no evidence for this type of dynamic inconsistency in subsidy choice.

[15]Note that we chose to use healthy food *spending* as our outcome of interest because it is a verifiable data source and serves as a good proxy for healthy food *consumption*. In sections that follow, we use consumption data from our baseline, pre-intervention, and endline, post-intervention, to verify that increases in healthy spending are highly predictive of increases in healthy consumption.

amples include employee productivity (Gosnell, List, and Metcalfe 2020), resource management (Hahn, Metcalfe, Novgorodsky, and Price 2016), education (Brownback and Sadoff 2019; Levitt, List, Neckermann, and Sadoff 2016), health (Volpp, John, Troxel, Norton, Fassbender, and Loewenstein 2008), and tax collection (Hallsworth, List, Metcalfe, and Vlaev 2017). Our investigation of how subsidies can be used to overcome behavioral biases fits into the growing public economics literature on sin taxation with behavioral agents that experience "internalities" from their biases (Allcott et al. 2019b,c; Farhi and Gabaix 2020; Lockwood 2020). We add to this line of work by designing and implementing novel interventions in a policy-relevant domain: subsidizing healthy food purchases. Our interventions are unique in that they impose no additional pecuniary costs (relative to restricted healthy subsidies) and can be easily incorporated into existing delivery mechanisms.

Our results also contribute to the research on using non-pecuniary interventions to spur behavior change (Grüne-Yanoff and Hertwig 2016; Johnson, Shu, Dellaert, Fox, Goldstein, Häubl, Larrick, Payne, Peters, Schkade et al. 2012; Jung and Mellers 2016; Thaler and Sunstein 2009). We add to this work by demonstrating that combining non-pecuniary behavioral interventions with standard incentives substantially increases the effect of the latter. To the best of our knowledge, ours is the first paper to examine how increased agency and the introduction of waiting periods (both together and apart) can improve the effectiveness of a pecuniary incentive program. Additionally, as with a "nudge" (Thaler and Sunstein 2009), our approach is less paternalistic than an outright purchase restriction.

The paper proceeds as follows. Section 2 describes the experimental design, including details about the mobile platform and procedural details. Section 3 outlines our hypotheses. Section 4 presents the results. Section 5 considers the welfare consequences of our interventions. We discuss our findings in Section 6 and conclude.

2 Experimental Design

The details of our experimental design and analysis were pre-registered on AsPredicted.org.[16]

Our experiment was conducted on a smartphone-based market-research platform called Field Agent. This platform is designed for crowd-sourced consumer research. Shoppers nationwide can find paid "jobs" to complete through their smartphones while conducting their typical shopping trips. This format allowed for unique experimental protocols that facilitated the study of shopping behavior in a natural environment. All of our shoppers across all treatments were able to conduct their shopping trips at whatever time and place they wanted without prior approval from us. Upon completion of a shopping trip and submission of all documents, Field Agent confirmed that all protocols were followed according to the assigned treatment.

There were three primary features of Field Agent that permitted our unique experimental design. First, Field Agent geo-tagged all responses and cross-referenced these location tags with the locations of grocery stores across the U.S. to guarantee that our shoppers were at the point of purchase while completing any of their in-store tasks. Second, Field Agent had an established electronic payments protocol with their users. This gave our experimental procedures credibility with the shoppers and guaranteed prompt transfers of subsidies and experimental payments. Finally, Field Agent collected pictures of all grocery shopping receipts uploaded from shoppers' smartphones. These receipts were tabulated by workers on Amazon's Mechanical Turk website who were blind to the experiment. These workers viewed the receipt photos while tabulating overall spending, spending in different food categories, and Shopping Trip characteristics such as EBT use. This procedure was identical across treatments during the

[16] We pre-registered the data collection and analysis for the primary treatments we focus on in this paper (Part 1). After conducting this experiment, we ran a follow up study to separately examine the impact of a waiting period without agency (Part 2). Treatments associated with Part 2 will be highlighted later in this section. We analyze these data in accordance with our pre-registered analysis plan with fixed effects for the data collection wave that capture any fixed differences between Part 1 and Part 2. Our pre-registration includes individual-level analysis of the relationship between time preferences on shopping and consumption behavior that is forthcoming in a companion paper.

study and for both baseline and endline food spending measurements. Using these tabulated receipts, we calculated subsidy payments and measured treatment effects. The receipt timestamps also verified that shoppers completed their tasks prior to finalizing their purchases. According to Field Agent's internal protocol, shoppers who violated the procedures—either timing or location—were first given a warning, and upon second-offense eliminated from the study. The experiment had no influence over this process.

To maintain a natural shopping environment, we asked all shoppers to continue with their normal shopping schedule—shopping at their regular store at their regular time. Our treatments then varied the delivery of food subsidies to shoppers' smartphones during these trips. The study involved "healthy" subsidies offering 30% off purchases of fruits and vegetables (FV) and "unhealthy" subsidies offering 30% off purchases of baked goods (BG). Both subsidies were capped at $10 per trip.[17] According to Consumer Expenditure Survey (CES) data, these categories represent roughly equal percentages of food purchases by SNAP recipients and offer a clear valence to define a "healthy" and "unhealthy" option.[18]

2.1 Recruitment

Field Agent has over 1 million registered users on its smartphone-based shopping survey platform in the U.S. alone.[19] We restricted our population to the subset of recently active users and recruited shoppers with a gross household income less than 185% of the federal poverty line (FPL) to take part in our study.[20] During recruitment, we

[17] Fruits and vegetables include fresh, canned, or frozen fruits or vegetables without added salt or sugar. Baked goods include bread, biscuits and rolls, muffins, cakes and cupcakes, pies and tarts.

[18] CES data span 1994-2003, limited to households that self-report as SNAP participants. FV and BG represent 8% and 9% of total food expenditure, respectively, and are purchased on 76% and 82% of days with at least $20 in food spending, respectively. In our control group, shoppers spend more on BG than FV, so *ex-ante*, the BG subsidy is more valuable.

[19] Smartphone ownership is a common pre-condition for studying behavior of low-income participants (e.g., Smith, Morgan, Plotnikoff, Dally, Salmon, Okely, Finn, and Lubans (2014)) since the majority meet the requirement. For example, over 67% of Americans with incomes less than $30,000 own smartphones (Pew Research Center: www.pewinternet.org/fact-sheet/mobile).

[20] Gross income of 185% of the FPL is the federal maximum for households to meet the income qualification for WIC subsidies. SNAP eligibility begins beneath 130% of the FPL. U.S. Department

told shoppers that, "the goal of [the] study is to understand your grocery shopping habits." We made no reference to healthy shopping or subsidies until after we revealed treatment assignment during the first shopping trip. Qualified shoppers were then invited to complete our baseline survey.

We conducted our study in two parts. In Part 1, we collected data from four treatments and a control group. In Part 2, we repeated one of the treatments from Part 1 and included a new treatment. In both Parts 1 and 2, we targeted approximately 150 shoppers for all but two treatments, which were given a slightly greater weight to increase statistical power. Our analysis will include fixed effects for Parts 1 and 2 to account for potentially non-random variation in behavior between the two.

For Part 1, we recruited shoppers in eight separate waves beginning in March 2018. The final endline surveys were completed by July 2018. Shoppers were randomized within each wave and allowed to complete the study at their own pace. This randomization will be discussed in more detail in Section 2.3. For Part 2, we recruited all shoppers in a single wave beginning in March 2019. The final endline surveys were completed by April 2019. The randomization occurred once at the beginning of March. All shoppers were given approximately eight weeks to complete four shopping trips. Within those eight weeks, shoppers could complete the study at their own pace.

2.2 Treatments

All shoppers who completed the baseline in either Part 1 or Part 2 were assigned to a treatment that was fixed throughout the study, generating a fully between-subjects design. Shoppers in Part 1 were randomly assigned to one of **Control, Restricted, Agency, Waiting Period (Delayed Choice), Waiting Period (Early Choice)**. Shoppers in Part 2 were randomly assigned to either **Restricted** or **Waiting Period (No Agency)**. All treatment instructions can be found in Appendix Section B. Our six treatments are:

of Agriculture Food and Nutrition Service: www.fns.usda.gov/wic/wic-eligibility-requirements, www.fns.usda.gov/snap/recipient/eligibility.

- C: **Control** — Shoppers submitted photographs of their receipts, but received no subsidies.
- T_1: **Restricted** — All shoppers received the healthy subsidy. Subsidy information was delivered in the store before the purchase decision.
- T_2: **Agency** — Shoppers chose between the healthy and unhealthy subsidies. Subsidy information was delivered in the store. Shoppers made their subsidy choice in the store before the purchase decision.
- T_3: **Waiting Period (Delayed Choice)** — Shoppers chose between the healthy and unhealthy subsidies. Subsidy information was delivered 4 to 48 hours before shopping. Shoppers made their subsidy choice in the store before the purchase decision.
- T_4: **Waiting Period (Early Choice)** — Shoppers chose between the healthy and unhealthy subsidies. Subsidy information was delivered and shoppers made their subsidy choice between 4 and 48 hours before shopping. Shoppers were asked to recall their subsidy choice in the store before the purchase decision.
- T_5: **Waiting Period (No Agency)** — All shoppers received the healthy subsidy. Subsidy information was delivered 4 to 48 hours before shopping. They were reminded about the subsidy in the store before the purchase decision.

Shoppers had the opportunity to complete four separate "Shopping Trips" under their assigned treatment. The Shopping Trips were designed to be as natural as possible with their timing and location decided by the shopper.[21] Our only requirement was that the Shopping Trips be at least five days apart. After completing all four Shopping Trips, shoppers concluded the study with an endline survey measuring many of the same characteristics as the baseline survey. In addition to any subsidy or time-preference elicitation payments, shoppers were paid a flat fee of $1 per completed survey with a bonus payment that guaranteed $30 total if they completed all of the surveys in the study.

Upon assignment, shoppers learned the procedures for their treatment. Shoppers followed their assigned procedures and submitted surveys and receipts to Field Agent

[21] In the endline survey, we asked shoppers if they changed their shopping schedule to accommodate the study. 62% said they integrated the study into their normal routine, and this response was balanced across treatments.

for verification. Any Shopping Trips or survey submissions that failed the verification checks were not reimbursed and our partners contacted the shoppers to explain the failure and how to correct it on future Shopping Trips. Our partners were blind to the treatments and conducted these checks without any input from the researchers.

There are three types of tasks associated with Shopping Trips—Pre-Shopping Tasks, Shopping Tasks, and Post-Shopping Tasks. The timing of the tasks depended on the assigned treatment. Table 1 clarifies the different tasks and timing for each treatment.

Table 1. Shopping Trip Timeline by Treatment

	4 to 48 hours before grocery shopping *"Pre-Shopping Task"*	While shopping (in-store before purchase) *"Shopping Task"*	In-store after purchase *"Post-Shopping Task"*
C: Control			Submit receipt for participation payment
T_1: Restricted		Subsidy information delivered	Submit receipt for participation payment and shopping reimbursement
T_2: Agency		Subsidy information delivered and subsidy choice made	
T_3: Waiting Period (Delayed Choice)	Subsidy information delivered	Subsidy reminder and subsidy choice made	
T_4: Waiting Period (Early Choice)	Subsidy information delivered and subsidy choice made	Subsidy choice reminder	
T_5: Waiting Period (No Agency)	Subsidy information delivered	Subsidy reminder	

For shoppers assigned to a Waiting Period treatment, the Shopping Trip began with a Pre-Shopping Task. Any shopper assigned a Pre-Shopping Task needed to complete this task 4 to 48 hours prior to shopping. This window enforces a minimum waiting period on the shopper while also remaining flexible enough to limit its burden. As with all aspects of our study, this was completed at the shopper's convenience. We encouraged shoppers to complete their Pre-Shopping Tasks at a time that allowed them to continue their normal shopping patterns within the specified 4- to 48-hour window. Failure to complete the Pre-Shopping task during this window would invalidate a submission and the shopper would be not eligible for the subsidy. In this event, they were asked to complete their trip as usual, and were then given one chance to complete a "make-up Pre-Shopping Task" before a later Shopping Trip. In our analysis, per our pre-registration, we include all data collected during these make-up trips.

Shoppers assigned to the Waiting Period (Early Choice) treatment made their selection between the healthy and unhealthy subsidy during the Pre-Shopping Task. Shoppers assigned to the Waiting Period (Delayed Choice) treatment received information about the subsidy choice they would later make as a part of their Shopping Task on the upcoming Shopping Trip. Finally, Shoppers assigned to the Waiting Period (No Agency) treatment simply received information about the healthy subsidy they were assigned to receive on their upcoming Shopping Trip.

Shoppers in every treatment but Control were required to complete a Shopping Task inside the grocery store prior to finalizing the grocery purchases. This ensured similar levels of attention and engagement across the subsidized treatments. During the Shopping Task, shoppers in the Agency and Waiting Period (Delayed Choice) treatments made a selection between the healthy and unhealthy subsidies. For shoppers in the Restricted and Waiting Period (No Agency) treatments, the Shopping Task simply delivered information about the healthy subsidy they were assigned to receive. Finally, shoppers in the Waiting Period (Early Choice) treatment were reminded of the subsidy choice they made on the Pre-Shopping Task. Field Agent used the timestamp and geo-tag from the completed Shopping Task to determine that the task was completed at the grocery store and prior to finalizing grocery purchases.[22]

Shopping Trips in every treatment finished with a Post-Shopping Task. This task required the shopper to submit a picture of their shopping receipt. For any treatment with a subsidy, this receipt determined their subsidy payment. For the control group, this finalized their Shopping Trip and ensured their participation payment.

Differences in the Post-Shopping Task highlight one complication in estimating the impact of the subsidies: we cannot observe every grocery shopping trip that a shopper makes. A specific concern is that our subsidies may incentivize shoppers to concentrate their shopping into fewer trips to maximize their subsidy payments, or that the

[22]We cannot confirm that the Shopping Tasks were completed at the beginning of the Shopping Trip, but we believe that any delay between starting the trip and completing the Shopping Task should bias our estimated treatment effects downward because a greater share of the shopping would have been completed prior to our intervention.

introduction of a waiting period may change the types of shopping trips we observe. For example, shoppers in Control may be more likely to conduct FV shopping on an extra-experimental shopping trip because they are not sacrificing subsidy payments by doing so. This biases the estimated level effects of the subsidies upward. However, this concern does not apply to the primary focus of this paper: estimating the differential impacts of the behavioral interventions that are paired with subsidies, which are measured against the Restricted healthy subsidy.[23] On the other hand, shoppers with agency can select out of the FV subsidy. If these shoppers prefer to buy FV outside of the study, this could cause a *downward* bias in our estimated impact on FV spending. As shown below, the subsidy choice rates in the Waiting Period treatments were nearly identical to the Agency treatment. Because the interventions did not affect the relative price of FV, they therefore do not affect incentives for extra-experimental shopping trips, which we believe limits this concern.

2.3 Randomization

As each recruitment wave was completed, we conducted a stratified randomization with separate strata for 1) SNAP participation and 2) a stated desire to improve fruit and vegetable consumption.[24] The proportion of shoppers assigned to each treatment was similar but not constant for each recruitment wave, so our analysis will include fixed effects for the recruitment wave.[25]

Shoppers were informed of their randomly-assigned treatment along with the respective procedures when they opened their first task for the first Shopping Trip after

[23] Subsidies may also distort the timing of Shopping Trips. For example, they may incentivize longer intervals between Shopping Trips. However, in Section A.8, we show our treatments have no detectable effect on the intervals between Shopping Trips. Another potential timing bias comes from the 4-48 hour waiting periods rules. In Section A.8 we show that the timing of shopping trips in the Waiting Period treatments is similar to that of shopping trips in the other treatments.

[24] We could not stratify based on baseline FV purchases as the receipts took too long to be tabulated.

[25] After Waves 1-6 of Study 1, we increased the weight on the Agency and Waiting Period treatments in our random assignment in order to improve power for testing our main interventions of interest. We did this in response to the lower than expected completion rates between the Baseline survey and the first Shopping Trip, and prior to running any analysis of our dependent variable: FV spending. To address any concerns about the post re-balancing results, we replicate all of our main findings using only the waves prior to re-weighting in Appendix A.4

the baseline. Thus, while differential attrition is a potential concern after the first Shopping Trip, differential selection at the time of assignment is not.

2.4 Shopper Characteristics

In Part 1, 802 shoppers successfully completed the baseline survey. Treatment assignment was balanced on observables, which are reported in the left side of Table 2. In Part 2, 300 shoppers completed the baseline survey. Treatment assignment was again balanced on observables. This balance is recorded in the right side of Table 2.

Table 2. Balance of Shopper Characteristics by Treatment

	Part 1: Mar–Jul, 2018						Part 2: Mar–Apr, 2019		
	Control	Restricted	Agency	Waiting Period (Delayed)	Waiting Period (Early)	F-Test	Restricted	Waiting Period (No Agency)	F-Test
Reported Dietary Satisfaction									
Want more fruits & vegetables (1-3)	1.809	1.859	1.815	1.805	1.793	0.602	1.839	1.821	0.691
	(0.036)	(0.029)	(0.029)	(0.029)	(0.035)		(0.030)	(0.033)	
Have enough food to eat (1-4)	2.319	2.303	2.265	2.216	2.310	0.642	2.195	2.159	0.667
	(0.059)	(0.055)	(0.052)	(0.053)	(0.050)		(0.059)	(0.058)	
Can afford food (1-5)	2.908	2.923	2.741	2.719	2.841	0.374	2.879	2.715	0.206
	(0.099)	(0.096)	(0.084)	(0.086)	(0.098)		(0.088)	(0.095)	
Can afford fruits & vegetables (1-5)	2.433	2.408	2.370	2.286	2.483	0.693	2.336	2.113	0.141
	(0.106)	(0.110)	(0.092)	(0.092)	(0.108)		(0.107)	(0.107)	
Shopper Characteristics									
SNAP participant	0.397	0.401	0.407	0.443	0.400	0.906	0.349	0.351	0.971
	(0.041)	(0.041)	(0.036)	(0.037)	(0.041)		(0.039)	(0.039)	
Male	0.184	0.211	0.153	0.151	0.152	0.590	0.154	0.166	0.792
	(0.033)	(0.034)	(0.026)	(0.026)	(0.030)		(0.030)	(0.030)	
Household Size	3.596	3.662	3.635	3.703	3.586	0.971	3.732	3.662	0.743
	(0.147)	(0.137)	(0.129)	(0.122)	(0.143)		(0.147)	(0.149)	
Time discounting (standardized)	-0.081	-0.022	-0.101	-0.088	0.005	0.851	0.021	-0.096	0.308
	(0.083)	(0.085)	(0.068)	(0.071)	(0.084)		(0.084)	(0.078)	
Baseline Grocery Receipts									
Fruits & vegetable purchases ($)	6.114	4.249	5.663	4.378	4.458	0.255	6.263	6.393	0.891
	(0.797)	(0.677)	(0.911)	(0.514)	(0.493)		(0.609)	(0.717)	
Baked goods purchases ($)	11.025	13.486	12.367	10.580	7.666	0.335	1.937	1.520	0.221
	(1.805)	(2.468)	(2.272)	(1.725)	(1.840)		(0.257)	(0.223)	
Use EBT card on baseline purchase	0.468	0.528	0.587	0.605	0.549	0.116	0.242	0.212	0.541
	(0.042)	(0.042)	(0.036)	(0.036)	(0.042)		(0.035)	(0.033)	
Observations	141	142	189	185	145		149	151	

F-test conducted as a joint test of equality across all treatments (robust standard errors), with the p-values reported.

