Dynamic Inconsistency in Food Choice: Experimental Evidence from Two Food Deserts

Sally Sadoff*
University of California, San Diego

Anya Samek†
University of Southern California

Charles Sprenger‡
University of California, San Diego

September, 2014
This Version: April 3, 2018

Abstract

We conduct field experiments to investigate dynamic inconsistency and commitment demand in food choice. In two home grocery delivery programs, we document substantial dynamic inconsistency between advance and immediate choices. When given the option to commit to their advance choices, around half of subjects take it up. Commitment demand is *negatively* correlated with dynamic inconsistency, suggesting those with larger self-control problems are less likely to be aware thereof. We evaluate the welfare consequences of dynamic inconsistency and commitment policies with utility measures based on advance, immediate and unambiguous choices. Simply offering commitment has limited welfare (and behavioral) consequences under all measures.

JEL classifications: C91, D12, D81

Keywords: dynamic inconsistency, commitment demand, field experiment, behavioral welfare analysis

*University of California at San Diego, Rady School of Management, 9500 Gilman Drive, La Jolla, CA 92093; sadoff@ucsd.edu.

†University of Southern California, 635 Downey Way, Los Angeles, CA 90035; samek@usc.edu.

‡University of California at San Diego, Rady School of Management and Department of Economics, 9500 Gilman Drive, La Jolla, CA 92093; csprenger@ucsd.edu.

1 Introduction

Models incorporating temptation impulses and self-control are among the most prominent in behavioral economics (Strotz, 1955; Thaler and Shefrin, 1981; Laibson, 1997; O'Donoghue and Rabin, 1999; Gul and Pesendorfer, 2001; Fudenberg and Levine, 2006). The dynamic inconsistencies predicted by these models provide a reason for the observed difficulty of people to save more for the future, exercise more, eat healthier and quit smoking. Based on the insights generated by these models, prescriptions such as offering commitment devices have grown prominent in policy circles.

In this paper, we address two open questions in the literature on self-control. The first is the relationship between self-control problems and awareness thereof. Several experimental studies find weak positive correlations between hallmarks of dynamic inconsistency and take-up of products with commitment features (Ashraf, Karlan and Yin, 2006; Augenblick, Niederle and Sprenger, 2015; Kaur, Kremer and Mullainathan, 2015). This suggests at least a weakly positive correlation between self-control problems and awareness thereof, a finding that is confirmed by recent work eliciting both behavior and beliefs (Augenblick and Rabin, forthcoming).[1] In contrast, outside of controlled experimental settings, there is limited evidence for anything more than tepid demand for commitment devices (Laibson, 2015). This suggests that perhaps those with the worst self-control problems may not be aware of them.[2] Ultimately, relatively little is known about the relationship between behavior and beliefs in non-experimental settings. Given that the impact of commitment policies depends on this real-world relationship, data from field settings has the potential to provide substantial value.

The second open question is the assessment of the welfare consequences of commitment policies. This assessment depends on two critical factors. The first is the aforementioned relationship between self-control problems and awareness thereof, and the second is the chosen welfare criterion. Ambiguity in welfare evaluations may exist in the context of self-control problems since there is inconsistency between 'long-run' preferences measured absent temptation and 'short-run' preferences measured under temptation. A practice has emerged that bases welfare calculations on long-run preferences under the positive justification that short-run preference deviations represent mistakes (Herrnstein, Loewenstein, Prelec and Vaughan, 1993; Gruber and Kőszegi,

[1] In experimental settings, dynamic inconsistency can explain only about 5% of the variation in commitment demand (Augenblick et al., 2015) and individuals seem to understand less than 25% of their self-control problems (Augenblick and Rabin, forthcoming).

[2] Limited commitment demand could have other sources. Laibson (2015) demonstrates that with an uncertain environment and costly commitment, commitment demand may be limited even among agents who are aware of their self-control problems.

2001; O'Donoghue and Rabin, 2006). More recently, Bernheim and Rangel (2007, 2009) develop an alternative approach to behavioral welfare analysis based on unambiguous choice – i.e., using choices that are consistent across the long- and short-run – and provide a theoretical evaluation of dynamically inconsistent preferences. Yet to our knowledge, there exists no empirical evaluation of the welfare consequences of dynamic inconsistency and commitment policies recognizing potential disagreement across welfare criteria.

We combine field evidence on dynamic inconsistency and commitment demand with a behavioral welfare exercise that evaluates commitment policies through the lens of alternative welfare criteria. Our field experiments are conducted in a natural setting, and individuals are not told that they are in an experiment, which mimics naturally occurring markets. Further, our experiments test dynamic inconsistency over consumption using longitudinal decisions with limited scope for arbitrage, which aligns tightly with theoretical models. Finally, we collect within-subject data on dynamic inconsistency and commitment over time, which allows us to investigate stability of these measures.

Our setting is a food delivery service for low-income participants in two cities: Chicago, Illinois and Los Angeles, California. Three-hundred eighty-nine subjects completed a 3-4 week food delivery program. Subjects were given a budget and asked to construct a bundle from a list of 20 foods for home delivery one week later. On the day of delivery, the delivery-person brought the pre-ordered bundle and also surprised subjects with additional foods available for exchange. Subjects were given the opportunity to make up to 4 exchanges. Every bundle that could be constructed with immediate exchanges (on the day of the delivery) is one that was available at the time of advance choice (one week earlier). As such, dynamic inconsistencies are identified as violations of revealed preference between advance and immediate choices.

In the second and third weeks of the study, subjects again made advance choices. However, before the delivery, they were asked if they would like the option to make exchanges at delivery again, or whether they would like to stick to their pre-ordered choices. Commitment demand is identified as choosing to restrict oneself to the advance bundle. The correlation between dynamic inconsistency (in the first week) and subsequent commitment demand provides data on the relationship between self-control problems and awareness thereof that can be used to evaluate commitment policies.

We find that when commitment is not available, 46% of subjects exhibit dynamic inconsistencies, exchanging at least one item from their advance bundle. Regularities exist in the nature of these inconsistencies. Immediate bundles contain significantly

fewer fruits and vegetables and more calories (primarily from fat) than advance bundles.

When commitment is available, 53% of subjects take it up, preferring to restrict themselves to their advance bundle. Importantly, subjects who were previously dynamically inconsistent are *less* likely to demand commitment (44%) than subjects who were previously dynamically consistent (60%). This suggests a negative correlation between self-control problems and awareness thereof.

A structural estimation exercise that formulates utilities in terms of food characteristics indicates the value of fruits and vegetables is significantly lower in immediate versus advance choice. The structural estimates are built using standard random utility methods and allow for tests that inconsistencies would arise by chance under dynamically-consistent preferences. Tests of consistent preferences are rejected for the aggregate data and for inconsistent subjects at all conventional levels. Utility estimates from when commitment is not available show that subjects who ultimately commit have substantially smaller differences between advance and immediate preferences than those who ultimately do not demand commitment.

To understand the welfare consequences of dynamic inconsistencies and commitment policies, we evaluate welfare using three criteria: the advance utility estimated from foods chosen before making exchanges, the immediate utility estimated from foods chosen after making exchanges and the unambiguous utility estimated from foods that were never exchanged. In the spirit of Bernheim and Rangel (2007, 2009), the third welfare criterion allows for welfare evaluation based only on unambiguous choices.

Using the standard 'long-run' welfare criterion of advance utility, we find that aggregate welfare declines by about 2% between advance and immediate choice overall, and by 4-5% for inconsistent subjects. Interestingly, aggregate welfare costs to inconsistency are also found when using the unambiguous and immediate welfare criteria.[3] Despite similarity in the aggregate estimates, there is heterogeneity between and within individual-level welfare measures. The median inconsistent subject exhibits disagreement between their advance and immediate utility measures: advance preferences are more likely to show welfare costs to inconsistency and immediate preferences are more likely to show welfare benefits to flexibility. Where this disagreement exists, the conflict between advance and immediate welfare measures may be helpfully arbitrated by the unambiguous utility measure. Fifty percent of subjects have unambiguous welfare reductions due to inconsistency.

[3]This similarity across criteria may seem surprising. It is driven by a general agreement in both advance and immediate choice that fruits and vegetables are desirable. When inconsistencies occur, they come in the form of deviating from this general agreement by exchanging fruits and vegetables for less desirable items, lowering total utility.

We combine utility estimates and subsequent commitment demand to evaluate the welfare consequences of three potential policies: the standard policy of offering commitment to those who desire it, mandated advance choice and a tailored policy that mandates advance choice only for people who, by our estimates, exhibit unambiguous welfare costs to inconsistency. Given that few dynamically inconsistent subjects ultimately demand commitment, simply offering commitment is predicted to have limited welfare effects. Only 20% of subjects are predicted to be affected, roughly equally split between those made better and those made worse off by their commitment decision. Mandated advance choice would affect about 45% of subjects, again about evenly split between winners and losers. The tailored mandate would affect about 20% of subjects – those with unambiguous costs to inconsistency – with winners outnumbering losers according to the other (advance and immediate) welfare criteria by at least two-to-one.

We also evaluate these policies on the basis of behavior change, specifically on how they impact the nutritional value of foods chosen. Among the three policies, mandated advance choice is predicted to have the greatest effects, increasing the number of fruits and vegetables, and decreasing the number of calories consumed. Simply offering commitment is predicted to have virtually no effect given the observed negative correlation between commitment demand and prior inconsistency. This prediction can be tested in our data in weeks when commitment is available. Indeed, simply offering commitment has virtually no effect on the nutritional value of foods ultimately chosen.

Our two core findings: dynamic inconsistency reflecting changing preferences between advance and immediate choices; and a negative correlation between dynamic inconsistency and demand for commitment are observed at both study sites. The original version of this paper featured only data from Chicago. Los Angeles was added as a full-scale replication and extension of the previously documented findings. Replicating the findings – in particular, the demonstration in field data that those with the most substantial self-control problems may be the least aware thereof – helps to assure the results are not obtained simply by chance.

This paper provides contributions along three principal avenues. First, our data on commitment demand provide evidence on a central assumption around which policy prescriptions for behavioral consumers are built. We show demand for commitment, but find that agents who demand commitment have systematically smaller self-control problems than those who do not. Much of the previous literature on self-control has relied on tests of diminishing patience over monetary rewards rather than consumption, and has used decisions made at a single point in time rather than longitudinally

(Sayman and Onculer, 2009; Halevy, 2015; Sprenger, 2015, provide discussion).[4] With the exception of Read and Van Leeuwen (1998), who studied snack choice among employees but did not study commitment, participants in these studies knew they were part of an experiment, which could affect their decisions. We study subjects in their natural setting, which could explain the difference in our results relative to the weakly positive correlation between self-control and awareness implied by prior research.

Second, our experimental populations sit in the cross-hairs of the food policy debate. Our neighborhoods are considered 'food deserts,' implying a high rate of poverty and limited access to fruits and vegetables.[5] Obesity and related diseases are at an all-time high in the United States, are largely driven by poor food choice, and disproportionately affect low-income communities.[6] Americans consume fewer than the recommended servings of fruits and vegetables, and too many high-calorie, low-nutrient foods. Food assistance programs such as the Supplemental Nutrition Assistance Program (SNAP) are one tool for improving healthfulness of food choice in low-income communities. A policy change is now being piloted that would allow retailers to accept SNAP dollars for pre-ordered food.[7] Our results add to an understanding of the impact of this policy change on behavior and welfare.