Our baseline data show that the majority of our sample experiences both food insecurity and a desire to improve the nutritional quality of their diet: 67% reported some measure of food insecurity and 83% reported that they would like to consume more fruits and vegetables.[26]

Our shoppers were geographically diverse, coming from across the United States. 40% of our sample reported participating in SNAP. The vast majority of our shoppers (85%) were female. While this is not representative of the overall population (or the SNAP primary-shopper population), the skew is representative of the gender disparity in grocery shopping that has been found by others (Bhattarai 2017). The majority of shoppers (53%) reported living in a city with fewer than 50,000 residents. Compared to the SNAP population, our average shopper also has a larger family with more children.[27]

2.5 Attrition

Attrition is a concern for any field experiment involving repeated observations. Ghanem, Hirshleifer, and Ortiz-Becerra (2021) provide an overview of attrition in field experiments as well as a useful framework for understanding when attrition may prevent identification of the local average treatment effect (LATE). Specifically, attrition prevents identification of the LATE if it is differential and selective based on outcomes of interest. If attrition is differential but not associated with our outcome of interest, then identification of the LATE-R—the LATE for the sub-population of respondents—is still possible. Since our outcome of interest is FV spending, it is important that any attrition not be associated with the levels of FV spending observed on the baseline survey. Such an association could indicate that high or low levels of demand for FV spending lead shoppers to leave the study or remain—suggesting selection-bias.[28]

[26] For comparison, Bartlett et al. (2014) surveyed SNAP to study the USDA's Healthy Incentives Pilot. The closest corresponding question asked if respondents encourage others to eat more fruits and vegetables. On that question, the average respondent slightly agreed (3.8 on a scale of 1-5).

[27] Based on 2018 SNAP Quality Control data.

[28] This is equivalent to a direct test of the assumptions about selective attrition made in common bounding exercises such as Lee (2009).

Table 3 quantifies the differences in attrition across the treatments and evaluates the criteria for identifying the LATE and the LATE-R. Each panel presents a different margin of possible attrition. For each margin, we test for i) differential attrition—differences across treatment groups, ii) selective attrition—associations between attrition and baseline FV spending, and iii) differentially selective attrition—differences across treatments in the associations between attrition and baseline FV spending.

In our study, shoppers learned their treatment assignment at the beginning of their first Shopping Trip. Therefore, attrition between the baseline and the first Shopping Trip cannot be attributed to the treatment assignment. Partial completion of the first survey, however, is a potential concern. Panel A of Table 3 addresses this and shows no differential retention through Trip 1 nor any selective attrition based on baseline FV spending nor any differences in this type of attrition across treatments. Thus, our data from the baseline to Trip 1 satisfy the requirements for identifying the LATE in our entire population.

Table 3. Shopping Trip Completion and Attrition by Treatment

	Part 1: Mar–Jul, 2018						Part 2: Mar–Apr, 2019		
	Control	Restricted	Agency	Waiting Period (Delayed)	Waiting Period (Early)	**P-value**	Restricted	Waiting Period (No Agency)	**P-value**
Panel A: Attrition between Baseline and Shopping Trip 1									
Mean retention:	73.8%	73.9%	72.0%	66.5%	67.6%	**0.449**	82.6%	75.5%	**0.134**
Selective retention: (by baseline FV spending)			Pooled: 0.000			**0.904**		Pooled: 0.001	**0.641**
Differential selective retention:	0.002	0.001	-0.000	-0.002	-0.000	**0.984**	-0.001	0.003	**0.489**
Panel B: Attrition between Baseline and Shopping Trip 4									
Mean trips completed	2.752	2.486	2.503	2.178	2.497	**0.077**	3.040	2.563	**0.014**
Selective completion: (by baseline FV spending)			Pooled: 0.003			**0.669**		Pooled: 0.008	**0.464**
Differential selective completion:	0.001	0.010	0.004	-0.006	-0.005	**0.976**	-0.007	0.019	**0.276**
Panel C: Attrition between Baseline and Endline									
Mean retention:	62.4%	53.5%	54.0%	43.8%	53.8%	**0.020**	68.5%	47.7%	**0.000**
Selective retention: (by baseline FV spending)			Pooled: 0.001			**0.798**		Pooled: 0.002	**0.510**
Differential selective retention:	-0.003	0.004	0.001	-0.002	-0.000	**0.860**	-0.007	0.019	**0.278**

Panels A and C report results from linear probability estimates of dropping out of the sample from the baseline to Trip 1, and the baseline to the endline, respectively (robust standard errors). Panel B reports the linear regression coefficient for the total number of shopping trips completed (from 0 to 4). The first row of each panel reports means for each treatment condition along with a joint F-test of equality across all treatments, with the *p*-values reported. The second row of each panel reports the results of a regression of the retention/completion measure on baseline FV spending (pooled across treatments) to test whether attrition is associated with a shopper's baseline healthy spending. The p-values reported are a significance test for the coefficient modifying baseline FV spending. The final row of each panel reports the coefficient estimates for the interactions between each treatment and baseline FV spending to test whether different treatments are associated with different rates of selective attrition based on a shopper's baseline healthy spending. The p-values reported are derived from a joint F-test of all coefficients being equal to zero.

In Panel B of Table 3, we explore potential attrition after learning of the treatment procedures. Here, we present the mean number of Shopping Trips completed by treatment. Depending on the treatment, shoppers in Part 1 complete 2.2–2.8 Shopping Trips, and shoppers in Part 2 complete 2.6–3.0 Shopping Trips. We find higher completion rates amongst the Control group in Part 1 and the Restricted treatment in Part 2. Importantly, we see no significant differences in attrition for T_1 through T_4 in Part 1 ($p = 0.251$), which are used to test our main hypotheses of interest—the effects of agency and waiting periods on food purchases. Moreover, in the second row, we show there is no evidence of selective attrition in the completion of Shopping Trips 1-4. In the third row, we show that there is also no evidence of differentially selective attrition across treatments. Combined, this points to the internal validity for our main hypotheses, which compare FV spending conditional on receipt of subsidies (that is, differences between T_1—T_5 rather than differences between treatments and Control). Additionally, these tests demonstrate that we are able to recover the LATE-R for all tests, even those that include the Control group.

Panel C of Table 3 tests for differences in the retention of shoppers from baseline through to completion of the endline survey. This longer margin of attrition shows sharper differential attrition across treatments. Nonetheless, we still find that there is no statistically significant differential attrition between treatments T_1 through T_4 in Part 1 ($p = 0.148$), suggesting that we can cleanly identify differences between treatments. Additionally, this attrition is not selective based on baseline FV spending and any selective attrition does not differ across treatments, indicating that we should be able to recover the LATE-R for our primary hypotheses.

We also implement the randomization test procedure from Ghanem et al. (2021) to evaluate the potential for attrition bias in our study.[29] Keeping Study 1 and 2 separate, we test for attrition bias on the three samples that match Table 3: Baseline to Shopping

[29]The model is specified with FV spending from the baseline survey as our baseline outcome, non-missing FV spending from the shopping trips as our response variable, the shopping trip number as the time variable, accounting for SNAP participation and a desire for more FV as our stratified randomization variables, and with standard errors clustered at the shopper level.

Trip 1, Baseline to Shopping Trip 4, and Baseline to Endline. The Ghanem et al. (2021) procedure produces two p-values for each test, one for testing the recovery of the LATE and one for the LATE-R. In all 12 tests, we fail to reject the null hypothesis of uniform attrition. The smallest LATE p-value is 0.671, and the smallest LATE-R p-value is 0.481.

Given these results on our ability to identify the LATE and LATE-R, we will proceed with our main analysis in accordance with our pre-analysis plan. We will also present specifications that only use data from the first Shopping Trip, where the LATE can be recovered for all hypotheses, and for the set of subjects that complete all four trips. When expanding to the full sample, the marginally significant attrition concerns of Panel B of Table 3 suggest that we may need to conservatively interpret results that compare treatments to Control as LATE-R—the impact of treatments on shoppers' willing to continue participating in the study. Critical for the main goal of our study, since attrition is not differential between treatments, we can safely interpret comparisons across treatments as the LATE. [30]

3 Hypotheses

In this section, we present four hypotheses about the way in which subsidies will affect FV purchases and how our behavioral interventions will enhance the effectiveness of these subsidies.

Our first hypothesis concerns how healthy food subsidies will affect FV purchases.

Hypothesis 1: *Healthy subsidies will increase purchases of fruits and vegetables.*

As with all ordinary economic goods, the law of demand states that demand for FV will increase as price falls.

Hypothesis 2: *Giving shoppers agency in choosing between healthy and unhealthy*

[30] There are 22 instances of our partner mistakenly advancing a shopper to the next trip following an invalid submission, giving the appearance of reverse attrition. Excluding these individuals has no impact on our results.

subsidies will not decrease FV purchases relative to the restricted subsidy.

This stands in contrast to the predictions of the standard neo-classical model. Our experiment allows us to compare shoppers who are restricted to FV subsidies to shoppers who have agency in choosing between FV and BG subsidies. The opportunity to choose between subsidies will weakly increase the price of FV and decrease the price of BG relative those restricted to FV subsidies. Standard economic theory predicts that this should decrease FV purchases.[31]

Behavioral theories such as a demand for consistency (Falk and Zimmermann 2018) and self-signaling Bénabou and Tirole (2004, 2011) predict that giving shoppers an *active choice* between subsidies could actually increase FV spending. Falk and Zimmermann (2018) argue that an initial active choice impacts how subsequent information is processed, such that information that runs counter to this choice is either downweighted or ignored. This asymmetry in information processing leads to behavior that is consistent with the initial choice. In the self-signaling framework proposed in Bénabou and Tirole (2004, 2011), people are uncertain about their true underlying preferences and look at prior choices to guide subsequent behavior. In this way, prior choices act as informative signals that spur individuals to act consistently.[32]

This consistency effect is illustrated in the "foot-in-the-door" paradigm of Freedman and Fraser (1966). The authors asked people to complete a relatively burdensome task. Prior to this appeal, one group was contacted with a simple request to which the vast

[31] Shoppers under the restricted subsidy experience both an income and substitution effect in favor of greater FV spending. Absent any behavioral mechanisms, shoppers with agency experience an income effect that could increase FV spending, but a substitution effect that clearly decreases it. It is, therefore, unlikely for a classical model to predict greater FV spending under agency than under the restricted subsidy. The income effect would not only need to dominate the substitution effect but would need to do so by more than the *combined* income and substitution effects of the restricted subsidy. In Appendix Section A.1, we show that the necessary conditions for such an effect to manifest are inconsistent with our data. In particular, FV subsidies decrease BG spending, excluding the possibility that increases in FV spending could be driven by a complementary relationship between FV and BG, where income effects dominate the higher FV prices.

[32] Prior choices can also act as strategic forward-looking signals to motivate oneself to behave consistently later on. For example, Andreoni, Kuhn, and Samuelson (2019) show that when subjects can select the parameters of a game they are about to play, they cooperate at higher rates than subjects placed into identically parameterized games without the choice.

majority consented. The other group was presented with the same burdensome appeal without the small request. Those who agreed in the first stage behaved consistently with their initial choice. As a result, there was a higher rate of compliance amongst those who were presented with a small initial request compared to those who were not. Subsequent work has shown that the tendency to act consistently is driven by a change in attitude towards the target act (Bem 1967; Gneezy, Imas, Brown, Nelson, and Norton 2012).

In our setting, where the vast majority of our shoppers state a desire to increase FV consumption, both the demand for consistency and self-signaling models predict that the *act of choosing* a FV subsidy will cause a shopper to spend more on FV than if they had passively received it. Thus, agency presents shoppers with a technology for leveraging their own behavioral responses that does not exist under a restricted healthy subsidy. A demand for consistency can counteract the price effects of increasing agency, potentially even dominating it. The self-signaling framework makes an additional prediction: that shoppers will be willing to choose the healthy subsidy even if it means leaving money on the table, since this would generate a costly signal of their intention to eat healthier. We examine this prediction in Section 4.3.

Hypothesis 3: *Waiting periods between the delivery of subsidy information and the shopping decision will increase FV purchases.*

Across three experiments, Imas et al. (2018) demonstrate that introducing waiting periods between information about a choice and the choice itself leads to more patient and far-sighted decisions. One possible psychological pathway for this effect is that waiting periods prompt deliberation and prospection of future utility outcomes (Gilbert and Wilson 2007; Wheeler, Stuss, and Tulving 1997). Gabaix and Laibson (2017) formalize this in a theoretical model where an individual is uncertain about the future utility consequences of her choices (e.g., increasing FV spending) and generates imperfect forecasts by mentally simulating the potential outcomes. The forecasted utility consequences of a choice are estimated with noise that increases with time;

outcomes further in the future have noisier distributions of potential consequences and as a result, are more heavily discounted. Mentally simulating these consequences because of the introduction of a waiting period generates unbiased signals of "true" utility, which, when combined with prior beliefs, reduce noise in the forecast. Under reasonable assumptions, this process leads to less myopic decision making.

Waiting periods may also affect food choices by allowing time for shoppers to set goals for their purchases. Research has shown that goal-setting and endogenously-established reference points can provide powerful motivation (Fishbach and Ferguson 2007; Hsiaw 2013; Koch and Nafziger 2011). Moreover, as shown theoretically by Koszegi and Rabin (2008), and empirically by Heffetz (2018), reference points may require time in order to "sink in." The same may be true for preference adaptation as outlined by Bernheim, Braghieri, Martinez-Marquina, and Zuckerman (2019). Our waiting period intervention can provide the needed time for preference adaptation or for reference points to sink in, which will impact shopping behavior accordingly. We focus on the deliberation model here as it motivated our study design.[33]

As with previous studies in behavioral economics, we consider the trade-off between the immediate pleasure and the delayed health consequences of food choices as an exercise in patience. For this reason, we predict that introducing a waiting period between the delivery of subsidy information and the grocery shopping will lead shoppers to take greater advantage of their healthy subsidies and spend more on FV.[34]

Hypothesis 4: *The early subsidy choice—before the waiting period—will increase FV spending compared to the delayed subsidy choice.*

[33] Thakral and To (2020a,b) suggest an alternative pathway for the effect of waiting periods—the role of anticipation. During delays, individuals to experience anticipatory utility, which reduces the preference for consuming temptation goods out of windfall. This mechanism is more applicable in explaining large, unexpected cash transfers, for example, but is less relevant for a healthy-food subsidy.

[34] Our waiting periods introduce a delay between information about the subsidy and the Shopping Trip on which the subsidy will be used. Thus, depending on the treatment assigned, this effect can operate through two margins: (1) increased selection of the healthy subsidy and (2) increased healthy spending conditional on the subsidy selected. In Section 4 we examine both dimensions, showing that waiting periods have no impact on subsidy selection. As a result, we focus on the capacity of the waiting period to prompt deliberation about the grocery shopping decision itself.

Allowing shoppers to make a subsidy choice before the waiting period begins, as in our Waiting Period (Early Choice) treatment, can increase FV spending through two channels. First, if immediate consumption generates temptation, then Early Choice may overcome present-bias in subsidy selection. Present-biased preferences may lead shoppers who prefer the FV subsidy before entering the store to nonetheless choose the BG subsidy after entering due to the immediate temptation of the unhealthy food. Early Choice allows shoppers to commit to the FV subsidy before facing this temptation, increasing FV spending by lowering average FV prices. However, conditional on the chosen subsidy, models of present-bias make no predictions about the impact of early choice on purchasing behavior.

A second channel motivates our Hypothesis 4. Early Choice allows shoppers to narrow the set of potential future outcomes and reduce uncertainty around relative prices. In the imperfect forecasting model of Gabaix and Laibson (2017), this acts to decrease the number of prospective states that the individual must consider when deliberating on future utility consequences. Waiting periods are thus predicted to be even more effective in reducing noise around forecasts of long-run health consequences, which is predicted to further increase FV spending.

Our data can distinguish between these two mechanisms by analyzing subsidy choice rates and purchases conditional on a subsidy choice. Under present bias, Waiting Period (Early Choice) will increase FV purchases through greater FV choice rates, but conditional on subsidy choice, food purchases should be similar to Waiting Period (Delayed Choice). In contrast, the imperfect forecasting model predicts an increase in FV spending compared to Waiting Period (Delayed Choice) even if subsidy choice rates remain the same.

We now proceed to test these hypotheses in our data.