Third, our exercise provides a demonstration of the value of combining structural methods and behavioral welfare analysis. Behavioral welfare measures require that researchers do not arbitrarily honor a given preference ranking without a clear reason to do so. In dynamically inconsistent choice, this delivers a natural intuition that virtually nothing concrete can be said with regards to welfare. We demonstrate that this is not necessarily the case. In our structural setting, the body of food choices are informative of how decision-makers value food characteristics. Through the lens of the model, we construct and compare welfare measures that deliver clear welfare implications. And we join a small list of empirical studies that investigate the welfare consequences of behavioral phenomena (Chetty, Looney and Kroft, 2009; Allcott, Mullainathan and Taubinsky, 2014; Allcott and Taubinsky, 2015; Rees-Jones and Taubinsky, 2016; Taubinsky and Rees-Jones, forthcoming). We join an even smaller list that recognizes the corresponding ambiguity in welfare estimates that may arise in the Bern-

[4]Related studies include Duflo, Kremer and Robinson (2011) for farmer fertilizer purchase; Augenblick et al. (2015) for effort choices in a laboratory experiment; and subsequent to our study, Augenblick and Rabin (forthcoming) also for effort choices in the laboratory.

[5]A food desert is defined as having a poverty rate of 20% or greater and at least 33% of the census tract lives more than one mile from a supermarket or large grocery store (http://apps.ams.usda.gov/fooddeserts/fooddeserts.aspx).

[6]See https://www.cdc.gov/obesity/data/adult.html.

[7]See https://www.fns.usda.gov/snap/online-purchasing-pilot.

heim and Rangel (2007, 2009) welfare framework (see Bernheim, Fradkin and Popov, 2015).

In what follows, Section 2 provides an overview of the experimental design and describes the structural analysis, Section 3 describes our results and Section 4 concludes.

2 Empirical Design

2.1 Experimental Setup

We conducted two field experiments with a total of 389 subjects at grocery stores in Chicago, Illinois and Los Angeles, California.[8] The first experiment was implemented with 218 subjects in 2014 at Louis' Groceries, a small-format neighborhood grocery store in the low-income community of Greater Grand Crossing in Chicago. The second experiment was implemented with 171 subjects in 2016-17 at Northgate Gonzalez Market, a large supermarket in low-income South-Central Los Angeles.[9]

The grocery stores carried out a promotion inviting customers to sign up for a free home food delivery program. Recruitment for both experiments was conducted on a rolling basis. Two research assistants worked at each grocery store to conduct the experiment and deliver the foods. Subjects for the study were recruited at a table set up at the store. We assured that foods were fresh and produce was not bruised at the time of delivery by working with the grocery stores and preparing deliveries as close to the delivery time as possible. In keeping with the natural field experiment methodology, subjects were not told that they were in an experiment.[10] In the Los Angeles study, to increase naturalism, research assistants partnered with a store associate to deliver items in the Northgate store delivery van. Thus, we were able to observe subjects in their natural environment as they made a series of food allocation decisions.

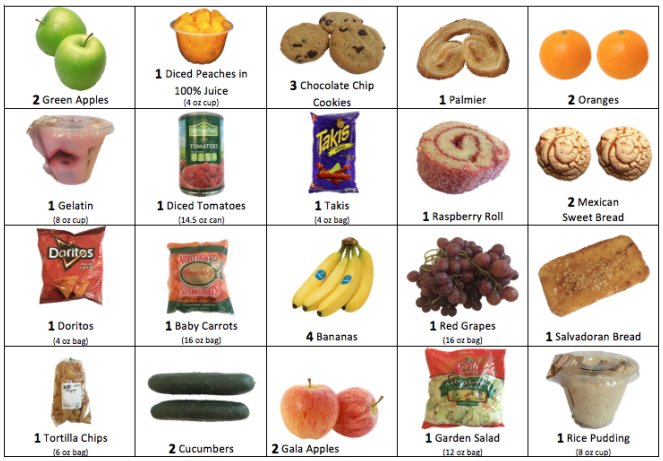
A total of 20 different foods were used in each experiment. Figure 1 displays the

[8]Four hundred and ten subjects were initially recruited into the study. Of these 410, 21 (5.12%) are considered attrited from the study due to not completing the full set of deliveries (17), never being offered a commitment decision due to experimenter error (3) or opting out after the study ended (1).

[9]According to the 2010 U.S. Census, Greater Grand Crossing has a population of 35,217, the majority of whom are African Americans (97.8%). South-Central Los Angeles has a population of 169,453. The majority of residents are Hispanic (74%) and African-American (24%). A larger share of our LA study participants were Hispanic (98%), since the store caters to Hispanic customers. Both neighborhoods have high rates of poverty (28.5%-33.6%).

[10]In the Chicago experiment, The University of Wisconsin-Madison Institutional Review Board (IRB) required us to notify subjects after the study was complete that they had participated and give them the option to withdraw their data. One subject chose to withdraw, and this subject's data is not in the dataset. The Los Angeles experiment was approved by the University of Southern California's IRB, which did not have this requirement.

(a) Chicago



(b) Los Angeles

Figure 1: Study Foods

promotion sheet of foods used. Foods were selected in consultation with store managers to determine which foods would be appealing to customers at each site. In each study, 10 of the foods were fruits or vegetables while the other 10 were sweets or salty snacks. Foods varied substantially in their caloric and nutritional content. Appendix Table A1 provides nutritional information for the foods included in each study.

Upon signing up for the program, subjects were asked whether they had eaten each of the 20 foods before and then rated those they had eaten on a Likert scale from 1 (least preferred) to 7 (most preferred). The use of Likert scales to rate foods has been promoted in the nutrition literature as a means of assessing dietary preferences (Geiselman, Anderson, Dowdy, West, Redmann and Smith, 1998).[11] Subjects were

[11]In Chicago, the question was worded as, *Please tell us how much you like the following foods, where 1 is DO NOT LIKE AT ALL and 7 is LIKE VERY MUCH.* The question was worded slightly

generally aware of and had eaten all 20 of the foods. On average, subjects rated 18.6 of 20 foods and the average food rating was 5.58 out of 7.[12]

In return for participating in the program – including selecting foods, receiving the weekly deliveries and completing surveys – subjects received a participation payment. This payment was a $20 cash voucher in the Chicago study and a $25 Northgate store gift card in the Los Angeles study.

2.2 Experimental Timeline

The experimental timeline is presented in Table 1. The Chicago study offered a 2-week food delivery program while the Los Angeles study offered a 3-week food delivery program. In Week 1, each subject decided on foods for delivery in Week 2. Upon receiving the delivery in Week 2, each subject was surprised with the option to make immediate exchanges. In Week 2, each subject also decided on foods for the second delivery in Week 3. All Chicago subjects subsequently made a commitment choice, deciding whether to have the option to make exchanges (i.e., not commit) or to stick to their pre-ordered choices (i.e., commit) for the second delivery. To investigate the stability of inconsistency and commitment demand, we randomly assigned half of the subjects in Los Angeles to receive commitment offers for both the second and third delivery. We assigned the other half to make a second surprise exchange and offered this group commitment only for the third delivery.

Week 1, Advance Choice: In Week 1, subjects received an order sheet and brochure listing available foods and decided on foods for their first delivery. All foods were also available at the store, and the fresh foods were visible to the subjects as they made their decisions. To simplify the selection process, each food was valued at $1, with cheaper foods bundled into several for $1 (e.g., 2 green apples together cost $1). All foods were priced as closely as possible to their respective market price. Subjects were asked to create a 'basket' of foods valued at $10 in total, by choosing from any of the 20 foods, including selecting the same food more than once. Subjects also selected

differently in Los Angeles. It was, *For foods that you have eaten, I'd like to know how much you like eating the food. When you answer how much you like eating the food, please think carefully about how much you enjoy the food, including aspects such as how the food tastes to you. [point to food] How much do you like eating the food? Do you not like it at all, do not like it, do not like it a little, have no preference, like it a little, like it or like it very much?*

[12]Completing a rating for all foods was voluntary; nevertheless, most subjects rated a large number of foods, with 357 of 389 (92%) rating 15 or more foods. In Chicago 191 of 218, or 88% rated at least 15 foods. In Los Angeles 166 of 171, or 97% rated at least 15 foods. This difference could be because in Chicago, subjects wrote down their responses while in Los Angeles, subjects responded verbally.

Table 1: Summary of Experiment

Week 1	Week 2	Week 3	Week 4 (L.A. only)
Pick Delivery 1 items	Get Delivery 1	Get Delivery 2	Get Delivery 3
	Decide about changes to Delivery 1	If no commitment: decide about changes to Delivery 2	If no commitment: decide about changes to Delivery 3
	Pick Delivery 2 items	Pick Delivery 3 items (L.A. only)	
Pre-Survey Food Ratings	Commitment choice for Delivery 2 (Chicago & half of L.A. subjects)	Commitment choice for Delivery 3 (L.A. only)	Post-Survey (Week 3 in Chicago)

their preferred dates and times of delivery.

Subjects were informed that they would need to be home during their delivery, and would need to show a picture ID to receive their basket. Delivery was scheduled as close to 7 days after sign-up as possible, taking into account the constraints faced by the research assistants (i.e., a maximum number of deliveries can be made in any day) and the availability of the subject. Subjects were required to give a current phone number and address to facilitate delivery. All subjects received a phone call to confirm enrollment upon sign-up, which also allowed us to validate their phone number.

Week 2, Immediate Choice: A few days before scheduled delivery in Week 2, we initiated a reminder call to ensure that subjects would be home at the pre-arranged time and then proceeded with delivery. Upon delivery, subjects were surprised with the opportunity to make up to 4 exchanges. In Chicago, we brought a customized box of 4 foods selected from the 20 that were available previously, whereby we tried to select foods that the subject liked. This box contained their highest rated fruit or vegetable, their highest rated fruit or vegetable not included in their original bundle, their highest rated sweet or salty snack and their highest rated sweet or salty snack not included in their original bundle. In Los Angeles, we brought a box with one of each of the 20 foods that were available previously, and subjects could make exchanges with any of these foods. As before, cheaper foods were bundled into several for $1. Subjects were not told in advance that they would have this opportunity to exchange. The opportunity to exchange was described by a research assistant serving as a delivery-person and was fully scripted as:

Hello, I am here with your basket. Please take a look [Bring open basket,

allow person to look through]. We also have some extra items available. If you like, you can exchange any one item in your basket for one of these items [show extra items on tray]. I brought 4 additional items, so you can make up to 4 exchanges. Do you want to make any exchange? [Great thanks, let me note that on your order sheet.][13]

After making any exchanges, subjects used a new order sheet to make a decision about the contents of their second delivery, scheduled for Week 3.

Weeks 2-3, Commitment Choice: We elicited demand for commitment by asking subjects whether they would like to have the option to make exchanges during the Week 3 delivery, or whether they would like to stick to their pre-ordered choices. We asked this of all subjects in Chicago and half of subjects in Los Angeles. The question was again fully scripted in both study locations. In Chicago, the script was:

Last time, we brought some extra items for you so you could exchange if you changed your mind from your previous choices. This time, we can also bring extra items, but I wanted to check if you'd like that or not. It is up to you: would you like me to bring extra items this time, or not?

In Los Angeles, the script was:

For this week's delivery, you had the option to change your mind by exchanging items in your basket. This time, you can choose whether you want the option to make exchanges, or whether you want to stick to your pre-ordered choices. It is no trouble for us either way, it is entirely up to you. Do you want to have the option to make exchanges, or do you want to stick to your pre-ordered choices?

In Chicago, the commitment question was asked via phone during the reminder call before the next delivery. In Los Angeles, the commitment question was asked in person immediately after the order for the next delivery was placed. If a subject answered that they wanted to have the option to make exchanges, additional items were presented at the next delivery as before. If a subject answered that they would like to stick to their

[13]In Los Angeles, the message was slightly different, *Here is your food delivery [show box]. Please take a look [bring open basket, allow person to look through]. We also have some extra items available. If you like, you can exchange any one item in your basket for one of these items [show extra items in tray]. I brought all the menu items, and you can make up to 4 exchanges. Do you want to make any exchange? [Great thanks, let me note that on your order sheet].*

pre-ordered choices, the box of additional items was not brought along with the delivery.

Weeks 3-4, Final Delivery and Commitment Choice: The subjects in Los Angeles not assigned to the commitment treatment were offered the opportunity to make exchanges in Week 3. The subjects in Los Angeles assigned to the commitment treatment only had the option to make exchanges if they previously chose not to commit. After delivery in Week 3, all Los Angeles subjects used a new order sheet to make a decision about the contents of their third delivery, scheduled for Week 4. After completing this order sheet, all subjects were asked the commitment question applied to their Week 4 delivery. At the final delivery (Week 3 for Chicago and Week 4 for Los Angeles), subjects completed a survey and received compensation for participating.

2.3 Design Considerations

Our Chicago and Los Angeles studies follow similar procedures. The Los Angeles study was constructed as a replication and extension and so allowed us to address potential concerns with respect to identifying dynamically inconsistent preferences and commitment demand. We are indebted to thoughtful comments from colleagues that helped guide these design alterations.

First, dynamic inconsistencies are identified from exchanges between advance and immediate food choice. An intuitive direction of inconsistency is exchanging objects such as fruits and vegetables for sweets and salty snacks. An interpretation that attributed such inconsistencies to changing preferences could be challenged by several concerns in the Chicago design. First, in the Chicago study, all fruit and vegetable items were perishable while no sweets and salty snacks were perishable. If perishable items wound up being damaged, spoiled or less attractive than expected upon delivery, exchange could be driven by such negative surprises rather than by inconsistent preferences. Recognizing this critique, the Los Angeles study was designed with primarily perishable items, only two non-perishable fruit and vegetable items (diced peach cup and canned diced tomatoes) and 2 non-perishable snack items (Doritos and Takis Chips). Additionally, 2 fruits and vegetables came in factory packaging (baby carrots and salad) while most snack items came from the bakery department without factory packaging (e.g., Salvadoran bread).

Second, in our Chicago study, we brought only 4 additional items selected based on subjects' rating data. Any lack of dynamic inconsistency could be driven by our inability to match subjects with tempting items for exchange. Though this suggests any exchanges would speak to a lower bound on inconsistent preferences, in the Los

Angeles study we improved on this design by making all 20 items available for exchange. To keep the designs as similar as possible, however, we retained the design element of allowing only up to 4 exchanges. In practice, this restriction rarely binds, with only 1 of 389 subjects making 4 exchanges at their first delivery.

Third, our Chicago subjects only made one exchange decision prior to being offered commitment. It may be that any observed dynamic inconsistency is ephemeral, a product of shocks or changing circumstances. These random shocks should not deliver a systematic direction for inconsistency. Nevertheless, having more data at the subject level as we do in the Los Angeles study allows us to further rule out that the inconsistencies are due to random shocks.

Fourth, the phrasing of our commitment offer in Chicago may have had the unintended effects of making commitment appear socially desirable and/or may have failed to emphasize that commitment induces a restriction to advance choice. Subjects who did not want to trouble the delivery person may have opted to commit to save him or her work. Subjects opting out of the exchange opportunity may not have realized that this was equivalent to a choice to commit to the advance bundle. For these reasons, the Los Angeles study script highlights that neither choice is more costly for the delivery person, and that the decision to commit is equivalent to sticking with advance choice.

In both of our studies, we observe choices but not consumption of food items. One may worry that subjects' choices do not represent their true preferences, but rather reflect their external opportunities to trade food items. For example, a subject who can trade tomatoes for chips more advantageously outside of the experiment may choose a bundle consisting only of tomatoes, conduct appropriate trades and generate for herself an opportunity set which dominates that provided by the researchers. Such arbitrage would imply that subject choices are not informative of preferences at all, but rather only of external constraints and the researchers' mis-pricing of items.[14] Several aspects of the experimental environment minimize the possibility of arbitrage. The prices in the stores are similar to those faced in the experiments. Hence, external exchanges are unlikely to be advantageous. Additionally, our stores are in 'food deserts,' and many study foods - e.g., fresh fruits and vegetables and bakery goods - are difficult to obtain elsewhere. Conducting exchanges with others in the neighborhood is also practically difficult given the cost of identifying interested parties and the perishability of some foods. Importantly, even if arbitrage opportunities exist, one would not expect them to change dramatically over a single week in our studies. Hence, if choice is driven

[14] A similar arbitrage argument is used to question the use of monetary payments in studies of intertemporal choice (Cubitt and Read, 2007; Chabris, Laibson and Schuldt, 2008; Andreoni and Sprenger, 2012; Augenblick et al., 2015).

by arbitrage strategies, dynamic inconsistencies should be rare. The data themselves can provide some indication of arbitrage strategies by examining the prevalence of completely concentrated bundles, consisting of only a single food. Such bundle concentration is never observed, with the average advance first week bundle having 9.3 unique items. Further, we rarely see a more limited version of concentration: subjects choosing exclusively fruits and vegetables or exclusively sweets and salty snacks. Only 14 of 389 advance bundles in the first week are concentrated this way.

An additional concern posed by not observing food consumption is that if foods are not consumed immediately, temptation may be limited. In our Los Angeles study, we measure the speed with which foods are consumed by including questions about consumption in our post-experiment survey. Subjects were asked, for the foods they ordered in their Week 3 delivery, how quickly they ate the foods - within 1-3 days, 4-7 days or in more than 7 days. Most foods were consumed within 1-3 days, ranging from 79% (for canned tomatoes) to 87% (for Palmiers). Importantly, the non-perishable foods are eaten within 1-3 days as frequently as the perishable foods. This suggests that most foods are indeed being consumed rapidly, within the time frames thought to be relevant for temptation. That subjects do not apparently store more long-lasting foods helps to alleviate the perishability issue discussed previously.

Finally, commitment demand may be an imperfect proxy for awareness about self-control problems. An alternative approach is to elicit beliefs about future behavior, as in Augenblick and Rabin (forthcoming). We did not elicit beliefs for two reasons. First, we wanted to maintain the naturalism of the study. Second, using incentives to elicit beliefs (to make the beliefs incentive compatible) is also a form of providing a commitment device because deviating from predicted behavior in immediate choice is costly (see Augenblick and Rabin, forthcoming, for discussion). Further, Augenblick and Rabin (forthcoming) find that participants may seek to match their behavior to earlier predictions, suggesting that predictions may affect future behavior rather than serving purely as an exogenous measure of self-awareness.[15]

2.4 Structural Analysis, Dynamic Inconsistency and Welfare

Subjects in our experiments choose a bundle of 10 foods from a set of 20 potential options. From such data, reduced form and structural analysis of dynamic inconsistency in food choice can be conducted. The structural method we propose follows

[15]To address these concerns, Toussaert (2015) elicits beliefs about the behavior of similar others rather than oneself. However, de Oliveira and Jacobson (2017) demonstrate that people may have systematically different beliefs about their own time preferences versus those of others.

standard random utility techniques, establishing the value of a given item as being derived from a set of characteristics. This allows for simple tests of dynamically inconsistent preference, recognizing the existence of random shocks. The estimated utilities lend themselves naturally to evaluation of commitment policies under different welfare criteria.

Following methodology from Beggs, Cardell and Hausman (1981), we define each food as a bundle of underlying attributes and analyze subject choices using rank order discrete choice methods.[16] Let the utility of each food, $j \in \{1, ..., J\}$, be written as a linear combination of attributes,

$$V_j = \mathbf{x}_j\beta + \epsilon_j \quad j = 1, ..., J,$$

where $\mathbf{x}_j$ represents a vector of food characteristics and ϵ_j represents a random utility shock drawn iid from a Type-1 extreme value distribution. The probability that a given food, j is preferred to alternatives $1, ..., J-K-1$ is

$$F_j[x_1, ..., x_{J-K-1}, x_j; \beta] = \frac{exp(\mathbf{x}_j\beta)}{exp(\mathbf{x}_j\beta) + \sum_{i=1}^{J-K-1} exp(\mathbf{x}_i\beta)}.$$

Consider a subject who chooses K unique food items. Order the foods as $r \equiv \{1, ..., J-K-1, J-K, J-K+1, ...J\}$, with the final K foods being the chosen items. The probability of observing such an ordering is thus

$$Prob(r, \mathbf{x}; \beta) = \prod_{j=J-K}^{J} F_j[x_1, ..., x_{J-K-1}, x_j; \beta],$$

where $\mathbf{x} \equiv \{\mathbf{x}_1, ..., \mathbf{x}_J\}$ is the matrix of attributes corresponding to the provided order. Indexing individuals by $i = 1, ..., N$, one constructs the log-likelihood of seeing a given N rankings as

$$L(\beta) = \sum_{i=1}^{N} \log(Prob(r_i, \mathbf{x}_i; \beta)). \tag{1}$$

This structure assumes that any chosen item is preferred to *all* unchosen items.

[16] An alternative structural methodology is to consider each bundle of 10 items as a potential option and consider the discrete choice problem of picking the best bundle. With 20 foods, there are $\binom{20}{10}$= 184,756 possible bundles of 10 unique items, and $\binom{20+10-1}{10}$ = 20,030,010 possible bundles of 10 items with repetitions. For both tractability and interpretability, we opt to formulate food and bundle utilities as being derived from a set of characteristics. Note, however, that our construction is not able to capture, for example, a preference for diversity in the bundle or complementarities between particular items.

Within the sets of chosen and unchosen items, no explicit ranking exists. In the language of rank order logit models, the ranks within these sets are 'tied' as all permutations of rankings within these sets would be consistent with observed behavior. Standard methodology exists for incorporating the probability of these ties into maximum likelihood estimates of the parameters of interest, β. We augment the probability of equation (1) with Efron's (1977) method for handling ties in rank order data, implemented in *Stata*.

2.4.1 Tests of Dynamic Inconsistency

Consider two rankings of foods: one from advance decisions and one from immediate decisions. Let r_A and r_I represent the advance and immediate rankings, respectively. Maximum likelihood estimation of attribute weights, β_A and β_I, based upon these rankings provide a means of comparing preferences across choice environments. Further, β_A and β_I can be estimated simultaneously and one can test the null hypothesis of dynamically *consistent* preferences, $\beta_A = \beta_I$, using standard χ^2 tests. Such tests establish the probability that observed exchanges would occur by chance under the extreme value error structure without dynamically inconsistent preferences.