4 Results

We find that each marginal intervention increases FV spending: in Part 1, the mean FV spending is \$4.03 in Control, \$8.67 in Restricted, \$9.31 in Agency, \$10.64 in Waiting

Period (Delayed Choice), and \$12.17 in Waiting Period (Early Choice). In Part 2, the mean is \$11.23 in Restricted, and \$12.20 in Waiting Period (No Agency). To explore the impact of our treatments on FV spending we use a linear, random-effects regression with shopper-specific random effects and standard errors clustered at the shopper level (as specified in the pre-registration). Our estimation equation is given by Equation 1 below:

$$\text{FV}_{i,t} = \alpha + \Gamma' Z_i + \beta \times \text{FV}_{i,0} + \delta_{w(i)} + u_i + \varepsilon_{i,t} \quad . \tag{1}$$

$\text{FV}_{i,t}$ is the FV spending for shopper i on Shopping Trip t.[35] Z_i is a vector of indicator variables for either assigned treatments, or partitions of the treatment set. $\text{FV}_{i,0}$ controls for FV spending from the baseline survey to increase precision. Because the randomization weights changed slightly over the course of the study, we use assignment-wave fixed effects, $\delta_{w(i)}$, to ensure robustness. We do this both at a coarse level, with a fixed effect for Part 1 or 2, and at a fine level, with a "wave fixed effect" for every randomization group within Part 1. u_i is the shopper-specific random effect. We present results with and without the baseline control and assignment-wave fixed effects to demonstrate that the treatment effects are unaffected by their inclusion.

4.1 Effect of Subsidies

We first demonstrate the effectiveness of subsidies on increasing FV spending. Table 4 pools across subsidy treatments to present these large, positive effects. Using the full specification in Column (3), the average effects of all subsidy treatments leads to a FV spending increase of 164%—or 0.98 SD—relative to the Control group.

Column (4) restricts the sample to just the first Shopping Trip, and we estimate a slightly larger effect of the subsidy. Column (5) then shows that the main effects are almost unchanged when we restrict the sample to only shoppers who complete all four Shopping Trips.

Increases in FV spending can arise through a substitution effect that shifts pur-

[35] This is the pre-subsidy, gross spending before subtracting off any subsidy payments.

Table 4. Effect of Subsidy Treatments on FV Spending ($)

Sample:	All Trips			Trip 1	Completers
Control Mean [SD]: 4.03 [6.71]	(1)	(2)	(3)	(4)	(5)
Subsidy	6.11***	6.33***	6.60***	7.89***	6.87***
	(0.60)	(0.60)	(0.61)	(0.89)	(0.68)
Baseline Survey FV Spending ($)		0.16***	0.15***	0.16***	0.16***
		(0.05)	(0.05)	(0.05)	(0.06)
Part 1, Part 2 Fixed Effects	Y	Y	Y	Y	Y
Wave Fixed Effects	N	N	Y	Y	Y
Observations	2767	2767	2767	781	2456
Clusters	805	805	805	N/A	625

*** $\Rightarrow p < 0.01$. Standard errors in parentheses are clustered at the shopper level. All specifications are linear random-effects models except Column (4), which is OLS. "Subsidy" pools all treatments featuring subsidies whether or not they also feature behavioral interventions ($T_1 - T_5$ from Table 1). "Baseline Survey FV Spending" is captured based on shopping receipts submitted on the baseline survey. Column (4) restricts the sample to only spending on the first Shopping Trip. Column (5) restricts the sample to spending across all Shopping Trips but only for shoppers who completed all four Shopping Trips in the study.

chases towards FV, through an income effect that increases spending across-the-board, or both. To show the effect of substitution towards FV, Table 5 presents the increase in the fraction of food spending on FV.[36] The subsidy treatments increase the share of food spending on FV by 15 percentage points from a mean of 13% in the control group. Appendix Table A.1 provides further evidence of the prominence of the substitution effect, finding no significant increases in total food spending. Additionally, Appendix Table A.2 shows that the marginal impact of a subsidy is, if anything, a negative for BG spending.[37]

Figure 1 shows that the subsidy treatments have the largest effect on the quasi-extensive margin. Subsidies move a large mass of people who spend very little on FV

[36] Our receipt tabulations include the receipt total and non-food spending. When both are present, we define total food spending as the difference. If (1) the non-food expenditure data is missing, (2) imputed total food spending is negative, or (3) total FV and BG spending exceeds imputed total food spending, then we use the sum of FV and BG spending as total food spending.

[37] Sadoff et al. (2020) finds that interventions increasing fruit and vegetable spending may encourage substitution away from sweets and salty snacks.

Table 5. Effect of Subsidy Treatments on FV Spending as a Fraction of Food Spending

Sample:	All Trips			Trip 1	Completers
Control Mean [SD]: 0.13 [0.22]	(1)	(2)	(3)	(4)	(5)
Subsidy	0.15***	0.15***	0.15***	0.19***	0.16***
	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Baseline Survey FV Fraction		0.16***	0.16***	0.25***	0.17***
		(0.04)	(0.04)	(0.06)	(0.05)
Part 1, Part 2 Fixed Effects	Y	Y	Y	Y	Y
Wave Fixed Effects	N	N	Y	Y	Y
Observations	2745	2710	2710	766	2410
Clusters	804	793	793	N/A	618

*** $\Rightarrow p < 0.01$. Standard errors in parentheses are clustered at the shopper level. All specifications are linear random effects models except Column (4), which is OLS. Baseline FV spending is captured based on shopping receipts submitted on the baseline survey. Total spending on FV is divided by total food spending to calculate the FV fraction. We lose 22 observations and 1 shopper relative to the count in Table 4 because of receipts with no food items (and, thus, a zero in the denominator). We similarly lose 35 observations and 11 shoppers in Columns (2) and (3) relative to the count in Column (1) due to baseline survey receipts with no food items. "Subsidy" pools all treatments featuring subsidies whether or not they also feature behavioral interventions ($T_1 - T_5$ from Table 1). Column (4) restricts the sample to only spending on the first Shopping Trip. Column (5) restricts the sample to spending across all Shopping Trips but only for shoppers who completed all four Shopping Trips in the study.

($0 to $5) toward spending intermediate amounts. Statistical tests show large and statistically significant differences in the distributions ($p < 0.001$). While changes in FV spending do not guarantee corresponding changes in FV consumption, in Section 5 we present evidence that this does seem to be the case for our study.

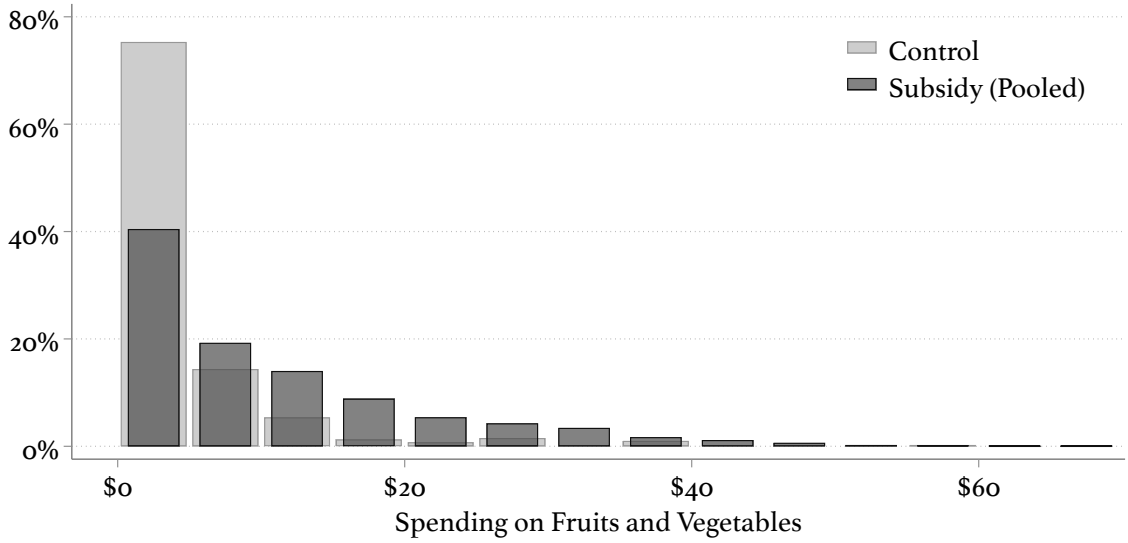

Figure 1. Histogram of fruit and vegetable spending.

4.2 Behavioral Food Subsidies

Panels A – D of Figure 2 summarize the impact of each of our treatments from Part 1. Along the x-axis, the treatments grow in the intensity of the intervention. Recall that Waiting Periods (Early Choice) combines all of our behavioral interventions: agency, waiting periods, and early choice. Panels A and B show that increasing the intensity of the interventions has a strong, positive effect on the mean FV spending and no effect on the mean BG spending. When all interventions are combined in the Waiting Periods (Early Choice) treatment, there is an $8.22 increase in FV spending. This is 61% larger than the effect of the restricted subsidy treatment and more than triple the mean FV spending in the control group of $4.03. Panel C corroborates these findings, estimating the treatment effects on FV spending after controlling for pre-intervention FV spending. Panel D performs a similar treatment effect estimation and shows a uniform null effect on BG spending across treatments after controlling for pre-intervention BG spending.[38]

[38]Figure A.3 in the Appendix reproduces Panel C for each of the four shopping trips in the study. All treatments represent significant improvements over the control at the 99% confidence level in all trips. The increasing effect with treatment intensity is most noticeable in Trips 1 and 2, and we

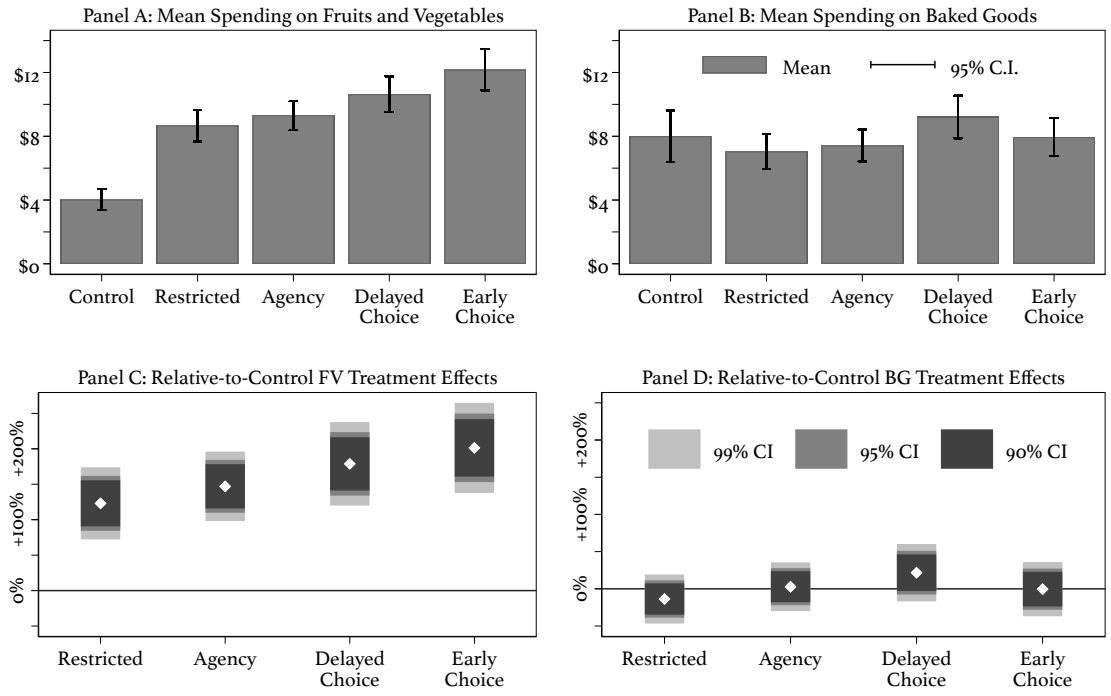

Figure 2. Panels A & B: Mean spending; Panels C & D: Treatment effects (Part 1 Only)

We now estimate the marginal impacts of agency and waiting periods on subsidy effectiveness. According to our hypothesis, a preference for consistency can only counteract the price effect of increased agency if the majority of shoppers choose the FV subsidy from the larger choice set. Indeed, we find that 78% of shoppers under the agency intervention select the FV subsidy.

Importantly, we find *no treatment effects on the FV subsidy choice rate*. The FV subsidy is selected 76% of the time in the Agency treatment, 79% of the time in the Waiting Period (Delayed Choice) treatment, and 79% of the time in the Waiting Period (Early Choice) treatment. We fail to reject the hypothesis that these rates are equal across all three treatments ($p = 0.82$). Equal selection rates means that any differences in average behavior imply differences in behavior *conditional* on subsidy choice, that is, differential subsidy effectiveness.

see the largest effects of the waiting periods in Week 1.

Table 6 shows the effects of our two primary behavioral interventions, agency and waiting periods. Columns (1)-(3) present the estimated effect of the subsidies and jointly estimates the marginal impact of agency and waiting periods. Column (4) reveals the immediate impact of the interventions by replicating this analysis for the first Shopping Trip only. This also has the benefit of restricting our analysis to a sample that has very limited attrition and no concern about differential completion rates (see Panel A in Table 3). Column (5) again replicates the analysis, but only for shoppers who completed all four Shopping Trips.

Table 6. Effect of Agency and Waiting Periods on FV Spending ($), Combined Data

Sample:	All Trips			Trip 1	Completers
Control Mean [SD]: 4.03 [6.71]	(1)	(2)	(3)	(4)	(5)
Subsidy	4.77***	5.06***	5.11***	6.90***	5.07***
	(0.82)	(0.83)	(0.81)	(1.27)	(0.87)
Agency	0.75	0.62	1.06	-1.14	1.55
	(0.87)	(0.88)	(0.87)	(1.34)	(0.94)
Waiting Period	1.61**	1.69**	1.46**	4.05***	1.35*
	(0.70)	(0.68)	(0.67)	(0.98)	(0.74)
Baseline Survey FV Spending ($)		0.16***	0.15***	0.16***	0.16***
		(0.05)	(0.05)	(0.05)	(0.06)
Part 1, Part 2 Fixed Effects	Y	Y	Y	Y	Y
Wave Fixed Effects	N	N	Y	Y	Y
Observations	2767	2767	2767	781	2456
Clusters	805	805	805	N/A	625

*** $\Rightarrow p < 0.01$, ** $\Rightarrow p < 0.05$, * $\Rightarrow p < 0.10$. Standard errors in parentheses are clustered at the shopper level. All specifications are linear random effects models except Column (4), which is OLS. "Subsidy" pools all treatments featuring subsidies whether or not they also feature behavioral interventions ($T_1 - T_5$ from Table 1). "Agency" pools all treatments that featured a choice between healthy and unhealthy subsidies ($T_2 - T_4$ from Table 1). "Waiting Period" pools all treatments that featured a waiting period between the pre-shopping task and the shopping task ($T_3 - T_5$ from Table 1). "Baseline Survey FV Spending" is captured based on shopping receipts submitted on the baseline survey. Column (4) restricts the sample to only spending on the first Shopping Trip. Column (5) restricts the sample to spending across all Shopping Trips but only for shoppers who completed all four Shopping Trips in the study.

Agency

We first turn to estimating the effect of agency on FV spending. With wave fixed effects and baseline spending controls in Column (3), we estimate that the agency treatments lead to \$1.06 more FV spending than the restricted subsidy—a 21% increase in the effect of the subsidy—though the coefficient is not statistically significant ($p = 0.22$). Importantly, 22% of shoppers with agency do not select the FV subsidy, increasing the price of FV relative to the Restricted treatment. The law of demand predicts that this will diminish FV spending, but our results suggest this does not occur. We can rule out any decrease in FV spending more than \$0.64 at the 5% level. That is, consistent with our second hypothesis, the behavioral effects associated with increased agency appear to compensate for the price effects of some shoppers choosing the BG subsidy.[39]

One concern with repeatedly eliciting subsidy choices from shoppers is that they may engage in inter-temporal substitution, alternating between subsidized FV purchases one week and subsidized BG purchases in a later week. We see no evidence of this in the subsidy choices: 58% of shoppers never change their subsidy choice throughout the study and only 7% change *and* change back to their original subsidy choice. For all shoppers, the previous subsidy choice has a large, positive, and statistically significant correlation with the current subsidy choice. Additionally, there are no aggregate trends with respect to subsidy choice across time.

Despite the positive average effect, for the minority of shoppers who choose the unhealthy subsidies, increased agency effectively incentivizes BG purchases. Thus, an important concern is whether our intervention has a negative impact on the FV spending of some of our shoppers. Appendix Figure A.4 in shows the CDFs of FV spending in the control treatment and the pooled agency treatments. The CDF of the agency treatments dominates the CDF of the control treatment. That is, regardless of

[39]This estimate is not stable when we restrict our sample. Looking at Column (4) and Column (5) suggests a neutral or a modestly positive effect that could be detected with greater power. Importantly, none of the specifications contradict our second hypothesis that agency does not decrease healthy food spending, despite the price effects.

a shopper's unsubsidized FV spending, our agency intervention does not decrease her FV spending, even if the FV subsidy is not chosen.[40]

Waiting Periods

Column (3) of Table 6 shows that the waiting period treatments generate \$1.46 more FV spending than treatments without a waiting period—a 29% increase in the size of the subsidy effect relative to the restricted subsidy.[41] The significant, positive effect of waiting periods provides evidence for our third hypothesis.[42] There is evidence that the impact of waiting periods carried over beyond the end of the subsidized shopping trips; among subsidized households, the endline-minus-baseline difference in FV spending is \$2.64 larger for households with waiting periods ($p = 0.04$).[43]

We now separately consider the impact of allowing shoppers to make their subsidy choice prior to the waiting period (Early Choice) compared to after the waiting period is over (Delayed Choice). We previously established that Early Choice has no impact on subsidy choice. Table 7 exhibits results on FV purchases, presenting the individual impacts of each of the two treatments above the impact of restricted healthy subsidies. First, we find that the waiting period with Delayed Choice causes FV spending to significantly increase by \$2.42, or 26% ($p = 0.020$), relative to the Restricted subsidy. The waiting period with Early Choice causes an even larger increase in FV spending relative to the Restricted subsidy—\$3.18, or 34% ($p = 0.003$).

[40] A related concern is that the impact varies by wealth, leaving the poorest households unaffected. This concern is mitigated by the fact that our entire sample is low-income ($< 185\%$ of the FPL). Moreover, Appendix Table A.4 shows that the subsidies do not have significantly different impacts by welfare status. We do find that welfare-recipient households are less likely to select the FV subsidy, but the difference in choice rates does not produce differential estimates of the treatment effects.