Two points related to our structural tests of dynamic consistency are worth noting. First, in both of our studies, subjects were only allowed to make up to 4 exchanges. This restriction limits the inconsistencies that can be observed between r_A and r_I. Though in practice, only 1 of 389 subjects made all 4 exchanges at their first delivery, this design feature could in principle, work against finding differences between β_A and β_I. Second, in our Chicago study, our design called for bringing only 4 additional items when making food deliveries. As such, r_I may be additionally restricted to be similar to r_A by our inability to provide subjects with sufficiently tempting alternatives, again working against finding differences between β_A and β_I. Our Los Angeles design does not suffer from this potential issue, as all foods were available for exchange when subjects made immediate choices. These points suggest that findings of dynamic inconsistency and the corresponding changes in preferences estimated in our study may be lower bounds.

2.4.2 Welfare Evaluation

Estimated utility weights, β_A and β_I, speak to two different potential welfare criteria based on advance and immediate preferences, respectively. One can construct the

deterministic utility portion of any proposed bundle under advance preferences as

$$V_A(\mathbf{q}) = \sum_{j=1}^{J} q_j \mathbf{x}_j \beta_A,$$

where $\mathbf{q} = \{q_1, ..., q_j, ...q_J\}$ is the proposed bundle with quantity q_j of food j.[17] Similarly, one can construct the immediate utility,

$$V_I(\mathbf{q}) = \sum_{j=1}^{J} q_j \mathbf{x}_j \beta_I.$$

These two measures can be used to evaluate the welfare consequences of dynamic inconsistency and commitment policies. If disagreement in choice, and hence potential differences between β_A and β_I exist, welfare statements may be ambiguous. $V_A(\cdot)$ and $V_I(\cdot)$ may disagree on the value of policies.

Where disagreement in choice exists across welfare relevant choice conditions, Bernheim and Rangel (2007, 2009) advocate for formulating welfare statements around an unambiguous choice relation that never contradicts choice. By examining only foods that were never exchanged, we can construct this unambiguous relation. Consider the ordering $r_U \equiv \{1, ..., J-E-K-1, J-E-K, J-E-K+1, ...J-E\}$ with the final K foods being the chosen items and E being the number of items that were ever exchanged from advance to immediate choice conditions. The likelihood

$$Prob(r_U, \mathbf{x}; \beta_U) = \prod_{j=J-K-E}^{J-E} F_j[x_1, ..., x_{J-K-1}, x_j; \beta_U]$$

can be used to estimate unambiguous utility values β_U, ignoring any exchanged items. If no items are ever exchanged, the rankings are identical and $\beta_U = \beta_A = \beta_I$. If exchanges are made, β_U can differ from both β_A and β_I. One can then construct the unambiguous utility of a proposed bundle $\mathbf{q}$,

$$V_U(\mathbf{q}) = \sum_{j=1}^{J} q_j \mathbf{x}_j \beta_U.$$

It is important to note that though β_U is estimated without foods that were ever

[17] Note that the intensive margin of choice represented by the quantities $\mathbf{q}$ is not a feature of the estimated likelihood, but is present in the determination of utility values. Given that most chosen bundles consist of only unique food items, the distinction between the extensive and intensive margin is rarely of importance in our setting.

exchanged, an unambiguous utility value is generated for exchanged foods. This means that though r_U does not contradict choice, β_U will potentially assign different utility values to two items that were exchanged for each other.[18] As such, β_U, informed by subjects' other decisions, may arbitrate between these two foods. If a subject unambiguously chooses fruits and vegetables over sweets and salty snacks, β_U will reflect this in utility weights that are positive to fruit and vegetable characteristics. Exchanging a bag of chips for a piece of fruit would be viewed as an improvement under β_U, while the opposite would be viewed as deleterious. We view the arbitration between conflicting advance and immediate welfare criteria as a valuable feature of our structural exercise and evaluate the consequences of commitment policies through the lens of all three measures, $V_A(\cdot)$, $V_I(\cdot)$ and $V_U(\cdot)$.

3 Results

We present the results in three sub-sections. Sub-section 3.1 discusses reduced form evidence on dynamic inconsistency and assesses the relationship between dynamic inconsistency and commitment. Sub-section 3.2 evaluates the welfare consequences of dynamic inconsistency and commitment policies. Sub-section 3.3 is dedicated to robustness tests and evaluation of additional data.

3.1 Reduced Form Evidence: Dynamic Inconsistency and Commitment Demand

3.1.1 Dynamic Inconsistency

Our analysis of dynamic inconsistency contrasts advance and immediate decisions when commitment is not available. In Chicago, 82 of 218 subjects (37.6%) exhibit dynamic inconsistency in the first week by making at least one exchange between advance and immediate choice. Similarly, in Los Angeles, 66 of 171 subjects (38.6%) exhibit dynamic inconsistency in the first week. Of the 256 allocations in Los Angeles where commitment is not offered, 121 (47.3%) exhibit inconsistencies. Pooling our study sites, 203 of 474 (43%) allocations made without commitment offered exhibit dynamic inconsistency. Of 389 total subjects, 177 (46%) ever exhibit such an inconsistent allocation.

[18] This is the sense in which our analysis is in the spirit of Bernheim and Rangel (2007, 2009). Whereas welfare statements constructed from an unambiguous choice relation will never contradict choice, welfare statements constructed from utility estimates based upon unambiguous choices may do so.

Figures 2 and 3 explore the nature of these inconsistencies at the aggregate and individual level. Though there are many ways in which the data can be examined, we begin by evaluating a simple observable characteristic: whether the chosen food is a fruit or vegetable, or a sweet or salty snack. Figure 2 graphs the frequency with which each food appears in immediate and advance bundles across study sites, where each point represents the raw frequency with which each food is chosen over all subjects in a location-week. Given one week of data prior to being offered commitment in Chicago and two weeks of data prior to commitment being offered to all subjects in Los Angeles, there are 60 total foods represented. Of the 30 fruits and vegetables, 22 are chosen less frequently in immediate choice. Of the 30 sweets and salty snacks, 23 are chosen more frequently in immediate choice. Figure 3 graphs the proportion of fruits and vegetables contained in chosen bundles, where each point now represents a subject-week prior to commitment being offered.[19] Among observations that change the proportion of fruits and vegetables between advance and immediate choice, 79%-96% show reductions in fruits and vegetables in immediate choice. A clear pattern emerges – fruits and vegetables are chosen more often in advance choice, while sweets and salty snacks are chosen more often in immediate choice.

The systematic patterns of inconsistencies discussed above are supported by the statistics in Table 2, which also includes analysis along additional nutritional dimensions. For each subject at each point in time, we aggregate bundle characteristics by summing over the chosen foods along observable and nutritional characteristics. We estimate differences between advance and immediate choice using Ordinary Least Squares (OLS) estimation with standard errors clustered at the individual level. We observe significant differences between advance and immediate bundles in almost every nutritional category at both study sites. Inconsistent subjects substitute lower calorie, lower fat and lower carbohydrate foods with higher calorie, higher fat and higher carbohydrate foods. These patterns largely come from exchanging fruits and vegetables for sweets and salty snacks.

3.1.2 Commitment Demand

Our design elicits commitment demand in the form of giving up the option to exchange foods for the next delivery date. Of 218 subjects in Chicago, 73 (33.5%) demand commitment for their second delivery. In Los Angeles, commitment demand is more frequent than in Chicago. Of 171 subjects in Los Angeles, 134 (78.4%) ever demand

[19] Appendix Figure A1 shows similar information for calories, fat grams, carbohydrate grams and protein grams.

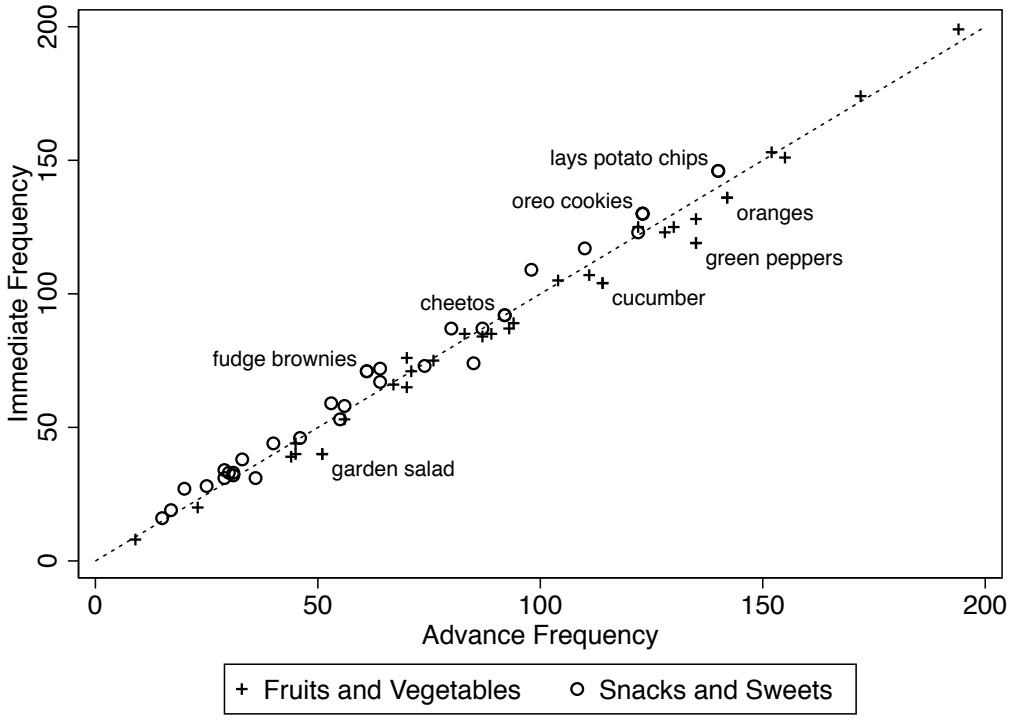


Figure 2: Frequency of Foods in Advance and Immediate Choice

Notes: Each point represents the frequency with which each food is chosen over all subjects in a location-week. This makes 60 points in total - 30 fruits and vegetables and 30 sweets and salty snacks. Foods appearing more frequently in advance versus immediate bundles lie below the 45° line. Of the 30 fruits and vegetables, 22 are chosen less frequently in immediate choice. Of the 30 sweets and salty snacks, 23 are chosen more frequently in immediate choice. While some foods are more popular than others, all foods are chosen with some frequency.

commitment, with 69 of 86 (80.2%) doing so in Week 2 and 127 of 171 (74.3%) doing so in Week 3. A potential reason for the difference across study sites is that we offered commitment to Chicago subjects a few days prior to the next delivery, while we offered commitment to Los Angeles subjects immediately after they made their advance choices for the next delivery. However, differences in the sample population and study design across sites make it difficult to identify the underlying reason for this difference.