[41] The waiting period continues to show a positive and significant effect when we restrict our sample. Column (4) shows the impact of waiting periods during only the first Shopping Trip. In this case, the effect of waiting periods is much larger and it retains precision ($p < 0.001$). Column (5) considers only shoppers who completed the entire study and finds that the impact of waiting periods is still positive but it is less precisely estimated ($p = 0.07$).

[42] Table 6 estimates the marginal impact of waiting periods across environments both with and without agency. We can estimate the marginal effect of waiting periods in the absence of agency by using a subsample of our data. This lower-powered test shows that the effect size falls to \$1.14 and is not significant ($p = 0.34$) due, in part, to increased standard errors.

[43] OLS estimate includes wave fixed effects and heteroskedasticity-robust standard errors.

While the difference between treatments is directionally in favor of Early Choice further increasing FV spending, it is not significant.[44] These results are therefore only suggestive evidence for our fourth hypothesis. Notably, because we do not find any impact of the interventions on subsidy choice rates, these positive effects point to Early Choice making the waiting periods more effective.

Table 7. Effect of Waiting Periods and Choice Timing on FV Spending ($)

Sample:	All Trips	Trip 1	Completers
	(1)	(2)	(3)
T_3: Waiting Period (Delayed Choice)	2.42**	1.98	2.53**
	(1.04)	(1.57)	(1.19)
T_4: Waiting Period (Early Choice)	3.18***	5.02***	3.98***
	(1.08)	(1.81)	(1.17)
Baseline Survey FV Spending ($)	0.10	0.10	0.07
	(0.07)	(0.09)	(0.08)
Constant (T_1: Restricted Subsidy)	9.32***	10.72***	9.69***
	(0.91)	(1.33)	(1.05)
Early Choice - Delayed Choice:	0.76	3.04	1.45
	($p = 0.522$)	($p = 0.103$)	($p = 0.287$)
Part 1, Part 2 Fixed Effects	N/A	N/A	N/A
Wave Fixed Effects	Y	Y	Y
Observations	1075	308	929
Clusters	327	N/A	242

*** $\Rightarrow p < 0.01$, ** $\Rightarrow p < 0.05$, * $\Rightarrow p < 0.10$. Standard errors in parentheses are clustered at the shopper level. Columns 1 & 3 present results from a linear random effects specifications. Column (2) presents results from a OLS specification. Sample is restricted to T_1: Restricted, T_3: Waiting Periods (Delayed Choice), and T_4: Waiting Periods (Early Choice), where T_1 constitutes the omitted group. "Baseline Survey FV Spending" is captured based on shopping receipts submitted on the baseline survey. Column (2) restricts the sample to only spending on the first Shopping Trip. Column (3) restricts the sample to spending across all Shopping Trips but only for shoppers who completed all four Shopping Trips in the study.

[44] Column (2) of Table 7 restricts the sample to only the first Shopping Trip and finds results that are largely consistent with the full sample. The impact of Early Choice does appear larger in the first Shopping Trip, and the marginal impact over Delayed Choice approaches statistical significance ($p = 0.103$). When restricting the sample to only shoppers who completed all Shopping Trips, we again find similar results, though the difference between Early Choice and Delayed Choice falls closer to the difference identified in the full sample.

4.3 Subsidy choice as a costly signal

Except in the case of the Waiting Period (Early Choice) treatment, shoppers selected their subsidies while in the grocery store at some point during their shopping trip. Thus, a money-maximizing shopper could ignore the subsidy, make her grocery shopping decisions, and then select the subsidy that would maximize the subsidy payment she would receive. Unlike many studies involving commitment, there is little scope for uncertainty about future shopping tastes.[45] Despite this available strategy, we find that many shoppers end up leaving money on the table.

Shoppers are significantly more likely to sacrifice money in order to select the healthy subsidy rather than the unhealthy one. We find that a minimum of 28% of shoppers selecting the FV subsidy would have walked away with more money had they chosen the BG subsidy *even without accounting for changes in behavior that would result from choosing the BG subsidy.* These shoppers sacrifice an average of $3.37 each (101% of the mean subsidy). This compares to 17% of shoppers sub-optimally selecting the BG subsidy, losing an average of $2.16 (64% of the mean subsidy).[46] The difference in proportions (28 vs. 17%) is statistically significant ($p < 0.01$). Shoppers selecting the healthy subsidy on average only earn $1.67 more than they would have with an unhealthy subsidy while the differential grows to $2.52 for those selecting the unhealthy subsidy ($p < 0.001$). Both results suggest that a significant fraction of the costly over-selection of healthy subsidies cannot be attributed to random choice errors, confusion, or misunderstanding.

This behavior offers an explanation for why adding the option of the BG subsidy did not decrease FV spending relative to the Restricted treatment. Shoppers may use

[45] The similarity of subsidy choice rates across treatments with and without the early choice suggests that even when the subsidy is selected 4-48 hours in advance, shopping taste uncertainty may not be an important contributor to subsidy choice, though of course average rates may be masking important heterogeneity.

[46] We assume that spending on the subsidized category is a conservative upper-bound for counterfactual spending without the subsidy. For example, if Shopper A selects the FV subsidy and spends $9 on FV and $10 on BG, we classify this choice as costly. If Shopper B selects the FV subsidy and spends $9 on both FV and BG, we do not classify this choice as costly, even though it most likely is.

the subsidy choice to self-signal and spur increased purchases of FV. That is, consistent with the self-signaling framework of Bénabou and Tirole (2004), shoppers may select the FV subsidy with the understanding that it will shift their purchases towards FV even if they will still ultimately spend more on BG than FV. This would counteract the price effect of the BG subsidy option for those who select it. Figure 3 captures the magnitude of this costly signaling for all shoppers who chose the FV subsidy. All shoppers below the 45-degree line would have earned more in subsidy payments had they chosen the BG subsidy; as shown in the figure, some forgo a significant amount of money by actively choosing the FV subsidy.

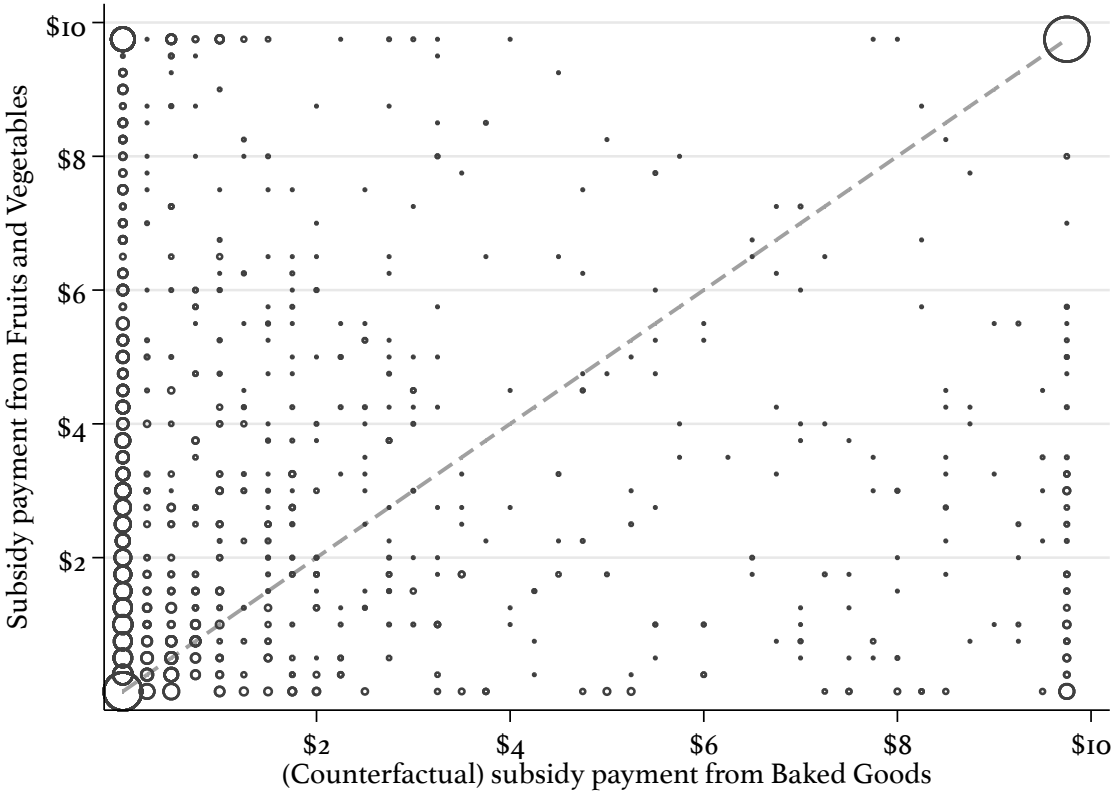

Figure 3. Healthy subsidy payments and counterfactual unhealthy subsidy payments for subjects who select the FV subsidy. Shoppers are grouped into $0.25 intervals for each axis. The marker size reflects the number of shoppers in the intersection of each interval.

5 Consumption and Welfare

We have thus far primarily concerned ourselves with the impact of subsidies and behavioral interventions on FV spending. In order to address nutritional concerns, we now seek to demonstrate that increases in FV spending lead to increases in FV consumption. Moreover, we will argue that these changes in shopper behavior are likely welfare increasing.

5.1 Fruit and Vegetable Consumption

Our data collection from each Shopping Trip did not include measures of consumption. Shopping receipts are the standard object of data collection by our implementing partners, since they are easily submitted as photographs and offer simple validation checks such as timestamps and addresses. These receipts presented us with a form of hard evidence for FV spending in a way that self-reports of consumption could not match. For these reasons, we only collected self-reported 24-hour consumption diaries on the baseline and endline surveys.

Since the endline survey was taken shortly after the intervention, we are able to evaluate the impact of our subsidies on self-reported FV consumption collected on that survey. These measures covered both the shopper and one child from the food household (if there was a child in the food household). Table 8 regresses the change in FV consumption from the baseline to the endline for both the shopper (Column (1)) and one child from their food household (Column (2)) on an indicator variable for receipt of any of our subsidies. While we do not find that the shopper themselves significantly changes their consumption, we do find a positive and significant increase in the FV consumption of the child in the household. Children in households receiving subsidies consumed 0.756 additional servings of FV on the endline food diary than on the baseline.

For additional evidence of how our interventions impact FV consumption, we can use the consumption data from the baseline and endline surveys to identify the associ-

Table 8. Effect of Subsidy Treatments on Changes in FV Consumption (Endline – Baseline)

	Shopper FV Consumption Change	Child FV Consumption Change
	(1)	(2)
Subsidy	0.234	0.756**
	(0.340)	(0.325)
Constant	1.041***	0.589*
	(0.371)	(0.317)
Wave Fixed Effects	Y	Y
Observations	399	215

** $\Rightarrow p < 0.05$, * $\Rightarrow p < 0.10$. Standard errors in parentheses are clustered at the shopper level. All specifications are OLS models. "Subsidy" pools all treatments featuring subsidies whether or not they also feature behavioral interventions ($T_1 - T_5$ from Table 1). Consumption change is calculated based on the difference between the FV consumption on the 24-hour food diaries submitted on the endline and baseline surveys. Sample sizes differ because not all households included at least one child.

ation between FV spending in a given shopping trip and self-reported FV consumption for a given 24-hour period. We can then extrapolate this association to estimate the additional servings of FV consumed as a result of the increased FV spending caused by the subsidies. Note that spending is measured by the Shopping Trip, which may cover groceries for a week or more, while consumption is measured for a 24-hour window. Table 9 presents the results of this analysis.

Each additional dollar of FV spending on the baseline and endline surveys is associated with an additional 0.025 servings of fruits and vegetables consumed in a 24-hour period. Combining this association with results from Table 6, we can estimate the impact of each intervention on FV consumption. We find that the Restricted healthy subsidy increases FV spending by \$5.11, thus generating 0.127 additional servings of FV consumption. Table 7 estimates that the the most effective behavioral intervention, the Waiting Period (Early Choice) treatment, would increase FV spending by an additional \$3.16 beyond the Restricted healthy subsidy. This should generate an additional 0.079 servings of FV consumption.

The USDA recommends consuming between 4 and 6 servings of FV daily for adults

Table 9. Correlation between FV Consumption and FV Spending ($)

Sample:	Baseline	Baseline + Endline
	(1)	(2)
FV Spending ($)	0.025**	0.025***
	(0.011)	(0.008)
Total Spending ($)	0.001	0.001
	(0.001)	(0.001)
Endline Survey		0.709***
		(0.123)
Constant	1.675***	1.897***
	(0.145)	(0.136)
Mean FV Consumption:	2.478	2.725
Wave Fixed Effects	Y	Y
Observations	760	1182
Clusters	N/A	783

*** $\Rightarrow p < 0.01$, ** $\Rightarrow p < 0.05$, * $\Rightarrow p < 0.10$. Standard errors in parentheses are clustered at the shopper level. All specifications are linear random effects models. Sample is restricted to baseline survey data (Column (1)) or baseline and endline survey data (Column (2)). FV spending and total spending are captured based on shopping receipts submitted on the baseline and endline surveys. "Endline Survey" is an indicator variable that equals one if the data come from the endline survey.

(USDA 2021), a goal that only 19% of our shoppers manage to meet. Thus, one approach to determining the impact of our subsidies would be to measure how much closer our shoppers come to meeting their FV consumption recommendations. Using a linear approximation based on Column (2) from Table 9 to identify shoppers with predicted FV consumption between 4 and 6 servings, we find that their mean FV spending is $60.36 or nearly 15 times the mean FV spending of the control group, with our subsidy closing roughly 10% of the gap.

5.2 Welfare Analysis

Given that 83% of our shoppers report wanting to consume more fruits and vegetables at baseline and the widely documented under-consumption of fruits and vegetables relative health benchmarks (Cole and Fox 2008; Rehm et al. 2016), there is reason to

believe that any increase in FV spending is likely a welfare improvement. However, as Bernheim and Taubinsky (2018) demonstrates, welfare evaluation in the presence of behavioral biases is not straightforward. They consider optimal taxation when consumers face "internalities" wherein the true consumption value of an object is shaded by a lack of attention, a lack of salience, incorrect beliefs, imperfect self-control, or other behavioral biases.[47]

Since our subsidies reflect a negative tax on FV spending, Bernheim and Taubinsky (2018) provide a straightforward criteria for identifying the welfare-maximizing level of subsidization—the subsidy should exactly offset the marginal internality. They derive the first-order condition for welfare maximization with respect to a tax, t: $\frac{\partial W}{\partial t} = (t - \bar{\gamma})\frac{\partial D}{\partial t}$, where $\bar{\gamma}$ is the average marginal bias and D is the demand curve. To adjust this model to our context, we consider γ to be the *undervaluation* of FV, and substitute a subsidy, s, in place of a tax. These changes flip the signs for the respective terms, giving us the first-order condition:

$$\frac{\partial W}{\partial s} = (\bar{\gamma} - s)\frac{\partial D}{\partial s} \tag{2}$$

This yields the same result for the optimal subsidy—the subsidy should exactly offset the degree of undervaluation of FV. It also provides a more general result: welfare will be increasing in the subsidies as long as the subsidy does not exceed the average marginal bias. Beyond such bias-correction, expansion of subsidies may over-correct for the internality and subsidize FV purchases whose true marginal benefit falls short of the marginal subsidy cost. Within limits, such over-correction may improve welfare relative to the baseline, but a sufficiently large subsidy will begin to diminish aggregate welfare. Our study adds a layer of complexity to welfare calculations because our behavioral interventions are designed to enhance the effectiveness of the subsidies themselves. This means that it is not sufficient to show that the subsidies alone do not over-correct for

[47] For applications, see Allcott et al. (2019c); Allcott and Taubinsky (2015); Chetty, Looney, and Kroft (2009).

the internality. Rather, we must show that the enhanced "behavioral subsidies" do not over-correct for the internality. In terms of the first-order conditions from Bernheim and Taubinsky (2018), we must include a modifier for the tax term, t, to reflect its changing effectiveness.[48] Call this modifier σ. The first-order condition with respect to the variable, σ becomes:

$$\frac{\partial W}{\partial \sigma} = (\bar{\gamma} - \sigma s)\frac{\partial D}{\partial \sigma} \tag{3}$$

The conclusions of this modified model are more ambiguous than the original model and must be split into multiple cases. First, if $s < \bar{\gamma}$ and the subsidy does not fully correct for the bias, welfare will be increasing as σ helps make up the difference. Indeed the optimal value of σ exactly equates the two terms: $\sigma^* = \frac{\bar{\gamma}}{s}$. Beyond this point, increases in subsidy effectiveness may be counterproductive. Below, we argue that this is the only relevant condition for our context. Second, if $s > \bar{\gamma}$, then the subsidies over-correct for the bias and it is actually optimal to make the subsidies less effective $\sigma^* < 1$. In either case, subsidies and behavioral interventions are substitutable. Thus, optimal subsidization can be achieved through multiple combinations of subsidization and interventions to change the subsidy effectiveness. While increasing the subsidy effectiveness may be less costly than increasing the subsidy in terms of program budgeting, the welfare impact may favor increasing the subsidy if such redistributions improve social welfare.

Our subsidy size (30%, capped at \$10 per trip) was not calibrated to the degree of bias, but rather chosen to mirror that of the USDA Healthy Incentives Pilot program. Nonetheless, we believe that the subsidy size is still small enough that our welfare analysis can conclude that $\sigma s < \bar{\gamma}$ and, therefore, that both the subsidies and the behavioral interventions are welfare increasing. As mentioned in Section 5.1, the goal of our subsidies is not merely to encourage greater FV spending, but to encourage greater FV consumption. By considering consumption rather than spending, we can leverage public-health guidance in establishing a notion of "optimal" consumption. This allows

[48] We thank an anonymous referee for suggesting this theoretical approach.

us to estimate our shoppers' internalities based on how far their consumption falls short of these targets. At baseline, the FV consumption for our shoppers has a mean of just under 2.5 FV servings in a 24-hour period with a median of 2 servings. This mean FV consumption is between 1.5 and 3.5 servings short of the USDA's daily recommendation (USDA 2021). The welfare-maximizing FV subsidy would increase FV spending by an amount sufficient to induce shoppers to consume 1.5—3.5 more servings of FV. Based on estimates from Table 5, this would require an increase in FV spending \$60—\$140, well outside of the range of any of our treatment effects. This strongly suggests that both the subsidies and our behavioral interventions are welfare-increasing.

A final welfare consideration is the psychic costs of each choice. For example, Butera, Metcalfe, Morrison, and Taubinsky (2019) find that interventions that promote healthy behavior such as exercise may still diminish welfare if image concerns outweigh health concerns. In our case, decisions are private, so signaling concerns should be limited. Moreover, we present choices neutrally, without any attempt to assign a negative connotation to unhealthy choices. For these reasons, we believe the expansion of choice and the non-restrictive nature of our interventions are welfare increasing.

6 Conclusion

In this paper, we test the effectiveness of behavioral food subsidies in encouraging healthy food spending. The vast majority of our shoppers state a desire to consume more FV, and prices appear to be an important reason why that desire is not satisfied. Relative to our control group, offering a restricted healthy subsidy more than doubles FV spending (a 124% increase). While we believe this to be an upper bound on the subsidy effect, the finding strongly replicates the quilitative findings of the USDA's Healthy Incentives Pilot (Bartlett et al. 2014). Our main contribution is to show that the effect of healthy food subsidies on FV spending can be increased by 61% by leveraging behavioral interventions. Because the behavioral interventions have no direct costs to the provider relative to the restricted subsidy, incorporating them into a subsidy program could have an large positive impact on the cost-benefit ratio of the

program.

Within the SNAP program, benefit disbursement provides a clear time point when these behavioral interventions can be paired with subsidies. According to Castner and Henke (2011), more than 20% of SNAP benefits are spent on the day they are disbursed, and more than half are spent by the end of the first week. In the days or hours leading to disbursement, there is an opportunity to contact the recipient with information that induces the type of waiting period we have found to be effective here: a notification/reminder of the subsidy that the upcoming benefits will be subject to, delivered in advance of actual benefit receipt; this creates a waiting period that can prompt deliberation. Online SNAP/EBT portals and third-party apps (such as Providers) already offer a place to view benefit balances, SNAP transactions, and disbursement schedules. A linkage between EBT cards and user input on these platforms could permit subsidy choices for the upcoming benefit disbursement as well—giving recipients agency with early choice.[49] We note that waiting periods in particular were most impactful during the first Shopping Trip (see Figure A.3), suggesting that care needs to be taken to maintain the salience of behavioral food subsidies with each disbursement cycle.

Our findings have theoretical implications for behavioral economists as well. We implement agency and waiting periods in the context of food choice, a complex decision with many temptations that may cause a well-intentioned shopper to stray from their goal of eating more fruits and vegetables. Our interventions may help shoppers overcome these short-run temptations and encourage more future-focused food choices. In the case of agency, we find evidence that expanding the choice set allows shoppers to signal their intentions through a (potentially costly) active decision in favor of the healthy subsidy, which we propose generates consistent shopping behavior. We also propose that waiting periods impact behavior by prompting shoppers to deliberate over their upcoming choices and adopt a more future-focused mindset. When the choice of subsidy is made at the beginning of the waiting period, there is suggestive evidence

[49]The HIP demonstrates that EBT-card specific, item-specific subsidies can be deployed (Bartlett et al. 2014).

that this deliberation is more effective. Interestingly, we find no evidence that being in the presence of tempting purchases while selecting a subsidy, or having a waiting period prior to subsidy selection, influences subsidy choices themselves. The act of selecting prices does not appear to be subject to the same decision-making factors as the act of making selections from a particular price set.

Future research should explore the effectiveness of these non-restrictive interventions for promoting goal-oriented decisions in other domains. For example, wellness programs may be enhanced by presenting participants with a small choice set of incentive plans. Moreover, our agency intervention was framed as a choice between healthy and unhealthy subsidies. However, the behavioral mechanism that we believe drives this impact—a preference for consistency—does not necessarily imply that subsidy choices must be between healthy and unhealthy options. Shoppers also faced the same choice on all four Shopping Trips in the study. We believe there would be substantial value in future research to understand the use of agency to increase user buy-in, for example, by testing the impact of a subsidy choice between two healthy options or the influence of changes to the subsidy choice set over time.

References

Allcott, H., R. Diamond, J.-P. Dubé, J. Handbury, I. Rahkovksy, and M. Schnell (2019a): "Food Deserts and the Causes of Nutritional Inequality," *Quarterly Journal of Economics*, 134, 1793–1844.

Allcott, H., B. B. Lockwood, and D. Taubinsky (2019b): "Regressive sin taxes, with an application to the optimal soda tax," *The Quarterly Journal of Economics*, 134, 1557–1626.

——— (2019c): "Should We Tax Sugar-Sweetened Beverages? An Overview of Theory and Evidence," *Journal of Economic Perspectives*, 33, 202–227.

Allcott, H., S. Mullainathan, and D. Taubinsky (2014): "Energy policy with externalities and internalities," *Journal of Public Economics*, 112, 72–88.

Allcott, H. and D. Taubinsky (2015): "Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market," *American Economic Review*, 105, 2501–38.

Alvaredo, F., L. Assouad, and T. Piketty (2019): "Measuring inequality in the Middle East 1990-2016: The World's Most Unequal Region?" *Review of Income and Wealth*, 65, 685–711.

Andreoni, J., M. A. Kuhn, and L. Samuelson (2019): "Building Rational Cooperation on Their Own: Learning to Start Small," *Journal of Public Economic Theory*, 21, 812–825.

Ariely, D. and K. Wertenbroch (2002): "Procrastination, deadlines, and performance: Self-control by precommitment," *Psychological science*, 13, 219–224.

Bartlett, S., J. Klerman, P. Wilde, L. Olsho, C. Logan, M. Blocklin, M. Beauregard, and A. Enver (2014): "Evaluation of the Healthy Incentives Pilot (HIP) Final Report," U.S. Department of Agriculture, Food and Nutrition Service.

Bem, D. J. (1967): "Self-perception: An alternative interpretation of cognitive dissonance phenomena." *Psychological review*, 74, 183.

Bénabou, R. and J. Tirole (2004): "Willpower and personal rules," *Journal of Political Economy*, 112, 848–886.

——— (2011): "Identity, Morals, and Taboos: Beliefs as Assets *," *The Quarterly Journal of Economics*, 126, 805–855.

Bernheim, B. D., L. Braghieri, A. Martinez-Marquina, and D. Zuckerman (2019): "A Theory of Chosen Preferences," Working paper.

Bernheim, B. D. and D. Taubinsky (2018): "Behavioral public economics," *Handbook of Behavioral Economics: Applications and Foundations 1*, 1, 381–516.

Bhattarai, A. (2017): "Grocery stores are adapting to more male shoppers - whom they treat like knuckleheads," Washington Post.

Brownback, A. and S. Sadoff (2019): "Improving College Instruction through Incentives," Journal of Political Economy, Forthcoming.

Burlando, A., M. A. Kuhn, and S. Prina (2020): "Digital Credit Delivery Speed and Repayment Rates," Working paper.

Bushong, B., L. M. King, C. F. Camerer, and A. Rangel (2010): "Pavlovian processes in consumer choice: The physical presence of a good increases willingness-to-pay," *American Economic Review*, 100, 1556–71.

Butera, L., R. Metcalfe, W. Morrison, and D. Taubinsky (2019): "Measuring the welfare effects of shame and pride," Tech. rep., National Bureau of Economic Research.

Carrera, M., H. Royer, M. Stehr, J. Sydnor, and D. Taubinsky (2019): "Who Chooses Commitment? Evidence and Welfare Implications," Tech. rep., National Bureau of Economic Research.

Castner, L. and J. Henke (2011): "Benefit Redemption Patterns in the Supplemental Nutrition Assistance Program," United States Department of Agriculture Food and Nutrition Service.

Chetty, R., A. Looney, and K. Kroft (2009): "Salience and taxation: Theory and evidence," *American economic review*, 99, 1145–77.

Cole, N. and M. Fox (2008): "Diet quality of Americans by Food Stamp participation status: Data from the National Health and Nutrition Examination Survey, 1999–2004," U.S. Department of Agriculture, Food and Nutrition Service.

Dai, X. and A. Fishbach (2013): "When waiting to choose increases patience," *Organizational Behavior and Human Decision Processes*, 121, 256–266.

DeJarnette, P. (2018): "Temptation Over Time: Delays Help," Working paper.

DellaVigna, S. and U. Malmendier (2006): "Paying not to go to the gym," *american economic Review*, 96, 694–719.

Drewnowski, A. and S. Specter (2004): "Poverty and obesity: the role of energy density and energy costs," *The American Journal of Clinical Nutrition*, 79, 6–16.

Falk, A. and F. Zimmermann (2017): "Consistency as a Signal of Skills," *Management Science*, 63, 2197–2210.

——— (2018): "Information Processing and Commitment," *The Economic Journal*, 128, 1983–2002.

Farhi, E. and X. Gabaix (2020): "Optimal taxation with behavioral agents," *American Economic Review*, 110, 298–336.

Fishbach, A. and M. J. Ferguson (2007): "The goal construct in social psychology." in *Social psychology: Handbook of basic principles*, ed. by A. W. Kruglanski and E. T. Higgins, The Guilford Press, 490–515.

Freedman, J. L. and S. C. Fraser (1966): "Compliance without pressure: the foot-in-the-door technique." *Journal of personality and social psychology*, 4, 195.

Gabaix, X. and D. Laibson (2017): "Myopia and Discounting," National Bureau of Economic Research Working paper No. 23254.

Ghanem, D., S. Hirshleifer, and K. Ortiz-Becerra (2021): "Testing attrition bias in field experiments," .

Gilbert, D. T. and T. D. Wilson (2007): "Prospection: Experiencing the Future," *Science*, 317, 1351–1354.

Gneezy, A., A. Imas, A. Brown, L. D. Nelson, and M. I. Norton (2012): "Paying to Be Nice: Consistency and Costly Prosocial Behavior," *Management Science*, 58, 179–187.

Gneezy, A., A. Imas, and A. Jaroszewicz (2020): "The Impact of Agency on Time and Risk Preferences," *Nature: Communications*.

Gosnell, G. K., J. A. List, and R. D. Metcalfe (2020): "The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains," *Journal of Political Economy*, 128, 1195–1233.

Grüne-Yanoff, T. and R. Hertwig (2016): "Nudge versus boost: How coherent are policy and theory?" *Minds and Machines*, 26, 149–183.

Hahn, R., R. D. Metcalfe, D. Novgorodsky, and M. K. Price (2016): "The Behavioralist as Policy Designer: The Need to Test Multiple Treatments to Meet Multiple Targets," National Bureau of Economic Research Working paper No. 22886.

Hallsworth, M., J. A. List, R. D. Metcalfe, and I. Vlaev (2017): "The behavioralist as tax collector: Using natural field experiments to enhance tax compliance," *Journal of Public Economics*, 148, 14–31.

Harnack, L., J. M. Oakes, B. Elbel, T. Beatty, S. Rydell, and S. French (2016): "Effects of Subsidies and Prohibitions on Nutrition in a Food Benefit Program: A Randomized Clinical Trial," *Journal of the American Medical Association - Internal Medicine*, 176, 1610–1618.

Heffetz, O. (2018): "Are reference points merely lagged beliefs over probabilities?" Tech. rep., National Bureau of Economic Research.

Herman, D. R., G. G. Harrision, A. A. Afifi, and E. Jenks (2008): "Effect of a Targeted Subsidy on Intake of Fruits and Vegetables Among Low-Income Women in the Special Supplemental Nutrition Program for Women, Infants, and Children," *American Journal of Public Health*, 98, 98–105.

Houser, D., J. Liu, D. H. Reiley, and M. B. Urbancic (2021): "Checking out temptation: A natural experiment with purchases at the grocery register," *Journal of Economic Behavior and Organization*, 191, 39–50.

Hoynes, H., D. W. Schanzenbach, and D. Almond (2016): "Long-run impacts of childhood access to the safety net," *American Economic Review*, 106, 903–934.

Hsiaw, A. (2013): "Goal-setting and self-control," *Journal of Economic Theory*, 148, 601–626.

Imas, A., M. Kuhn, and V. Mironova (2018): "Waiting to Choose," Working Paper.

Johnson, E. J., S. B. Shu, B. G. Dellaert, C. Fox, D. G. Goldstein, G. Häubl, R. P. Larrick, J. W. Payne, E. Peters, D. Schkade, et al. (2012): "Beyond nudges: Tools of a choice architecture," *Marketing Letters*, 23, 487–504.

Jung, J. Y. and B. A. Mellers (2016): "American attitudes toward nudges." *Judgment & Decision Making*, 11.

Kaur, S., M. Kremer, and S. Mullainathan (2010): "Self-control and the development of work arrangements," *American Economic Review*, 100, 624–28.

——— (2015): "Self-control at work," *Journal of Political Economy*, 123, 1227–1277.

Klerman, J. A., S. Bartlett, P. Wilde, and L. Olsho (2014): "The Short-Run Impact of the Healthy Incentives Pilot Program on Fruit and Vegetable Intake," *American Journal of Agricultural Economics*, 95, 1372–1382.

Koch, A. K. and J. Nafziger (2011): "Self-regulation through goal setting," *Scandinavian Journal of Economics*, 113, 212–227.

Koenig, C. and D. Schindler (2018): "Dynamics in Gun Ownership and Crime -Evidence from the Aftermath of Sandy Hook," Working paper.

Koszegi, B. and M. Rabin (2008): "Reference-Dependent Consumption Plans," Working paper version of 2009 AER publication.

Laibson, D. (1997): "Golden eggs and hyperbolic discounting," *Quarterly Journal of Economics*, 112, 443–477.

——— (2015): "Why don't present-biased agents make commitments?" *American Economic Review Papers and Proceedings*, 105, 267–272.

Lee, D. S. (2009): "Training, wages, and sample selection: Estimating sharp bounds on treatment effects," *The Review of Economic Studies*, 76, 1071–1102.

Levitt, S. D., J. A. List, S. Neckermann, and S. Sadoff (2016): "The behavioralist goes to school: Leveraging behavioral economics to improve educational performance," *American Economic Journal: Economic Policy*, 8, 183–219.

Lockwood, B. B. (2020): "Optimal income taxation with present bias," *American Economic Journal: Economic Policy*, 12, 298–327.

Luca, M., D. Malhotra, and C. Poliquin (2017): "Handgun waiting periods reduce gun deaths," *Proceedings of the National Academy of Sciences*, 114, 12162–12165.

O'Donoghue, T. and M. Rabin (1999): "Doing it now or later," *American Economic Review*, 89, 103–124.

Olsho, L. E., J. A. Klerman, P. E. Wilde, and S. Bartlett (2016): "Financial incentives increase fruit and vegetable intake among Supplemental Nutrition Assistance Program participants: a randomized controlled trial of the USDA Healthy Incentives Pilot," *American Journal of Clinical Nutrition*, 104, 423–435.

Piketty, T. and E. Saez (2014): "Inequality in the long run," *Science*, 344, 838–843.

Read, D. and B. van Leeuwen (1998): "Predicting Hunger: The Effects of Appetite and Delay on Choice," *Organizational Behavior and Human Decision Processes*, 76, 189–205.

Rehm, C., J. Peñalvo, A. Afshin, and D. Mozaffarian (2016): "Dietary intake among us adults, 1999-2012," *Journal of the American Medical Association*, 315, 2542–2553.

Sadoff, S., A. Samek, and C. Sprenger (2020): "Dynamic Inconsistency in Food Choice: Experimental Evidence from Two Food Deserts," *Review of Economic Studies*, forthcoming.

Schilbach, F. (2019): "Alcohol and self-control: A field experiment in India," *American economic review*, 109, 1290–1322.

Schwartz, J., D. Mochon, L. Wyper, J. Maroba, D. Patel, and D. Ariely (2014): "Healthier by precommitment," *Psychological science*, 25, 538–546.

Schwartz, M. B. (2017): "Moving Beyond the Debate Over Restricting Sugary Drinks in the Supplemental Nutrition Assistance Program," *American Journal of Preventive Medicine*, 52, S199 – S205.

Smith, J. J., P. J. Morgan, R. C. Plotnikoff, K. A. Dally, J. Salmon, A. D. Okely, T. L. Finn, and D. R. Lubans (2014): "Smart-phone obesity prevention trial for adolescent boys in low-income communities: the ATLAS RCT," *Pediatrics*, 134, e723–e731.

Thakral, N. and L. T. To (2020a): "Anticipation and Consumption," Working paper.

——— (2020b): "Anticipation and Temptation," Working paper.

Thaler, R. H. and C. R. Sunstein (2009): *Nudge: Improving decisions about health, wealth, and happiness*, Penguin.

USDA, U. S. D. O. A. (2021): "What's on your plate?" .

Volpp, K. G., L. K. John, A. B. Troxel, L. Norton, J. Fassbender, and G. Loewenstein (2008): "Financial incentive–based approaches for weight loss: a randomized trial," *Journal of the American Medical Association*, 300, 2631–2637.

Wheeler, M. A., D. T. Stuss, and E. Tulving (1997): "Toward a theory of episodic memory: The frontal lobes and autonoetic consciousness." *Psychological Bulletin*, 121, 331–354.

Wilde, P., J. A. Klerman, L. E. Olsho, and S. Bartlett (2015): "Explaining the impact of USDA's Healthy Incentives Pilot on different spending outcomes," *Applied Economic Perspectives and Policy*, 38, 655–672.

Wolf, A. M. (2012): "What Is the Economic Case for Treating Obesity?" *Obesity Research*, 6, 2S–7S.