Figure 4 displays the association between dynamic inconsistency and subsequent commitment demand. In Chicago, 55 of 136 (40.4%) dynamically consistent subjects demand commitment, while only 18 of 82 (22.0%) dynamically inconsistent subjects do so. In Los Angeles, 95 of 105 (90.5%) of subjects who are dynamically consistent in their first delivery ever demand commitment, while only 39 of 66 (59.1%) dynamically inconsistent subjects do so. Of 256 total allocations made in Los Angeles prior to being offered commitment, 123 of 135 (91.1%) dynamically consistent obser-

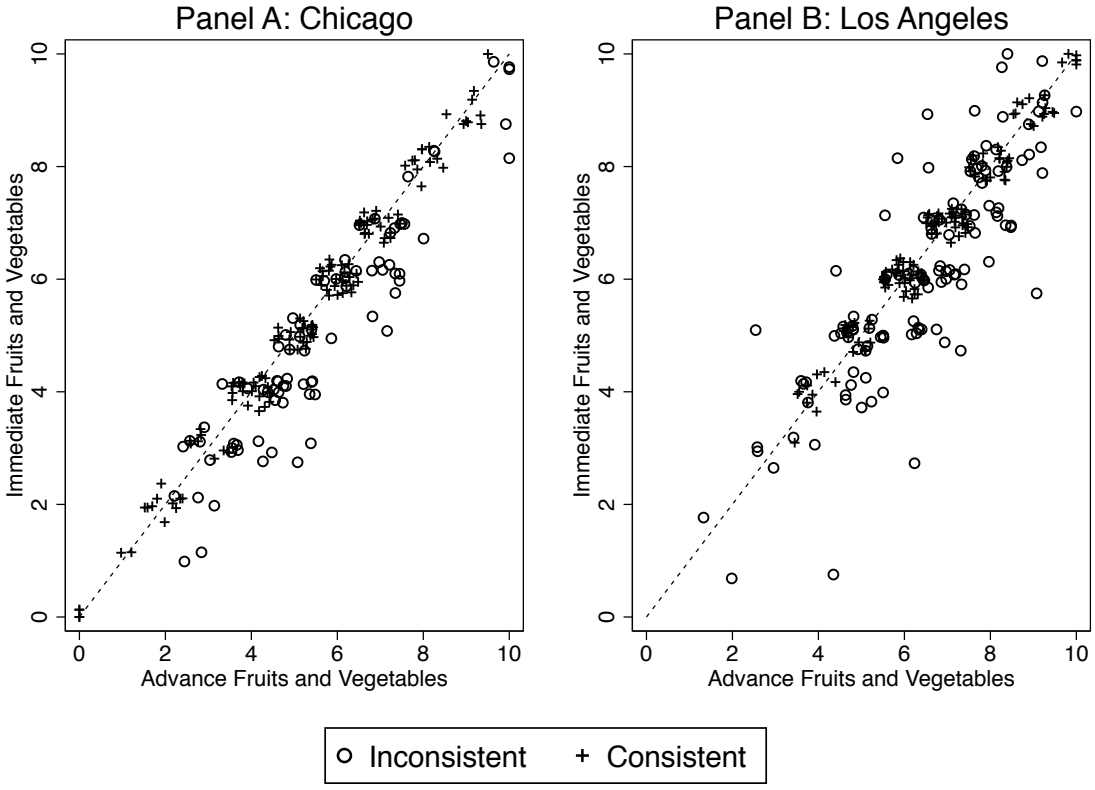


Figure 3: Bundle Composition in Advance and Immediate Choice

Notes: Each point represents a subject-week. Bundles with more fruits and vegetables in advance versus immediate choice lie below the 45° line. This graph includes a 5% jitter, which is the reason why all consistent observations do not appear on the 45° line. In Chicago (Panel A), 82 observations are inconsistent, with 46 of 82 (56%) changing the number of fruits and vegetables from advance to immediate choice. Among the 46 who change the number of fruits and vegetables, 44 (96%) reduce the number of fruits and vegetables in immediate choice. In Los Angeles (Panel B), 121 observations are inconsistent, with 66 of 121 (55%) changing the number of fruits and vegetables from advance to immediate choice. Among the 66 who change the number of fruits and vegetables, 52 (79%) reduce the number of fruits and vegetables in immediate choice.

vations and only 69 of 121 (57.0%) dynamically inconsistent observations are linked to subsequent commitment demand. The correlation between commitment demand and dynamic inconsistency at both study sites is negative and statistically significant at conventional levels - $\rho = -0.19$ ($p < 0.01$) in Chicago, and $\rho = -0.37$ ($p < 0.01$) and $\rho = -0.39$ ($p < 0.01$) in Los Angeles. Hence, though levels of commitment differ across study sites, the negative relationship between commitment demand and prior inconsistency is reproduced at both locations.

Table 3 provides OLS regressions on bundle characteristics for committing and non-committing subjects in advance and immediate choice for all allocations made prior to commitment being offered. At both study sites, committing subjects exhibit

Table 2: Bundle Characteristics

	(1) Fruits/Veg	(2) Sweets	(3) Salty Snacks	(4) Calories	(5) Fat (g)	(6) Carb (g)	(7) Protein (g)
Panel A: Chicago Study							
Immediate Choice	-0.220***	0.161***	0.060**	61.573***	4.051***	5.661***	0.338**
	(0.034)	(0.029)	(0.024)	(12.429)	(0.716)	(1.856)	(0.148)
Constant	5.390***	2.628***	1.968***	2723.890***	89.658***	462.236***	39.414***
	(0.140)	(0.103)	(0.078)	(40.233)	(2.783)	(5.129)	(0.444)
# Observations	436	436	436	436	436	436	436
# Subjects	218	218	218	218	218	218	218
Panel B: Los Angeles Study							
Immediate Choice	-0.168***	0.141***	0.027	57.686**	3.263**	6.254	1.092**
	(0.042)	(0.039)	(0.031)	(25.598)	(1.359)	(3.825)	(0.473)
Constant	6.745***	2.263***	0.986***	3354.537***	67.616***	665.328***	55.596***
	(0.116)	(0.099)	(0.060)	(60.199)	(3.155)	(8.921)	(1.071)
# Observations	512	512	512	512	512	512	512
# Subjects	171	171	171	171	171	171	171
Week Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel C: Pooled Data							
Immediate Choice	-0.192***	0.150***	0.042**	59.474***	3.626***	5.981***	0.745***
	(0.028)	(0.025)	(0.020)	(14.932)	(0.803)	(2.231)	(0.265)
Constant	6.757***	2.258***	0.979***	3353.643***	67.435***	665.464***	55.769***
	(0.116)	(0.098)	(0.060)	(59.508)	(3.119)	(8.803)	(1.064)
# Observations	948	948	948	948	948	948	948
# Subjects	389	389	389	389	389	389	389
Week Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Ordinary least squares regression. Dependent variable reported for each column. Standard errors clustered on individual level in parentheses. Levels of significance: * 0.10, ** 0.05, *** 0.01.

different behavior in both advance and immediate choice. Though more pronounced in Los Angeles, committing subjects construct advance bundles with more fruits and vegetables, fewer sweets and salty snacks, and fewer calories. Non-committing subjects exhibit substantial inconsistencies along these dimensions, exchanging fruits and vegetables for sweets and salty snacks. Committing subjects carry inconsistencies of smaller magnitude, in line with the correlations noted previously.

3.2 Structural Evidence: Welfare Consequences and Policy Evaluation

In this subsection, we use structural estimation to evaluate the utility and welfare consequences of dynamic inconsistency and commitment demand. We also introduce three potential commitment policies and evaluate them on the basis of welfare and behavior change.

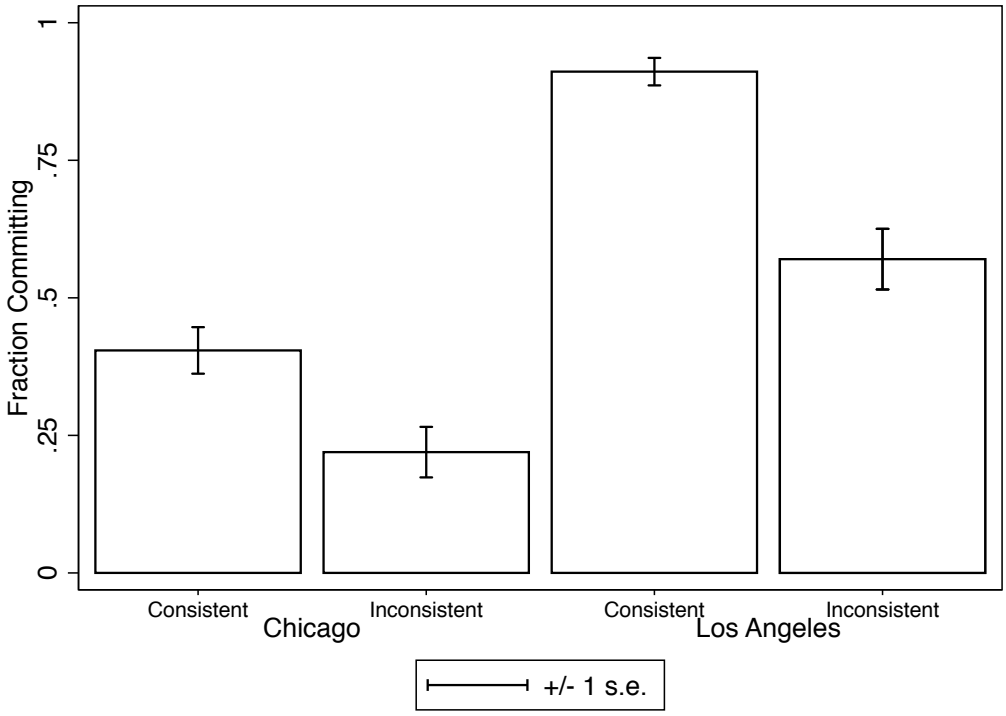


Figure 4: Fraction of Committing Subjects by Prior Inconsistency

Notes: This figure displays the fraction of participants who demand commitment, split by whether they were previously dynamically inconsistent.

3.2.1 Welfare Consequences of Dynamic Inconsistency

In Section 2.4, we used a random utility model to link food choices at each point in time, summarized by the advance and immediate orderings, r_A and r_I, to utility parameters, β_A and β_I. Table 4 (top panel) provides structural estimates for each study site. We assume that observable characteristics, such as being a fruit or vegetable and being perishable, and nutritional characteristics, such as grams of fat, carbohydrates and protein, are potential utility drivers.[20] We stack all orderings obtained when commitment is not available and estimate β_A and β_I simultaneously following the likelihood established in equation (1). Standard errors are clustered by individual, but week-location controls are not included given the formulation of covariates as utility drivers.

We estimate preferences for Chicago subjects in column (1), preferences for Los Angeles subjects in column (2) and preferences in the pooled data in column (3). Results

[20]Calories are not included as a utility driver as they are collinear with nutritional characteristics. There are 9 calories in 1 fat gram, 4 calories in 1 carbohydrate gram and 4 calories in 1 protein gram. Hence, calories = 9*Fat (g) + 4*Carb(g) + 4*Protein(g).

Table 3: Prior Bundle Characteristics and Commitment Demand

	(1) Fruits/Veg	(2) Sweets	(3) Salty Snacks	(4) Calories	(5) Fat (g)	(6) Carb (g)	(7) Protein (g)
Panel A: Chicago Study							
Immediate Choice	-0.290***	0.207***	0.083**	80.200***	5.491***	6.722***	0.333
	(0.044)	(0.039)	(0.033)	(16.773)	(0.933)	(2.517)	(0.206)
Committer	0.444	-0.368*	-0.116	-54.762	-9.502	11.149	-1.168
	(0.288)	(0.205)	(0.163)	(85.118)	(5.914)	(10.540)	(0.974)
Immediate X Committer	0.207***	-0.138***	-0.069	-55.625**	-4.300***	-3.170	0.017
	(0.064)	(0.053)	(0.045)	(22.953)	(1.365)	(3.481)	(0.267)
Constant	5.241***	2.752***	2.007***	2742.228***	92.840***	458.503***	39.806***
	(0.175)	(0.133)	(0.096)	(49.544)	(3.378)	(6.473)	(0.522)
# Observations	436	436	436	436	436	436	436
# Subjects	218	218	218	218	218	218	218
Panel B: Los Angeles Study							
Immediate Choice	-0.281**	0.297***	-0.016	108.238	7.522*	6.921	3.124**
	(0.129)	(0.111)	(0.093)	(72.759)	(3.897)	(10.999)	(1.276)
Committer	0.774**	-0.657**	-0.121	-280.291*	-16.782*	-24.072	-5.461**
	(0.310)	(0.255)	(0.135)	(150.516)	(8.587)	(19.605)	(2.753)
Immediate X Committer	0.151	-0.208*	0.057	-67.403	-5.678	-0.890	-2.710**
	(0.134)	(0.117)	(0.097)	(76.501)	(4.086)	(11.559)	(1.349)
Constant	6.136***	2.782***	1.080***	3575.314***	80.862***	684.206***	59.920***
	(0.275)	(0.231)	(0.116)	(130.723)	(7.400)	(17.506)	(2.400)
# Observations	512	512	512	512	512	512	512
# Subjects	171	171	171	171	171	171	171
Week Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel C: Pooled Data							
Immediate Choice	-0.287***	0.234***	0.053	88.786***	6.113***	6.783*	1.187***
	(0.050)	(0.043)	(0.037)	(25.162)	(1.359)	(3.784)	(0.437)
Committer	0.612***	-0.522***	-0.113	-170.697*	-13.365**	-6.557	-3.559**
	(0.211)	(0.163)	(0.106)	(86.895)	(5.223)	(11.116)	(1.483)
Immediate X Committer	0.170***	-0.151***	-0.019	-52.430*	-4.449***	-1.435	-0.791
	(0.058)	(0.052)	(0.043)	(30.711)	(1.646)	(4.621)	(0.542)
Constant	6.258***	2.684***	1.069***	3493.293***	78.407***	670.763***	58.647***
	(0.205)	(0.166)	(0.099)	(89.460)	(5.149)	(12.133)	(1.590)
# Observations	948	948	948	948	948	948	948
# Subjects	389	389	389	389	389	389	389
Week Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Location Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Ordinary least squares regression. Standard errors clustered on individual level in parentheses. Levels of significance: * 0.10, ** 0.05, *** 0.01.

are remarkably similar across study sites. The vector of utility weights, β_A, shows that fat significantly decreases a food's value while carbohydrates and protein are weighed positively. Controlling for nutritional characteristics, a food being a fruit or vegetable has positive utility weight. Interaction effects identify whether food characteristics are weighed differently in immediate choice, estimating the difference between β_A and β_I. Echoing the reduced form evidence on inconsistencies, the utility weight of fruits

Table 4: Utility Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
		All Subjects			*Inconsistent Subjects*	
	Chicago	Los Angeles	Pooled	Chicago	Los Angeles	Pooled
Fruit/Vegetable	0.043	0.073**	0.229***	0.064	0.091*	0.217***
	(0.048)	(0.031)	(0.028)	(0.084)	(0.048)	(0.041)
Perishable		0.491***			0.398***	
		(0.038)			(0.053)	
Fat	-0.007***	-0.002**	-0.004***	-0.007*	-0.002	-0.005***
	(0.002)	(0.001)	(0.001)	(0.004)	(0.002)	(0.001)
Carbohydrates	0.001**	0.003***	0.002***	0.000	0.003***	0.002***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Protein	0.031***	-0.008**	-0.001	0.037***	-0.010*	0.000
	(0.006)	(0.004)	(0.003)	(0.010)	(0.005)	(0.004)
Immediate Choice						
X Fruit/Vegetable	-0.072***	-0.051***	-0.050***	-0.200***	-0.114***	-0.117***
	(0.015)	(0.013)	(0.008)	(0.036)	(0.027)	(0.017)
X Perishable		-0.010			-0.013	
		(0.012)			(0.025)	
X Fat	-0.001	-0.000	0.000	-0.002	-0.000	0.001
	(0.001)	(0.001)	(0.000)	(0.002)	(0.001)	(0.001)
X Carbohydrates	0.001***	-0.000	-0.000	0.001***	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
X Protein	-0.004	0.001	-0.001	-0.012	0.004	-0.002
	(0.003)	(0.002)	(0.001)	(0.008)	(0.004)	(0.003)
# Observations	8720	10240	18960	3280	4840	8120
# Rankings	436	512	948	164	242	406
# Clusters	218	171	389	82	95	177
Log-Likelihood	-18437.60	-21306.22	-39949.47	-6934.57	-10117.24	-17121.39
H_0: Dynamic Consistency	$\chi^2(4) = 47.63$ ($p < 0.01$)	$\chi^2(5) = 29.60$ ($p < 0.01$)	$\chi^2(4) = 67.89$ ($p < 0.01$)	$\chi^2(4) = 73.33$ ($p < 0.01$)	$\chi^2(5) = 34.63$ ($p < 0.01$)	$\chi^2(4) = 85.43$ ($p < 0.01$)
$V_A(\mathbf{q_A})$	1.347	4.632	2.215	1.274	4.123	2.173
$V_A(\mathbf{q_I})$	1.327	4.562	2.168	1.200	3.991	2.064
$\frac{V_A(\mathbf{q_A})-V_A(\mathbf{q_I})}{V_A(\mathbf{q_A})}$	0.015	0.015	0.021	0.058	0.032	0.050
$V_I(\mathbf{q_A})$	0.966	4.161	1.890	0.182	3.158	1.399
$V_I(\mathbf{q_I})$	0.961	4.101	1.853	0.214	3.069	1.343
$\frac{V_I(\mathbf{q_A})-V_I(\mathbf{q_I})}{V_I(\mathbf{q_A})}$	0.005	0.014	0.020	-0.176	0.028	0.040
$V_U(\mathbf{q_A})$	1.187	4.640	2.125	0.769	4.083	1.942
$V_U(\mathbf{q_I})$	1.174	4.571	2.081	0.747	3.955	1.850
$\frac{V_U(\mathbf{q_A})-V_U(\mathbf{q_I})}{V_U(\mathbf{q_A})}$	0.011	0.015	0.021	0.029	0.031	0.047

Notes: Rank Order Logit regression results. Standard errors clustered on individual level in parentheses. Week and location controls are not included given the formulation of covariates as utility drivers. Calories not included as a utility driver as they are collinear with nutritional characteristics. Levels of significance: * 0.10, ** 0.05, *** 0.01. Null hypothesis tests stationarity of preferences from interacted rank order Logit regression of choices on nutritional characteristics with different coefficients for immediate choice. Test corresponds to all interaction terms being equal to zero.

and vegetables decreases significantly – by around 150% – from advance to immediate choice. Importantly, as can be seen from column (2) which uses Los Angeles data and incorporates perishability, inconsistencies do not appear to be linked to perishability, with perishable items receiving indistinguishable weight under both β_A and β_I. These results help to ensure that the possible spoilage of foods does not drive aggregate results of dynamic inconsistency. The hypothesis test of dynamic consistency, $\beta_A = \beta_I$,

which corresponds to a test of all interaction terms being equal to zero, is rejected at all conventional levels – $\chi^2(4) = 47.63,\ (p < 0.01)$ in Chicago and $\chi^2(4) = 29.6,\ (p < 0.01)$ in Los Angeles.

Columns (4) through (6) of Table 4 repeat the structural analysis for the subgroup of inconsistent subjects (203 of 474 allocation observations and 177 of 389 total subjects). Though inconsistent subjects are similar to the full sample in terms of advance preferences, immediate preferences show stark reductions in the value of fruits and vegetables. Relative to advance preferences, the utility weight of fruits and vegetables declines by around 50% for inconsistent subjects.

Table 4 (bottom panel) provides an initial examination of the aggregate welfare consequences of dynamic inconsistency. We evaluate the advance and immediate bundles, $\mathbf{q_A}$ and $\mathbf{q_I}$, under the three utility measures presented in Section 2.4: $V_A,\ V_I$ and V_U.[21] The welfare costs from dynamic inconsistency under advance preferences, $\frac{V_A(\mathbf{q_A})-V_A(\mathbf{q_I})}{V_A(\mathbf{q_A})}$, are 1-2% in general and 3-6% for inconsistent subjects. The welfare costs under the immediate preferences, $\frac{V_I(\mathbf{q_A})-V_I(\mathbf{q_I})}{V_I(\mathbf{q_A})}$, are smaller in percentage terms and negative for inconsistent subjects in Chicago. Under the unambiguous preference measure, V_U, we find intermediate utility consequences of dynamic inconsistency. Directionally, inconsistency is costly, but the estimates are generally less extreme than those identified under either advance or immediate preferences.[22]

Though the aggregate results of Table 4 are helpful for understanding time inconsistency and welfare under homogeneity assumptions, welfare consequences are best evaluated on an individual basis. Individual analysis allows for comparison of welfare costs based only on a single subject's estimated preferences.[23] We estimate equation (1) at the individual level using the advance, immediate and unambiguous orderings, $r_{A,i},\ r_{I,i}$ and $r_{U,i}$. Every allocation is considered in isolation such that subjects who make two

[21] Estimates for β_U constructed by eliminating exchanged foods are provided in Appendix Table A2.

[22] One might expect the immediate values, $\frac{V_I(\mathbf{q_A})-V_I(\mathbf{q_I})}{V_I(\mathbf{q_A})}$, to be negative, with the immediate bundle having a higher utility value under immediate preferences. Similarly, one might expect the unambiguous values, $\frac{V_U(\mathbf{q_A})-V_U(\mathbf{q_I})}{V_U(\mathbf{q_A})}$, to be effectively zero, making no statement as to relative value of advance and immediate bundles. These points highlight an important aspect of our aggregate estimation strategy. We infer utility values from the body of chosen and unchosen foods. The fact that on aggregate, fruits and vegetables remain frequently chosen in immediate choice leads to apparent utility decreases from inconsistencies that tend to replace fruits and vegetables with sweets and salty snacks. Where this is not the case — i.e., in Chicago for inconsistent subjects — the welfare evaluation differs depending on the perspective taken. For inconsistent subjects in Chicago, advance bundles have a majority of fruits and vegetables (5.54 (clustered s.e. = 0.21) fruits and vegetables out of 10 items), while immediate bundles have a minority of fruits and vegetables (4.96 (0.21) fruits and vegetables out of 10 items).