A Online Appendix

A.1 Classical Utility Maximization

Consider the consumers utility maximization problem. Suppose preferences over FV and BG spending are given by a Cobb-Douglas utility function: $U(FV, BG) = FV^{\alpha} \cdot BG^{1-\alpha}$. Demand for FV is given by $FV^* = \alpha \cdot \frac{M}{p_F}$, where p_F is the price of FV and $\$M$ is the money dedicated to spending on the two goods. When we switch from subsidizing FV to offering a choice between FV and BG subsidies, one of two things will happen. Either the shopper selects the FV subsidy, and p_F and FV^* are unchanged, or the shopper selects the BG subsidy, which increases p_F, and thus decreases FV spending. This optimization process is demonstrated graphically in Appendix Figure A.1. Here, we show these budgets and the utility-maximizing indifference curves for two individuals: one with $\alpha = 0.75$ who will choose the FV subsidy, and one with $\alpha = 0.25$ who will choose the BG subsidy. In Panel A we show how both individuals behave when faced with the FV subsidy, and in Panel B, we show how they respond to the subsidy choice. The consumer with $\alpha = 0.25$ selects the BG subsidy, and in doing so, substitutes to reduce FV spending from $F2$ to $F3$.

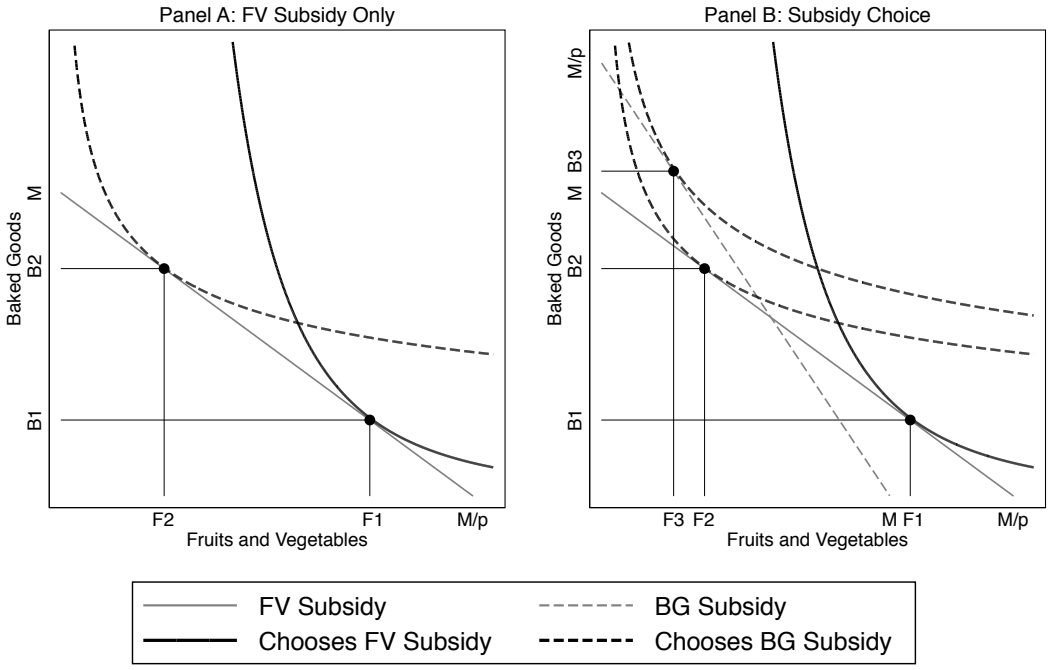

Figure A.1. Cobb-Douglas Utility Maximization Behavior from Subsidy Budgets

To consider a situation in which the consumer selecting the BG subsidy might increase their FV spending as a result, we need a utility function that allows for more complementarity. Consider the CES utility function: $U(FV, BG) = (\alpha FV^{\gamma} + (1 - \alpha)BG^{\gamma})^{\frac{1}{\gamma}}$. Here, $\sigma = \frac{1}{1-\gamma}$ is the elasticity of substitution and α is the share parameter.

We use $M = 12$, which is roughly the average sum of FV and BG spending for shoppers that do not receive a subsidy, and the subsidy price of \$0.70 (as opposed to an unsubsidized price of \$1) that we use in the study.

The set of parameters that allows subsidy choice to result in increased FV spending is small and shown in Figure A.2. Shoppers will not select the BG subsidy unless $\alpha < 0.5$. However, there is no non-zero elasticity of substitution for which the selection of the BG subsidy will increase FV spending until the share parameter falls to about 0.1. Once the share parameter is in that range, low elasticities will permit this, although there is no non-zero share parameter for which this will happen if the elasticity of substitution is 0.5 or higher.

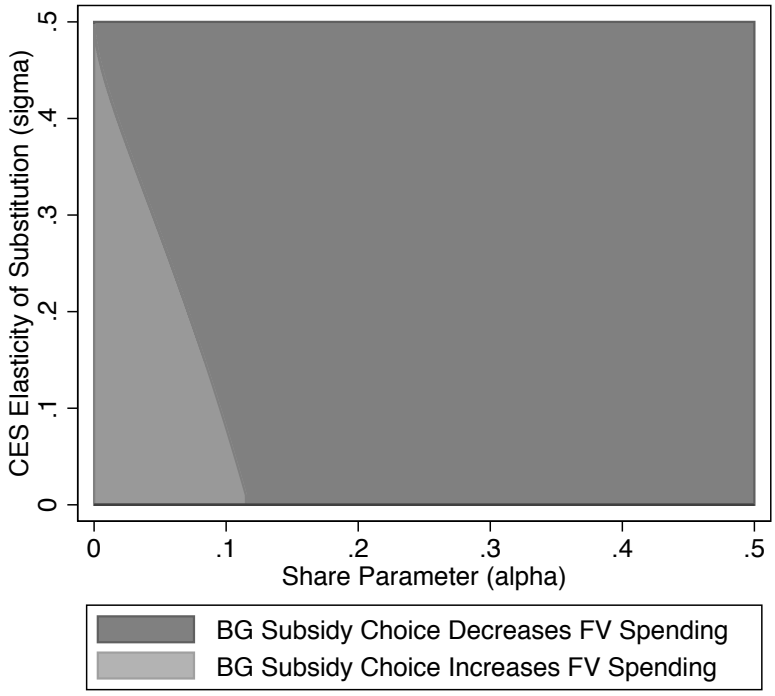

Figure A.2. Impact of Subsidy Choice on FV Spending by CES Parameters

A.2 Impact of Treatments on Total Food Spending

To consider the income effect, we look at the total spending on food items in the pooled subsidy groups. We do not observed a statistically significant spending increase, meaning that our effects are largely driven by substitution towards FV.

Table A.1. Effect of Subsidy Treatments on Total Food Spending ($)

Sample:	All Trips			Trip 1	Completers
Control Mean [SD]: 51.37 [50.21]	(1)	(2)	(3)	(4)	(5)
Subsidy	0.99	0.99	1.50	3.70	1.67
	(3.83)	(3.87)	(3.98)	(6.68)	(3.83)
Baseline Survey Food Spending ($)		0.12***	0.12***	0.09**	0.13***
		(0.03)	(0.03)	(0.04)	(0.03)
Part 1, Part 2 Fixed Effects	Y	Y	Y	Y	Y
Wave Fixed Effects	N	N	Y	Y	Y
Observations	2767	2767	2767	781	2456
Clusters	805	805	805	N/A	625

*** $\Rightarrow p < 0.01$, ** $\Rightarrow p < 0.05$. Standard errors in parentheses are clustered at the shopper level. All specifications are linear random effects models except Column (4), which is OLS. "Subsidy" pools all treatments featuring subsidies whether or not they also feature behavioral interventions ($T_1 - T_5$ from Table 1). "Baseline Survey Food Spending" is captured based on shopping receipts submitted on the baseline survey. Column (4) restricts the sample to only spending on the first Shopping Trip. Column (5) restricts the sample to spending across all Shopping Trips but only for shoppers who completed all four Shopping Trips in the study.

A.3 Impact of Treatments on BG Spending and Spending Fractions

Appendix Table A.2 shows in Columns (1)-(3) that the impact of allowing BG subsidy choice on BG spending ranges from $1.07 to $1.51 across specifications. The difference-in-differences impact of agency by BG vs. FV spending is not statistically significant, despite the natural prior here would be that allowing for the BG subsidy increases BG spending by more than it increase FV spending.

A consequence of the increases in both FV and BG spending that result from agency and waiting periods is that there is a muted effect of these interventions on the fraction of food spending devoted to FV. Columns (4)-(6) of Table A.2 shows these estimates. Indeed, to the extent that agency or waiting periods increase the fraction of spending on FV, it appears to occur through the Waiting Period (Early Choice) treatment, which has a positive and significant effect on the fraction of food spending on FV.

Table A.2. Effect of Treatments on BG Spending ($) and Spending Fractions

	Total BG Spending			Fraction Spent on FV		
	(1)	(2)	(3)	(4)	(5)	(6)
Subsidy	-1.06	-0.25	-1.05	0.150***	0.152***	0.150***
	(1.02)	(0.94)	(1.02)	(0.023)	(0.018)	(0.023)
Agency	1.07**		1.30*	0.007		0.004
	(0.66)		(0.68)	(0.022)		(0.024)
Waiting Period		1.02**	0.65		0.006	0.005
		(0.43)	(0.45)		(0.016)	(0.018)
Baseline BG Spending ($)	-0.01	-0.01	-0.01			
	(0.01)	(0.01)	(0.01)			
Baseline FV Fraction (%)				0.157***	0.158***	0.158***
				(0.043)	(0.044)	(0.044)
Control Mean [SD]:	8.00 [16.15]			12.9% [21.6]		
Part 1, Part 2 Fixed Effects	Y	Y	Y	Y	Y	Y
Wave Fixed Effects	Y	Y	Y	Y	Y	Y
Observations	2767	2767	2767	2710	2710	2710
Clusters	805	805	805	793	793	793

*** $\Rightarrow p < 0.01$, ** $\Rightarrow p < 0.05$, * $\Rightarrow p < 0.10$. Standard errors in parentheses are clustered at the subject level. All specifications are linear random effects models. "Subsidy" pools all treatments featuring subsidies whether or not they also feature behavioral interventions ($T_1 - T_5$ from Table 1). "Agency" pools all treatments featuring a choice between subsidies ($T_2 - T_4$ from Table 1). "Waiting Period" pools all treatments featuring a waiting period between a pre-shopping survey and the shopping trip ($T_3 - T_5$ from Table 1). "Baseline Survey FV Spending" is captured based on shopping receipts submitted on the baseline survey. In Columns (4)-(6), we lose 22 observations and 1 subject relative to the count in Table 6 because of receipts with no food items. We lose another 35 observations and 11 subjects due to baseline survey receipts with no food items.

A.4 Main Results Excluding Post-re-weight Data

In order to prioritize our power to detect agency and waiting periods effects (as well as differences in subsidy choice rates between treatment groups) we re-weighted our randomization to favor those treatments after Wave 6 within Study 1 (at the expense of Control and Restricted Subsidy). Here, we show version of our main results with the data collected under the adjusted randomization (Waves 7 and 8) excluded. Table A.3 shows the main specifications from Tables 4-7 (Column (3) in Tables 4-6, and Column (1) in Table 7) estimated without those waves. Results are very similar to those with Waves 7 and 8 included.

Table A.3. Main Results with Waves 7 and 8 Excluded

	Table 4	Table 5	Table 6	Table 7
Subsidy	6.83***	0.16***	5.32***	
	(0.69)	(0.02)	(0.90)	
Agency			0.99	
			(1.00)	
Waiting Period			1.65**	
			(0.77)	
T_3: Waiting Period (Delayed Choice)				2.50**
				(1.21)
T_4: Waiting Period (Early Choice)				3.38***
				(1.19)
Baseline Survey Dependent Variable	0.17***	0.18***	0.17***	0.08
	(0.06)	(0.05)	(0.06)	(0.08)
Part 1, Part 2 Fixed Effects	Y	Y	Y	Y
Wave Fixed Effects	Y	Y	Y	Y
Observations	2359	2308	2359	875
Clusters	678	667	678	261

*** $\Rightarrow p < 0.01$, ** $\Rightarrow p < 0.05$. Standard errors in parentheses are clustered at the shopper level. All specifications are linear random effects models. For Tables 4-6 we use the specifications from column (3), and for Table 7 we use the specification from column (1). Refer to the relevant tables in the main text for additional detail on each specification.

A.5 Impact of Treatments by Shopping Trip

We expect the treatment effects to be largest during the first shopping trip, and indeed they are. Figure A.3 reproduces Panel C of Figure 2 separately for each shopping trip using coefficients from a single model where treatment indicator variables from Study 1 are interacted with shopping trip indicator variables. We also include un-interacted shopping-trip indicator variables to allow FV spending in Control to vary by trip. We

exclude the random effects to allow each week to be fully independent but maintain the control for baseline FV spending and wave fixed effects. Unlike Panel C of Figure 2, we leave the vertical axis in $ to avoid showing percentage effects with a base that shifts across panels. Mean control FV spending fluctuates non-monotonically: $4.63 in in Trip 1, $3.33 in Trip 2, $4.61 in Trip 3, and $3.48 in Trip 4. Therefore, in percentage terms, the smaller absolute effects in Trip 2 are actually similar to those in Trip 1, whereas the even smaller absolute effects in Trip 3 are substantially smaller than the others, and the effect sizes rebound somewhat in Trip 4.

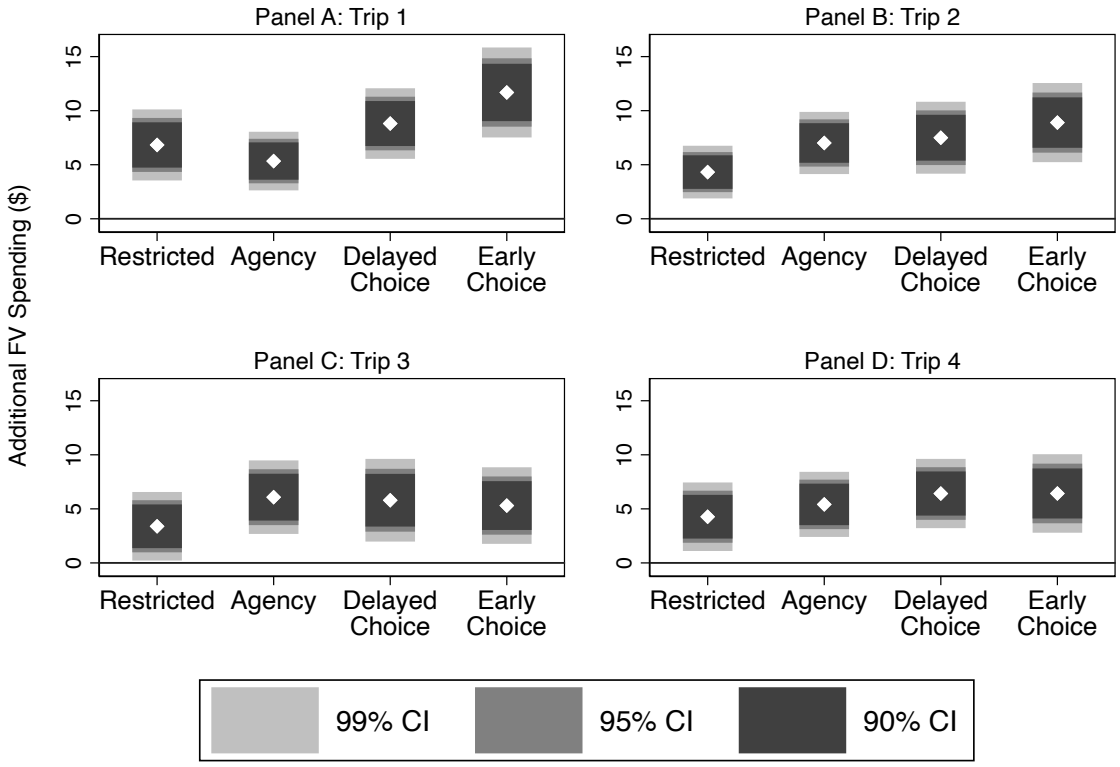

Figure A.3. Relative-to-Control FV Treatment Effects, Study 1

A.6 FV Spending of Shoppers Choosing BG Subsidy

We find that the average shopper in the agency treatment—pooled across those choosing the FV and BG subsidies—spends more on FV than without the subsidy or subsidy choice. While we focus on this pooled group for identification purposes, it is important to consider the 22% of households that select the less healthy BG subsidy; are they worse off in terms of FV spending than they were without the subsidy? If so, a subsidy choice program might not be desirable even if it has a positive average effect. In Figure A.4, we show the CDFs of FV spending in the control treatment and the agency treatments.

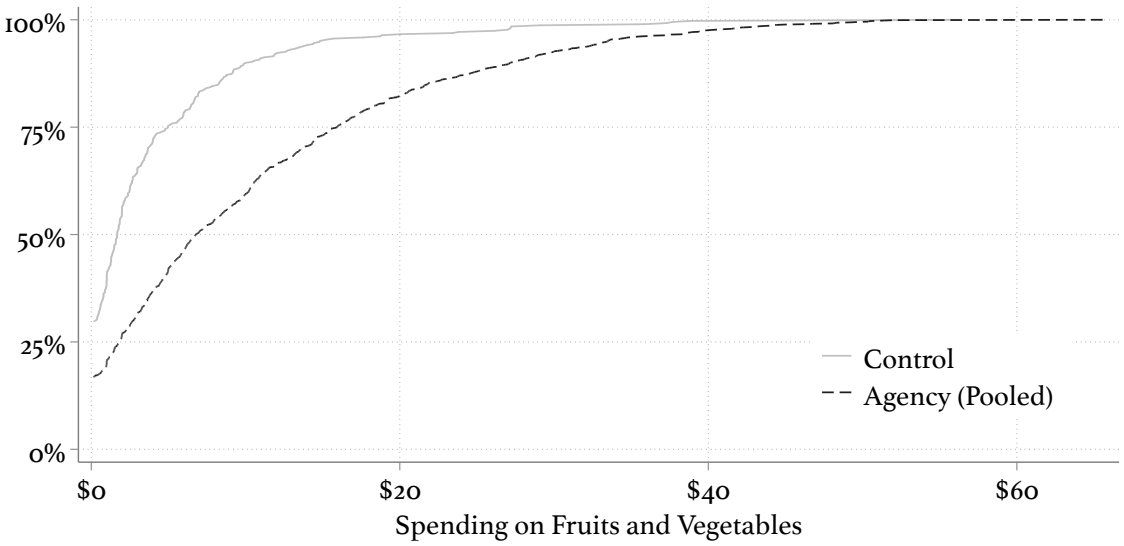

Figure A.4. Cumulative Distribution of FV Spending

The agency treatments distribution dominates the control treatment distribution, indicating that the BG subsidy does not lead to lower FV spending than would prevail without any subsidy. In other words, if the BG subsidy leads to a substitution effect that lowered FV spending relative to control, and we assume that selection into the BG subsidy is monotonic in the FV spending level, the CDFs should cross at a density of 0.22: the fraction of the sample in the agency treatments selecting the BG subsidy. However, this is not the case.