[23] Beggs et al. (1981) also provide individual estimates for stated preferences over electric cars and compare individual and aggregate results.

allocation decisions in the Los Angeles study site will have two values of each of $r_{A,i},\ r_{I,i}$ and $r_{U,i}$. The individual rank order logit follows the form of Table 3, column (6) with 'Fruit/Vegetable,' 'Fat,' 'Carbohydrates' and 'Protein' as utility drivers. From this, we construct individual measures of welfare consequences of dynamic inconsistency under all three preference orderings.

Figure 5, Panel A provides histograms of $\frac{V_{A,i}(\mathbf{q_{A,i}})-V_{A,i}(\mathbf{q_{I,i}})}{|V_{A,i}(\mathbf{q_{A,i}})|}$ and $\frac{V_{I,i}(\mathbf{q_{A,i}})-V_{I,i}(\mathbf{q_{I,i}})}{|V_{I,i}(\mathbf{q_{A,i}})|}$, the individual advance and immediate welfare consequences of dynamic inconsistency, for the 203 (of 474 total) inconsistent observations.[24] There is wide heterogeneity both between and within welfare measures for the consequences of dynamic inconsistency. Under advance preferences, $\frac{V_{A,i}(\mathbf{q_{A,i}})-V_{A,i}(\mathbf{q_{I,i}})}{|V_{A,i}(\mathbf{q_{A,i}})|}$ has a median [25th-75th percentile] value of 0.044 [-0.031, 0.146]. Under immediate preferences, $\frac{V_{I,i}(\mathbf{q_{A,i}})-V_{I,i}(\mathbf{q_{I,i}})}{V_{I,i}(\mathbf{q_{A,i}})}$ has a median [25th-75th percentile] value of -0.055 [-0.169, 0.026]. The median disagreement between advance and immediate preferences is intuitive. Advance preferences suggest costs to inconsistency and immediate preferences suggest benefits to flexibility. Indeed, there is broad distributional disagreement in the advance and immediate welfare measures, with greater costs to inconsistency under the advance welfare measure and greater benefits to flexibility under the immediate measure, *Mann-Whitney* $z = 10.13,\ (p < 0.01)$.

Figure 5, Panel B relates advance and immediate welfare measures for the inconsistent observations. Though this relationship generally falls below the 45° line of perfect agreement, a significant correlation does exist, $\rho = 0.28,\ (p < 0.01)$. The line of best fit highlights the general pattern of disagreement, with immediate welfare measures tending to suggest more benefits to flexibility than advance measures. Sixty-eight of 203 individual observations (33.5%) exhibit disagreement in sign between advance and immediate measures. All but 1 of these 68 disagreements are in the direction of the medians, with 67 observations having $\frac{V_{A,i}(\mathbf{q_{A,i}})-V_{A,i}(\mathbf{q_{I,i}})}{|V_{A,i}(\mathbf{q_{A,i}})|} > 0$ and $\frac{V_{I,i}(\mathbf{q_{A,i}})-V_{I,i}(\mathbf{q_{I,i}})}{|V_{I,i}(\mathbf{q_{A,i}})|} < 0$. The welfare measures for the remaining 135 observations agree in sign, with 69 (34.0%) exhibiting unanimous costs to dynamic inconsistency, and 66 (32.5%) exhibiting unanimous benefits to flexibility.

Figure 6, Panel A presents the unambiguous individual welfare measure, $\frac{V_{U,i}(\mathbf{q_{A,i}})-V_{U,i}(\mathbf{q_{I,i}})}{|V_{U,i}(\mathbf{q_{A,i}})|}$, constructed from the unambiguous orderings, $r_{U,i}$. The unambiguous welfare measure has median [25th-75th percentile] value of 0.003 [-0.106, 0.107], with 102 of 203 (50.3%) observations exhibiting unambiguous welfare costs to inconsistency. As in the aggregate exercise, individuals' unambiguous choices imply intermediate wel-

[24]The absolute value of the denominator is used because a small number of observations have estimated utility parameters that imply negative bundle values. The absolute value ensures that we correctly capture the direction of change for our proportional measure. The utility measures are top and bottom-coded at +/- 1.

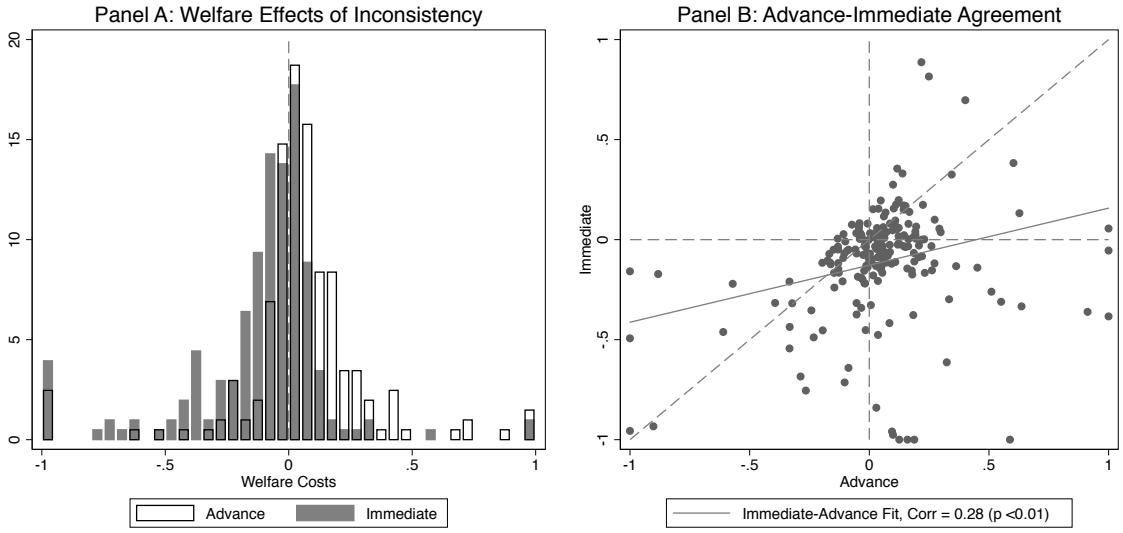


Figure 5: Advance and Immediate Welfare Consequences of Dynamic Inconsistency

Notes: Panel A provides a histogram of individual estimates of the welfare costs of inconsistency under advance and immediate preferences. Panel B provides a scatterplot of agreement between advance and immediate welfare measures for the inconsistent observations.

fare effects of dynamic inconsistency. However, relying only on an individual's own preference ranking rather than the body of aggregate choices brings the central tendency of unambiguous welfare consequences closer to zero.

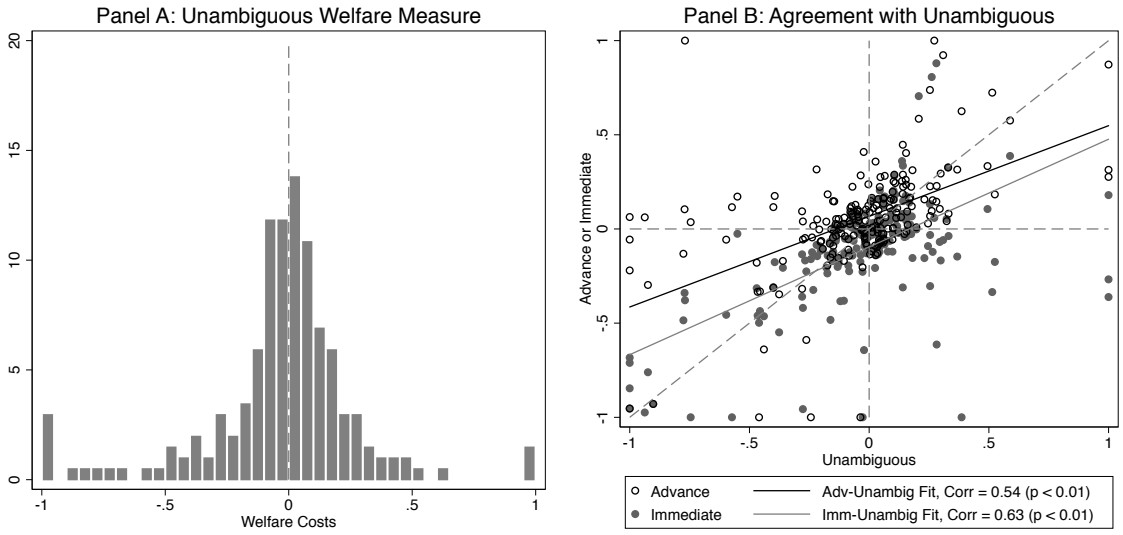


Figure 6: Unambiguous Welfare Consequences of Dynamic Inconsistency

Notes: Panel A provides a histogram of individual estimates of the welfare costs of inconsistency under unambiguous preferences. Panel B provides a scatterplot of agreement between unambiguous and advance or immediate welfare measures.

When advance and immediate welfare measures agree, so too does the unambiguous measure with all 135 values of $\frac{V_{U,i}(\mathbf{q_{A,i}})-V_{U,i}(\mathbf{q_{I,i}})}{|V_{U,i}(\mathbf{q_{A,i}})|}$ sharing the same sign. When the advance and immediate welfare measures disagree, 34 of 68 (50%) have values of $\frac{V_{U,i}(\mathbf{q_{A,i}})-V_{U,i}(\mathbf{q_{I,i}})}{|V_{U,i}(\mathbf{q_{A,i}})|} > 0$, implying welfare costs to inconsistency.

Figure 6, Panel B relates the unambiguous to the advance and immediate welfare measures. Both the advance and immediate welfare measures are substantially more correlated with the unambiguous measure than they are with each other, $\rho = 0.54$ ($p < 0.01$) and $\rho = 0.63$ ($p < 0.01$), respectively. These patterns of connection are intuitive: though disagreement exists between advance and immediate orderings, their commonalities are respected by the unambiguous ordering, and hence, the unambiguous welfare measure. Further, the lines of best fit highlight the general tendency of advance measures to exceed, and immediate measure to fall below, the unambiguous welfare consequences of inconsistency.

We next evaluate the utility estimates and welfare consequences separately for committing and non-committing subjects.[25] For each welfare measure we calculate whether the advance and immediate bundles carry equal value (e.g., $\frac{V_{A,i}(\mathbf{q_{A,i}})-V_{A,i}(\mathbf{q_{I,i}})}{|V_{A,i}(\mathbf{q_{A,i}})|} = 0$ for advance preferences), whether there are benefits to flexibility (e.g., $\frac{V_{A,i}(\mathbf{q_{A,i}})-V_{A,i}(\mathbf{q_{I,i}})}{|V_{A,i}(\mathbf{q_{A,i}})|} < 0$) or whether there are costs to inconsistency (e.g., $\frac{V_{A,i}(\mathbf{q_{A,i}})-V_{A,i}(\mathbf{q_{I,i}})}{|V_{A,i}(\mathbf{q_{A,i}})|} > 0$). Figure 7 provides corresponding results.