A.7 Differential Effects by Welfare Recipient Status

While all shoppers in our study are from households with income under 185% of the FPL, that range still permits considerable variation across households in terms of wealth. If the subsidies and behavioral interventions are ineffective for the neediest households, especially those already receiving food assistance, then the scope for using these subsides is smaller. While we do not have income information for many shoppers in the sample, we can measure whether the shopper is a participant in a welfare program. There are two ways we can do this. First, shoppers are asked if they receive SNAP funds in the initial screening. Second, we can observe whether any transaction was paid for using Electronic Benefit Transfer (EBT) in our study. SNAP recipients may not use EBT every time they shop, and EBT can be used to access funds from programs other than SNAP. Therefore, we define a welfare recipient as someone who falls into either category. 55% of shoppers in our sample are identified as welfare recipients.

We find that welfare recipients are roughly 8% less likely to select the FV subsidy in the agency treatments ($p = 0.020$).[50] To evaluate whether this difference in choice

[50]Estimate is from a linear regression of a shopper's mean subsidy choice on their welfare-recipient

rates is substantial enough to make our interventions less effective at increasing FV spending, in Table A.4 we re-estimate a number of important treatment effects in the paper, allowing for a heterogeneous impact by welfare recipient status.

First, we note that when we split the sample in half, we lose power to identify strong, statistically significant effects of agency and waiting periods for just the non-welfare recipients, even though the effects remain sizeable. However, for welfare recipients, the effects are slightly larger, and statistically significant; agency increases FV spending by \$2.11 ($p = 0.019$, Column (2)) on average, and a waiting period increases FV spending by \$1.93 ($p = 0.020$, Column (3)). Without agency or a waiting period, welfare-recipients households spend roughly \$1 less on FV than non-recipients per trip, but this difference is not statistically significant. The subsidy effect remains very large and statistically significant in Column (1), and while it is slightly less effective for welfare recipients, the difference in effects is not significant. Overall, there is no significant evidence of differential effects by welfare-recipient status.

status with study part and randomization wave fixed effects, and a control for baseline FV spending, with robust standard errors.

Table A.4. Heterogeneous Effect of Interventions on FV Spending by Welfare Status

Sample:	All	Subsidy	
	(1)	(2)	(3)
Welfare Recipient	0.023	-1.037	-1.020
	(1.058)	(0.979)	(0.835)
Subsidy	7.178***		
	(0.023)		
Subsidy × Welfare	-0.927		
	(1.254)		
Agency		1.693	
		(1.189)	
Agency × Welfare		0.420	
		(1.303)	
Waiting Period			1.522*
			(0.914)
Waiting Period × Welfare			0.409
			(1.245)
Baseline FV Spending ($)	0.146***	0.160***	0.162***
	(0.051)	(0.059)	(0.057)
Part 1, Part 2 Fixed Effects	Y	Y	Y
Wave Fixed Effects	Y	Y	Y
Observations	2767	2379	2379
Clusters	805	700	700

*** $\Rightarrow p < 0.01$, * $\Rightarrow p < 0.10$. Standard errors in parentheses are clustered at the subject level. All specifications are linear random effects models. "Subsidy" pools all treatments featuring subsidies whether or not they also feature behavioral interventions ($T_1 - T_5$ from Table 1). "Agency" pools all treatments featuring a choice between subsidies ($T_2 - T_4$ from Table 1). "Waiting Period" pools all treatments featuring a waiting period between a pre-shopping survey and the shopping trip ($T_3 - T_5$ from Table 1). "Baseline Survey FV Spending" is captured based on shopping receipts submitted on the baseline survey.

A.8 Impact of Treatments on Shopping Trip Timing

One pathway by which the treatments could increase FV spending is by inducing shoppers to wait longer between Shopping Trips in order to increase FV spending *per* Shopping Trip. To evaluate this concern, we consider the amount of time that has elapsed prior to the Shopping Trip in question. Note that this value can only be calculated for Shopping Trips 2-4, since we do not know the timing of the Shopping Trip that preceded the first Shopping Trip of the study. In Table A.5, we show that the treatments have a negligible impact on the length between each Shopping Trip.

Table A.5. Interval Length Between Shopping Trips by Treatment

	Part 1: Mar–Jul, 2018						Part 2: Mar–Apr, 2019		
	Control	Restricted	Agency	Waiting Period (Delayed)	Waiting Period (Early)	F-Test	Restricted	Waiting Period (No Agency)	F-Test
Days between Shopping Trips	11.14	12.40	10.53	12.29	11.79	0.34	8.31	8.40	0.81
Observations	283	253	337	247	236		330	262	

F-test conducted as a joint test of equality across all treatments (standard errors clustered by shopper), with the p-values reported.

For a visualization of the timing between Shopping Trips, we present a histogram of Shopping Trip intervals below in Figure A.5. The majority of Shopping Trip intervals are within a week with certain long intervals wherein shoppers did not participate in an observed Shopping Trip.

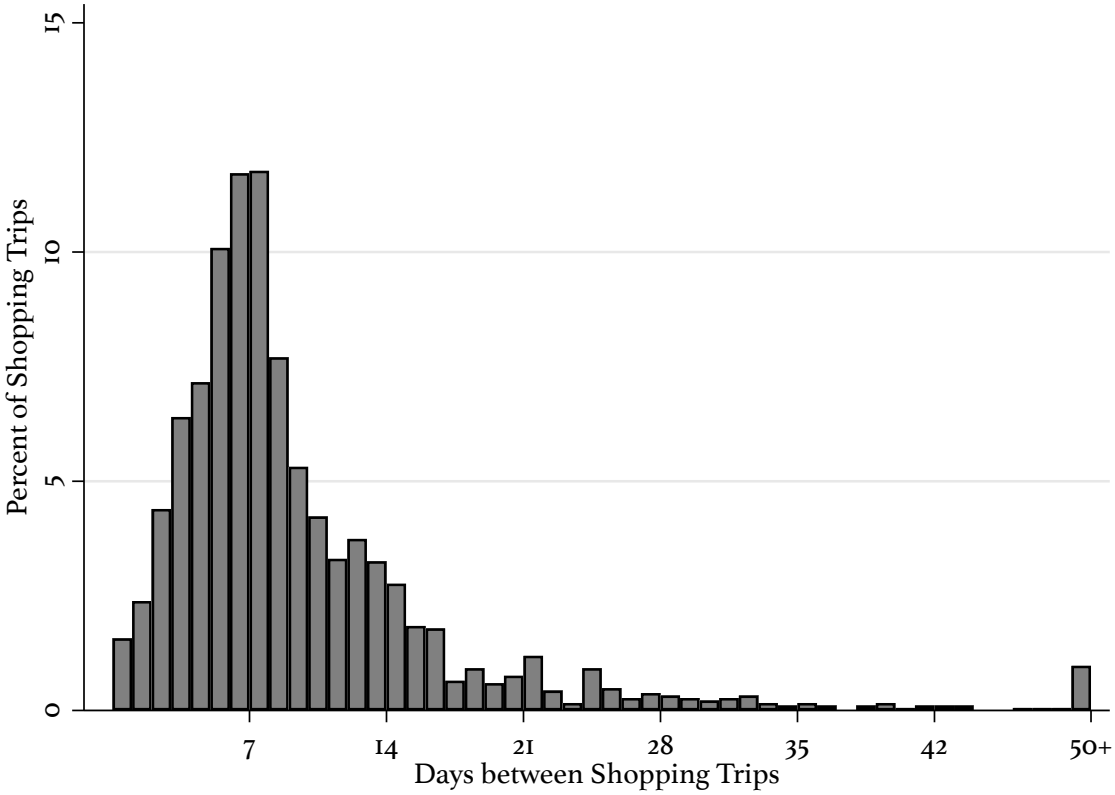

Figure A.5. Intervals between Shopping Trips (in Days)

A concern that is particular to the waiting period treatments (both delayed-choice and early-choice) is that the 4-48 hour mandatory waiting periods changed when people shopped—both hour of the day, and day of the week—in a way that interacted with preferences, food shopping location, or item availability. We use Kolmogorov-Smirnov tests to evaluate whether the implementation of a waiting period affects the distribution of shopping times and days. While shoppers with a waiting period appear to shop slightly earlier in the day and later in the week, the differences in distributions are not statistically significant even without any adjustment for repeated observations withing an individual ($p = 0.176$ for hour of day and $p = 0.144$ for day of week). Such an adjustment is not straightforward with a distributional test, as taking an individual-level mean or median can substantially distort time/day variables (e.g., someone who only ever shops at six a.m. and six p.m. never shops at noon, their average shopping hour). Therefore, we stick with this test as a likely lower-bound on the proper p-values.

B Instructions

In this section, we include the instructions and protocols for each of the surveys that the shoppers may have seen. For the baseline survey, that protocol includes the elicitation of time preferences. Assignment occurred after the baseline survey. The first instructions shoppers received for their assigned treatment were given at the beginning of their survey for Shopping Trip 1.

All shoppers saw an introduction identical to the control group. The instructions for submitting pictures of receipts was also identical across treatments. Where the individual treatments differed from the control group, we will indicate in their subsection.

B.1 Baseline

Hi Agents! This is the first survey of the 4-6-week grocery shopping study. You can take this survey anywhere. You are required to have the following BEFORE accepting this survey:

- A recent itemized receipt from a recent grocery shopping trip you took to a store with a produce section.
- A blank piece of paper and pencil/pen.

When you have both required items available, you can begin the survey. There are 10 questions and 2 photos.

Have fun!

Thank you for your interest in our study. Our study seeks to investigate grocery choices. The study will cover 6 weeks and will require that you submit receipts from your grocery purchases over that timeframe. You will submit these receipts in the form of a photo attached to a Field Agent Job. This Job may involve answering a series of questions both during your shopping trip and in between shopping trips.

The payment for participation is $30 total ($1 per survey with the balance paid at the end with the successful completion of all study requirements.

In addition to the participation payments, the majority of participants will have the opportunity to receive up to $40 in free groceries.

It is very important that all participants complete all 6 weeks of the study, so please dont agree to participate unless you have time to finish the entire study.

As a bonus for completing this study, we have included three "Bonus Questions" that offer you a chance to earn additional money. We will randomly choose agents to receive this bonus money, so payment for this question is NOT guaranteed.

For these Bonus Questions, we are going to show you 3 different scenarios (2 now, and 1 at the end of this survey) and ask you to select your preferred option in each scenario.

For each Bonus Question you answer today, you will have about a 1 in 50 chance of winning the amount determined by your selection. So, treat each Bonus Question as if it will determine your actual bonus payment.

Additional money earned from any Bonus Question will be deposited directly into your Field Agent account. However, you will have a choice about when to receive this bonus.

As a reminder, you will be paid the full $30 for successfully completing the full study no matter your answers to the bonus questions and whether or not you are selected for bonus payment.

1. **Bonus Question #1:** In this question, your bonus grows larger the longer you wait for it.

   The earliest you can choose to receive your bonus is 1 week from today. If you choose to receive it 1 week from today, it will be $50. If you choose to wait longer to receive your bonusup to a maximum of 27 weeks from todayit will grow by some amount.

   Below, there are a number of combinations of waiting times and bonus amounts that you can choose.

   Which is your preferred option?

   (a) Receive $50 in 1 week
   (b) Wait 2 weeks, receive $53
   (c) Wait 3 weeks, receive $54
   (d) Wait 5 weeks, receive $55
   (e) Wait 7 weeks, receive $56
   (f) Wait 9 weeks, receive $57
   (g) Wait 11 weeks, receive $58
   (h) Wait 13 weeks, receive $59
   (i) Wait 16 weeks, receive $60
   (j) Wait 19 weeks, receive $61
   (k) Wait 23 weeks, receive $62
   (l) Wait 27 weeks, receive $63

2. **Bonus Question #2:** This question is similar to Bonus Question #1, except shifted by a week.

   Here, you have the option of receiving the bonus today. Below are the combinations of waiting times and bonus amounts that you can choose from.

   Which is your preferred option?

   (a) Receive $50 right away
   (b) Wait 1 week, receive $53
   (c) Wait 2 weeks, receive $54
   (d) Wait 4 weeks, receive $55

(e) Wait 6 weeks, receive $56
(f) Wait 8 weeks, receive $57
(g) Wait 10 weeks, receive $58
(h) Wait 12 weeks, receive $59
(i) Wait 15 weeks, receive $60
(j) Wait 18 weeks, receive $61
(k) Wait 22 weeks, receive $62
(l) Wait 26 weeks, receive $63

Great! For the next few questions, we want you to think about the group of people with whom you share a budget for food.

If you are a member of a family living together, this is probably the entire household. If you live with roommates and dont pool money together for food, this would mean just you.

In general, think of the entire group you plan for when you go grocery shopping. We will refer to this group as your food household from now on. We are now going to ask some questions about your food situation. Please provide an answer as it relates to your food household.

3. Which of these statements best describes the food eaten in your food household in the last 30 days?

   (a) Enough of the kinds of food we want to eat
   (b) Enough, but not always the kinds of food we want to eat
   (c) Sometimes not enough to eat
   (d) Often not enough to eat

4. In the last 30 days, how often did you worry about whether your food would run out before you got money to buy more?

   (a) Almost always
   (b) Most of the time
   (c) About half of the time
   (d) Some of the time
   (e) Almost never

5. In the last 30 days, how often did you feel like your food household couldnt afford to eat well-balanced (healthy) meals because you couldnt afford it?

   (a) Almost always
   (b) Most of the time
   (c) About half of the time

(d) Some of the time
(e) Almost never

6. Do you think your food household eats the right amount of fruits and vegetables?

   (a) Yes, we eat the right amount
   (b) No, we should eat more
   (c) No, we should eat less

We are now going to ask some questions about your health situation.

7. How would you rate your physical health status?

   (a) Excellent
   (b) Very good
   (c) Good
   (d) Fair
   (e) Poor

8. Thinking about the past couple weeks, how often do you find that you have difficulty maintaining energy, focus, or attention?

   (a) Almost always
   (b) Most of the time
   (c) About half of the time
   (d) Some of the time
   (e) Almost never

Food Diary: You will need a piece of paper and a pencil for this activity. We would like to know what a typical day of food consumption is like for you. If there are any children in your food household, then we would also like to know what a typical day of consumption is like for one child. To make this easy, just think about the food you ate yesterday.

On a piece of paper, create four sections for yesterdays food: breakfast, lunch, dinner, and snacks. In each section, try to remember and list everything you ate. If you know the quantity you ate, please include it. If you do not remember the quantity, please still list the item.

If there are any children in your food household, please make a separate chart for one child.

See the example below, of what a food diary might look like. Once you complete your food diary, please take a readable photo of it. Make sure your photo is clear and in focus. Make sure that we can read your food diary entries.

My Food Diary

Breakfast	Lunch	Dinner	Snacks
Cheerios and milk	Pizza (2 slices)	Grilled chicken	Banana
Apple	Breadsticks	Baked potato	Popcorn (1 bag)
Coffee	Coke (1 can)	Mac and Cheese	
		Side salad with Ranch	
		Iced Tea	

Figure B.6. Example Food Diary

Shopping Receipt: In your study confirmation email, we asked you to begin saving your receipts from your grocery shopping trips. Please take a photo of an itemized receipt from your most recent Stock-Up grocery-shopping trip.

As a reminder, a "Stock-Up" shopping trip is a trip to a grocery storesomewhere that has a produce section with multiple aisleswhere you purchased food or ingredients to prepare multiple meals at home.

Some additional requests:

- If your receipt is too long to fit into a readable photo, please neatly tear it into two or more sections to put next to one another for the photo.
- Make sure your photo is clear and in focus. Make sure that we can read the items listed on the receipt.
- Please take your picture with a vertical (portrait) orientation.

9. **Bonus Question #3:** Now that youve had a little more time to think about Bonus Question #2, you have another chance to make a selection from the same set of options in Bonus Question #2. This will not replace your answer to Bonus Question #2: it is a different question and your choice from it may count separately from Bonus Question #2.

   Recall that, for this question, a $50 bonus is available today. Also recall that if you wait to receive your bonusup to a maximum of 26 weeks from todayit will increase as you wait longer.

   Below are your options with different waiting times and bonus amounts that you can choose.

   Taking time to think about it, which is your preferred option?

   (a) Receive $50 right away
   (b) Wait 1 weeks, receive $53
   (c) Wait 2 weeks, receive $54
   (d) Wait 4 weeks, receive $55
   (e) Wait 6 weeks, receive $56
   (f) Wait 8 weeks, receive $57
   (g) Wait 10 weeks, receive $58

(h) Wait 12 weeks, receive $59
(i) Wait 15 weeks, receive $60
(j) Wait 18 weeks, receive $61
(k) Wait 22 weeks, receive $62
(l) Wait 26 weeks, receive $63

That is all the questions we have for you today. In the next few days, you will see the first grocery-shopping survey of the study in your Field Agent app under the heading "Just For You." This survey will contain all of the instructions you will need to complete the study.

Make sure to OPEN AND READ THE INSTRUCTIONS for your next survey BEFORE your next grocery-shopping trip. This survey will include all instructions for the study. You do not have to alter your shopping schedule for this survey, just remember to begin your next survey before your next usual grocery shopping trip.

Remember that the store you go to MUST:

- Have a produce section with multiple aisles of fruits and vegetables.
- Provide itemized receipts.

If you have any questions or concerns about this study please email us at Support@Fieldagent.net. Put "Grocery Shopping Study" in the subject line.

See you soon!

B.2 Control

B.2.1 Common Introduction

Hi Agents!

Take this survey the next time you go on a grocery-shopping trip.

Remember that the store you go to must:

- Have a produce section with multiple aisles of fresh fruits & vegetables
- Provide an itemized receipt

WAIT until you arrive at the store to do this survey. If you are not at the grocery store right now, close the survey and re-open it when you arrive. If you are at the grocery store right now, please continue with the survey.

You cannot go backwards in the study, so read and answer all questions carefully.

Have fun!

B.2.2 Common Receipt Submission

Now, go ahead and complete your shopping trip. The next screen will ask for a picture of your itemized grocery receipt. Take a picture of the receipt before going home form the store (its fine to be in your car or the parking lot)

When you have finished shopping and have paid for your purchases continue to the next screen.

PHOTO 1: Take a photo of your itemized grocery shopping receipt while still at the store. You can be in your car, but DO NOT LEAVE THE LOT.

- The receipt must be clear and in focus Your reimbursement is dependent upon the receipt being clear and in focus.
- We must be able to read all of the items that you purchased today.
- Make sure we can see the name of the store you shopped at.
- Make sure we can see the date and time of your purchase.
- Please take the picture with a vertical orientation.
- If the receipt is too big to capture in one photo, carefully tear the receipt and arrange it so that it can all fit into 1 picture.

Take a picture of your itemized grocery receipt.

That is all the questions we have for you today.

You must wait at least 5 days to complete your next grocery shopping trip survey.

If you have any questions, email us at Support@fieldagent.net before you submit this survey. Put "Grocery Shopping Study" in the subject line.

B.3 Restricted

[Begin with Common Introduction]

B.3.1 Common Restricted Subsidy

For this grocery-shopping trip, we will reimburse you for some of your groceries. Please read the instructions carefully.

Any item in the fruits and vegetables category will qualify for the partial reimbursement. Fruits and vegetables are defined below:

- Fresh fruits and vegetables
- Canned fruits and vegetables
- Frozen fruits and vegetables
- Processed or prepared products that simply contain some fruits and vegetables do not qualify. For example, frozen fruit qualifies, but frozen juice concentrate does not because of the added sugar. Cabbage qualifies, but prepared coleslaw does not because of the added mayonnaise. Canned tomatoes qualify but canned pasta sauce does not because of the added salt.

You will be reimbursed 30% of the price on all items on your receipt in the fruits and vegetables category, up to a maximum reimbursement of $10.

For example:

- If you spend $30 on fruits and vegetables, you will receive a reimbursement of (30% of $30) = $9.
- If you spend $25 on fruits and vegetables, you will receive a reimbursement of (30% of $25) = $7.50.
- If you spend $33.33 or on fruits and vegetables, you will receive the maximum reimbursement of $10.

Your reimbursement will be credited to your Field Agent account within 7 days of this shopping trip. We will use the receipt you submit at the end of this survey to calculate your reimbursement.

Do you understand?

- Yes [Proceed to next question]
- No

[Continue with Common Receipt Submission]

B.4 Agency

[Begin with Common Introduction]

B.4.1 Common Choice Procedure

For this grocery-shopping trip, we will reimburse you for some of your groceries. You will need to choose which types of foods to be reimbursed for.

Please read the instructions carefully.

There are two food categories for you to choose from:

1. Fruits and vegetables
2. Baked goods

Heres a list of what falls into each category:
Fruits and Vegetables:

- Fresh fruits and vegetables
- Canned fruits and vegetables
- Frozen fruits and vegetables
- Processed or prepared products that simply contain some fruits and vegetables do not qualify. For example, frozen fruit qualifies, but frozen juice concentrate does not because of the added sugar. Cabbage qualifies, but prepared coleslaw does not because of the added mayonnaise. Canned tomatoes qualify but canned pasta sauce does not because of the added salt.

Baked Goods:

- Bread, biscuits and rolls
- Muffins, cakes and cupcakes
- Pies and tarts
- Crackers, chips, pretzels and other snacks do not qualify.
- In general, if you can buy it at a bakery, it counts. For example, sandwich bread qualifies, but a prepared sandwich does not. Muffins, English muffins and bagels qualify but a package of crackers does not.

For whichever category you choose, you will be reimbursed 30% of the price on all items on your receipt in that category, up to a maximum reimbursement of $10.

For example:

- If you choose Fruits and Vegetables and spend $30 on items in that category, you will receive a reimbursement of (30% of $30) = $9.
- If you choose Baked Goods and spend $25 on items in that category, you will receive a reimbursement of (30% of $25) = $7.50.
- If you spend $33.33 or more on items in your chosen category, you will receive the maximum reimbursement of $10.

Your reimbursement will be credited to your Field Agent account within 7 days of this shopping trip. We will use the receipt you submit at the end of this survey to calculate your reimbursement.

Which category would you like to get the 30% reimbursement on? *[Randomized list]*

- Fruits and Vegetables
- Baked Goods

Now, go ahead and complete your shopping trip. The next screen will ask for a picture of your itemized grocery receipt. Take a picture of the receipt before going home form the store (its fine to be in your car or the parking lot)

When you have finished shopping and have paid for your purchases continue to the next screen.

[Continue with Common Receipt Submission]

B.5 Waiting Period

B.5.1 Pre-Shopping Survey

Hi Agents! We have 2 short surveys for you for your next grocery-shopping trip of over $25. Well call these surveys 1) your Planning Survey and 2) your "Shopping Survey."

This is the Planning Survey (1 of 2).

- You MUST complete this survey 4 to 48 hours BEFORE you go grocery shopping.
- If you are already at the store, or will shop sooner than 4 hours from now, just wait to do this survey until the next grocery-shopping trip. We will wait for you.
- If you arent going on a grocery shopping trip in the next 48 hours, close the survey and re-open it 4-48 hours before your next grocery trip.

DO NOT:

- Do not complete this survey in the parking lot of the store. Complete this survey at home. When you are at the store, you will complete the Shopping Survey.

WE WILL:

- Verify that you have completed this survey between 4 and 48-hours before your shopping trip, well compare the time you submit this Planning Survey to the timestamp on your shopping receipt.

NEXT:

- If you will complete a grocery-shopping trip in the next 4-48 hours, please continue with the survey.
- Remember that the store you go to must:
  - Have a produce section with multiple aisles of fruits and vegetables.
  - Provide itemized receipts.

You cannot go backwards in the study, so read and answer all questions carefully.
Have fun!

For your upcoming grocery-shopping trip, we will reimburse you for some of your groceries. Any item in the fruits and vegetables category will qualify for the partial reimbursement. Fruits and vegetables are defined below:

Fruits and Vegetables:

- Fresh fruits and vegetables
- Canned fruits and vegetables
- Frozen fruits and vegetables
- Processed or prepared products that simply contain some fruits and vegetables do not qualify. For example, frozen fruit qualifies, but frozen juice concentrate does not because of the added sugar. Cabbage qualifies, but prepared coleslaw does not because of the added mayonnaise. Canned tomatoes qualify but canned pasta sauce does not because of the added salt.

You will be reimbursed 30% of the price on all items on your receipt in the fruits and vegetables category, up to a maximum reimbursement of $10.

For example:

- If you spend $30 on fruits and vegetables, you will receive a reimbursement of (30% of $30) = $9.
- If you spend $25 on fruits and vegetables, you will receive a reimbursement of (30% of $25) = $7.50.
- If you spend $33.33 or more on fruits and vegetables, you will receive the maximum reimbursement of $10.

Your reimbursement will be credited to your Field Agent account within 7 days of this shopping trip. We will use the receipt you submit at the end of your Shopping Survey to calculate your reimbursement.

On your upcoming Shopping Survey:

You will be required to answer the first survey question BEFORE you finish shopping. Then, you can check out and take a photo of your receipt.

(NOTE: If you miss the 4-48-hour window, your Shopping Survey will direct you to a Make-Up Planning Survey. You will be required to re-take this Planning Survey and wait another 4-48 hours before shopping.)

Select "I understand" to indicate that you have read the instructions and understand the reimbursement offer.

B.5.2 Shopping Survey

Hi Agents! This is survey 2 of 2, your Shopping Survey.

COMPLETE THIS JOB IF:

- If it has been between 4-48 hours since submitting the Planning Survey (1 of 2), and you are at the grocery store. Accept this job and answer the first 2 questions BEFORE you shop for groceries and check out.
- If you missed the 4-48 hour time window Life happens. We get it. Accept this survey and let us know. You will be directed on how to take the Make Up Surveys.
- If you still need to get your groceries now, that if perfectly fine. You will need to wait until your next grocery-shopping trip to complete the Make-Up surveys. We will wait for you.

WE WILL:

- Verify that you have answered the first two questions before purchasing your groceries.

NEXT:

- Remember that the store you go to must:
  - Have a produce section with multiple aisles of fresh fruits and vegetables
  - Provide an itemized receipt

You cannot go backwards in the survey, so read and answer all questions carefully. Have fun!

We asked you in the Planning Survey to wait a minimum of 4 hours and a maximum of 48 hours before shopping and completing this Shopping Survey. Tell us about how long it has been since you completed your Planning Survey for this week. An estimate is OK. Well double check for you.

If it has not been 4 hours, please wait to complete this Shopping Survey until it has been at least 4 hours.

If it has been longer than 48 hours (2 days), we need you to complete a Make Up Planning Survey. Dont worry, well direct you there.

About how long has it been since you completed your Planning Survey for this grocery shopping trip?

- 4-48 hours
- 48 hours or more

Great! Lets get shopping.

[Continue with Common Restricted Subsidy]

[Continue with Common Receipt Submission]

B.6 Waiting Period with Agency

B.6.1 Pre-Shopping Survey

Hi Agents! We have 2 short surveys for you for your next grocery-shopping trip of over $25. Well call these surveys 1) your Planning Survey and 2) your "Shopping Survey." This is survey 1 of 2: your Planning Survey.

Each week you will take this Planning Survey at least 4 hours, and no more than 48 hours BEFORE you head to the grocery store for a grocery shopping trip to purchase food. Remember that the store you go to must:

- Have a produce section with multiple aisles of fruits and vegetables.
- Provide itemized receipts.

You MUST complete this Planning Survey at least 4 hours BEFORE you go to the store for your shopping trip. Do not wait to complete this survey in the parking lot of the store. When you are at the store, you will complete the Shopping Survey. If you are already at the store, or will shop sooner than 4 hours from now, close the survey and re-open it 4-48 hours before your next Stock Up grocery trip. If you arent going

on a Stock Up grocery shopping trip in the next 48 hours, close the survey and re-open it 4-48 hours before your next Stock Up grocery trip.

To verify that you have completed this survey between 4 and 48-hours before your shopping trip, well compare the timestamp on your shopping receipt to the time you submit this Planning Survey. If you will complete a grocery shopping trip in the next 4-48 hours, please continue with the survey.

You cannot go backwards in the study, so read and answer all questions carefully.

Have fun!

For your upcoming grocery-shopping trip, we will reimburse you for some of your groceries. When you begin your Shopping Survey this week, you will need to choose which types of foods to be reimbursed for. Please read the instructions carefully.

There are two food categories for you to choose from:

1. Fruits and vegetables
2. Baked goods

Heres a list of what falls into each category:
Fruits and Vegetables:

- Fresh fruits and vegetables
- Canned fruits and vegetables
- Frozen fruits and vegetables
- Processed or prepared products that simply contain some fruits and vegetables do not qualify. For example, frozen fruit qualifies, but frozen juice concentrate does not because of the added sugar. Cabbage qualifies, but prepared coleslaw does not because of the added mayonnaise. Canned tomatoes qualify but canned pasta sauce does not because of the added salt.

Baked Goods:

- Bread, biscuits and rolls
- Muffins, cakes and cupcakes
- Pies and tarts
- Crackers, chips, pretzels and other snacks do not qualify.
- In general, if you can buy it at a bakery, it counts. For example, sandwich bread qualifies, but a prepared sandwich does not. Muffins, English muffins and bagels qualify but a package of crackers does not.

For whichever category you choose, you will be reimbursed 30% of the price on all items on your receipt in that category, up to a maximum reimbursement of $10.

For example:

- If you choose Fruits and Vegetables and spend $30 on items in that category, you will receive a reimbursement of (30% of $30) = $9.
- If you choose Baked Goods and spend $25 on items in that category, you will receive a reimbursement of (30% of $25) = $7.50.
- If you spend $33.33 or more on items in your chosen category, you will receive the maximum reimbursement of $10.

Your reimbursement will be credited to your Field Agent account within 7 days of this shopping trip. We will use the receipt you submit at the end of your Shopping Survey to calculate your reimbursement.

We will ask you at the store which category you would like to receive the 30% reimbursement for (Baked goods or Fruits & Vegetables).

You will be required to answer that question BEFORE you finish shopping. Then, you can check out and take a photo of your receipt.

(NOTE: If you miss the 4-48-hour window, your Shopping Survey will direct you to a Make-Up Planning Survey. You will be required to re-take this Planning Survey and wait another 4-48 hours before shopping.)

Select "I understand" to indicate that you have read the instructions and understand the reimbursement offer.

B.6.2 Shopping Survey

Hi Agents! This is survey 2 of 2, your Shopping Survey. Take this survey once you are at the grocery store for your grocery shopping trip. You can go to any grocery store you typically shop at as long as it has a produce section and provides itemized receipts.

Remember that the grocery store you go to must:

- Have a produce section with multiple aisles of fruits and vegetables.
- Provide itemized receipts.

WAIT until you arrive at the store to do this survey. If you are not at the grocery store right now, close the survey and re-open it when you arrive. If you are at the grocery store right now, please continue with the survey and answer the questions BEFORE you finish shopping.

If you have missed your 4-48 hour window, you will be directed on how to take a Make Up Planning Survey 4-48 hours before a Make-Up Shopping Survey. If you still need to complete this shopping trip, that if perfectly fine, but wait until your next grocery shopping trip to complete the Make-Up surveys. Accept the Make-Up survey for further directions.

You cannot go backwards in the study, so read and answer all questions carefully.

Have fun!

We asked you in the Planning Survey to wait a minimum of 4 hours and a maximum of 48 hours before shopping and completing this Shopping Survey. Tell us about how

long it has been since you completed your Planning Survey for this week. An estimate is OK. Well double check for you.

If it has not been 4 hours, please wait to complete this Shopping Survey until it has been at least 4 hours.

If it has been longer than 48 hours (2 days), we need you to complete a Make Up Planning Survey. Dont worry, well direct you there.

About how long has it been since you completed your Planning Survey for this grocery shopping trip?

- 4-48 hours
- 48 hours or more

Great! Lets get shopping.

[Continue with Common Choice Procedure]

[Continue with Common Receipt Submission]

B.7 Commitment

B.7.1 Pre-Shopping Survey

Hi Agents! We have 2 short surveys for you for your next grocery-shopping trip of over $25. Well call these surveys 1) your "Planning Survey" and 2) your "Shopping Survey." This is survey 1 of 2: your Planning Survey.

Each week you will take this Planning Survey at least 4 hours, and no more than 48 hours BEFORE you head to the grocery store for a grocery shopping trip to purchase food. Remember that the store you go to must:

- Have a produce section with multiple aisles of fruits and vegetables.
- Provide itemized receipts.

You MUST complete this Planning Survey at least 4 hours BEFORE you go to the store for your shopping trip. Do not wait to complete this survey in the parking lot of the store. When you are at the store, you will complete the Shopping Survey. If you are already at the store, or will shop sooner than 4 hours from now, close the survey and re-open it 4-48 hours before your next grocery trip. If you arent going on a grocery shopping trip in the next 48 hours, close the survey and re-open it 4-48 hours before your next grocery trip.

To verify that you have completed this survey between 4 and 48-hours before your shopping trip, well compare the timestamp on your shopping receipt to the time you submit this Planning Survey. If you will complete a grocery shopping trip in the next 4-48 hours, please continue with the survey.

You cannot go backwards in the study, so read and answer all questions carefully.

Have fun!

For your upcoming grocery-shopping trip, we will reimburse you for some of your groceries. In this survey, you will need to choose which types of foods to be reimbursed for.

[Continue with Common Choice Procedure]

B.7.2 Shopping Survey

We asked you in the Planning Survey to wait a minimum of 4 hours before shopping and completing this Shopping Survey. Tell us about how long it has been since you completed your Planning Survey for this week. An estimate is OK. Well double check for you.

If it has not been 4 hours, please wait to complete this Shopping Survey until it has been at least 4 hours.

If it has been longer than 48 hours (2 days), we need you to complete a Make Up Planning Survey. Don't worry; we'll direct you there.

About how long has it been since you completed your Planning Survey for this grocery shopping trip?

- 4-48 hours
- 48 hours or more

Great! Lets get shopping.

For this grocery-shopping trip, we will reimburse you for some of your groceries. You previously chose which types of foods to be reimbursed for. It is important that you remember this correctly for us to reimburse you.

Which of the two categories did you choose to be reimbursed for?

1. Fruits and vegetables
2. Baked goods

Heres a list reminding you of what falls into each category:
Fruits and Vegetables:

- Fresh fruits and vegetables
- Canned fruits and vegetables
- Frozen fruits and vegetables
- Processed or prepared products that simply contain some fruits and vegetables do not qualify. For example, frozen fruit qualifies, but frozen juice concentrate does not because of the added sugar. Cabbage qualifies, but prepared coleslaw does not because of the added mayonnaise. Canned tomatoes qualify but canned pasta sauce does not because of the added salt.

Baked Goods:

- Bread, biscuits and rolls
- Muffins, cakes and cupcakes
- Pies and tarts
- Crackers, chips, pretzels and other snacks do not qualify.
- In general, if you can buy it at a bakery, it counts. For example, sandwich bread qualifies, but a prepared sandwich does not. Muffins, English muffins and bagels qualify but a package of crackers does not.

For whichever category you chose, you will be reimbursed 30% of the price on all items on your receipt in that category, up to a maximum reimbursement of $10.

Your reimbursement will be credited to your Field Agent account within 7 days of this shopping trip. We will use the receipt you submit at the end of this Shopping Survey to calculate your reimbursement.

[Continue with Common Receipt Submission]