By all three measures, committing subjects are disproportionately represented in the group with equal bundle values. For example, under the advance welfare measure, 178 of 265 (67%) of committing observations have equal advance and immediate bundle values compared to 93 of 209 (45%) of non-committing observations.[26] Further, for inconsistent subjects, commitment choice does not seem to target those with costs to inconsistency. Only 59 of 265 (22%) committing observations exhibit costs to inconsistency under the advance welfare measure, while 77 of 209 (37%) of non-committing observations exhibit such costs.[27]

[25] In Appendix Table A3 we re-conduct the aggregate utility estimation separately for committing and non-committing subjects. These aggregate utility estimates echo the reduced form results. We observe differences in advance preferences β_A, across the two groups, $\chi^2(4) = 46.15$ ($p < 0.01$). Non-committing subjects appear to value fruits and vegetables less than committing subjects in advance choice and also experience greater declines in value when comparing advance and immediate preferences. For both study sites separately and combined, we reject the null hypothesis of equal immediate preferences, β_I across committing and non-committing groups. Additionally, subjects who ultimately demand commitment experience smaller aggregate welfare consequences from inconsistency regardless of the measure.

[26] Pearson's χ^2 test for the independence of commitment and whether subjects have costs to inconsistency, benefits to flexibility or equal bundle values under the advance measure yield a test statistic of $\chi^2 = 24.6$, $(p < 0.01)$.

[27] Seen from the other margin, of 136 observations with costs to inconsistency under the advance

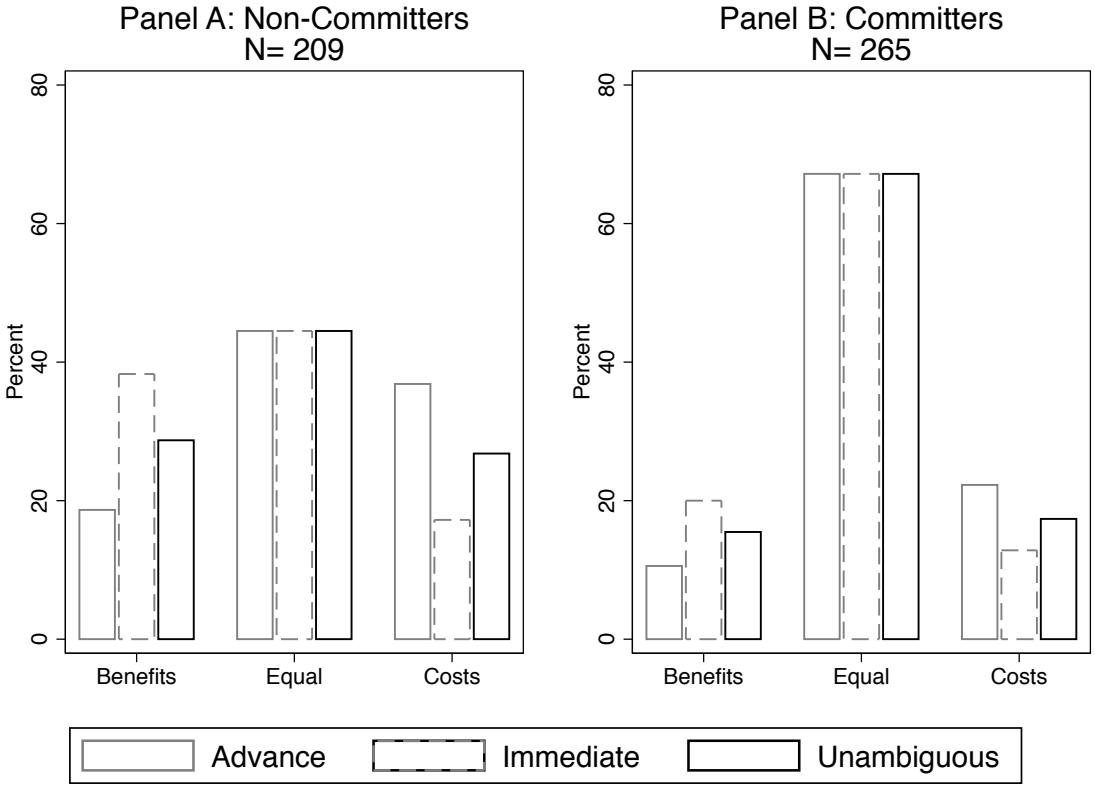


Figure 7: Commitment and Prior Costs to Inconsistency

Notes: This figure displays the percentage of individuals for whom the advance and immediate bundles carry equal value (equal), whether there are benefits to flexibility (benefits), or whether there are costs to inconsistency (costs), separately for each welfare criterion.

Our analysis yields a clear connection between commitment demand and prior dynamic inconsistency. At both study sites, inconsistency is negatively correlated with subsequent commitment. Corresponding welfare analysis shows that by all three utility measures, those individuals who subsequently demand commitment experience smaller welfare effects driven by their relative infrequency of inconsistency.

3.2.2 Policy Evaluation

We predict the welfare effects of commitment policies by contrasting the costs of dynamic inconsistency through the lens of the individual preference measures, $\beta_{A,i}$, $\beta_{I,i}$ and $\beta_{U,i}$. Specifically, we evaluate the proportion of observations where costs of inconsistency — e.g., $\frac{V_{A,i}(\mathbf{q}_{A,i})-V_{A,i}(\mathbf{q}_{I,i})}{|V_{A,i}(\mathbf{q}_{A,i})|}$ for advance utility — are predicted to increase, decrease or remain constant under different policies relative to complete flexibility.

measure, 59 (43.4%) commit and the remainder do not.

Figure 7 provided an initial examination of the policy of simply offering commitment. All non-committing and dynamically consistent individuals are predicted to be unaffected by such a policy. Given that 87 of 474 (18.4%) observations are both dynamically inconsistent and associated with subsequent commitment, offering commitment is predicted to affect a minority of individuals. Figure 8, Panel A summarizes the overall effects for each welfare measure. Even under the most favorable welfare criterion, the advance measure, only 59 of 474 (12.5%) observations would see welfare improvements from commitment, while 28 (5.9%) would see welfare reductions.[28]

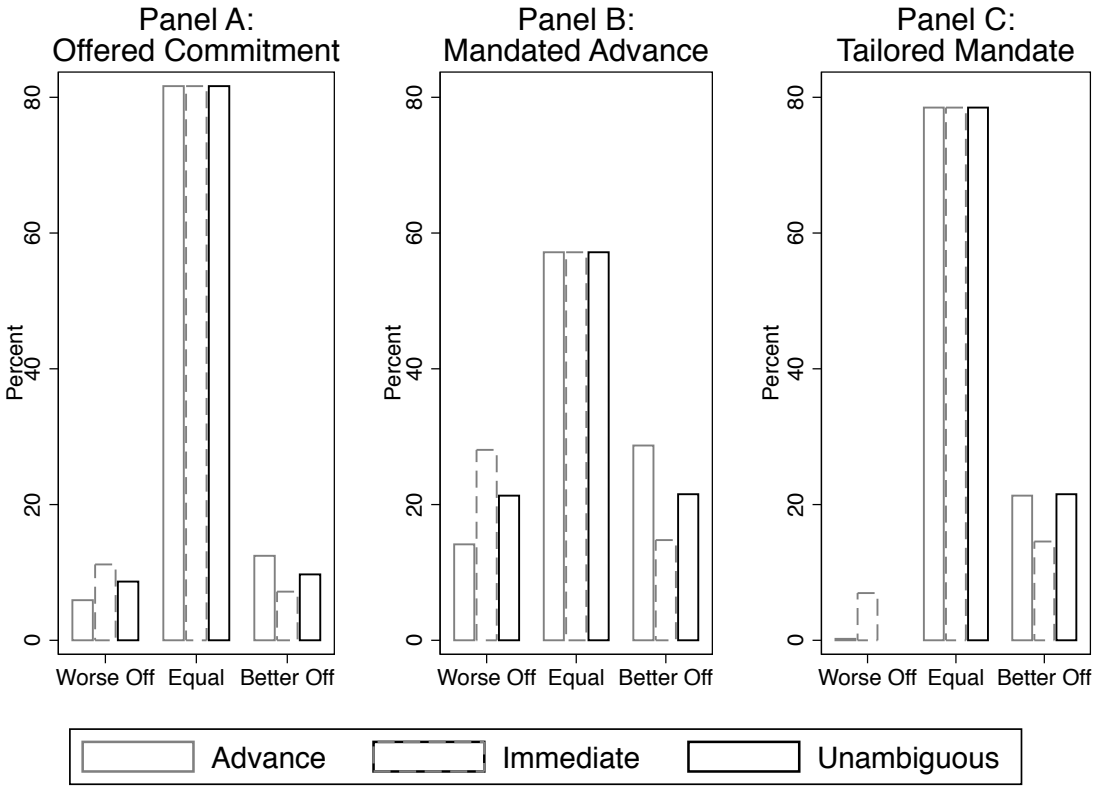


Figure 8: Policy Evaluation

Notes: This figure summarizes the percentage of individuals who would be worse, equal or better off under advance, immediate and unambiguous welfare criteria for each policy. Panel A displays the policy of simply offering commitment, Panel B displays the policy of mandated advance choice and Panel C displays the policy of a tailored mandate.

Figure 8, Panel B illustrates an alternative policy of mandated advance choice. As in Panel A, whether someone benefits from mandated advance choice depends on whether dynamic inconsistency is estimated to have positive, negative or no costs in the absence of the policy. Under advance preferences, 136 of 474 observations (28.7%) are

[28]Note that these calculations do not incorporate actual choices made when commitment is offered.

made better off by mandated advance choice, 67 (14.1%) are made worse off and 271 (57.2%) are equally well off. The other two welfare measures alter the relative portion of winners and losers from mandated advance choice, with those who are harmed by the program exceeding (almost identical to) those who benefit under the immediate (unambiguous) preferences.

Figure 8, Panel C provides a final policy of a tailored mandate. Based on the unambiguous utility measure, the policy mandates advance choice for any observation with $\frac{V_{U,i}(\mathbf{q}_{A,i})-V_{U,i}(\mathbf{q}_{I,i})}{|V_{U,i}(\mathbf{q}_{A,i})|} > 0$ and mandates flexibility for any observation with $\frac{V_{U,i}(\mathbf{q}_{A,i})-V_{U,i}(\mathbf{q}_{I,i})}{|V_{U,i}(\mathbf{q}_{A,i})|} \leq 0$. In effect, this policy honors the unambiguous preferences, $\beta_{U,i}$, and tailors contract terms depending on whether dynamic inconsistencies are estimated to be detrimental or beneficial. Of 474 observations, 102 (21.5%) have $\frac{V_{U,i}(\mathbf{q}_{A,i})-V_{U,i}(\mathbf{q}_{I,i})}{|V_{U,i}(\mathbf{q}_{A,i})|} > 0$ and so would have their advance choice mandated, while 372 (80%) have $\frac{V_{U,i}(\mathbf{q}_{A,i})-V_{U,i}(\mathbf{q}_{I,i})}{|V_{U,i}(\mathbf{q}_{A,i})|} \leq 0$ and would have flexibility mandated. Under this policy, no subjects can be worse off according to the unambiguous measure, and all observations with $\frac{V_{U,i}(\mathbf{q}_{A,i})-V_{U,i}(\mathbf{q}_{I,i})}{|V_{U,i}(\mathbf{q}_{A,i})|} > 0$ are made better off. However, subjects may grow better or worse off under the alternate measures. Under advance preferences, 101 of 474 observations (21.3%) are made better off, 1 (0.2%) is made worse off, and the remaining 372 (78.5%) are equally well off. Under immediate preferences, 69 of 474 observations (14.6%) are made better off, 33 (7%) are made worse off and the remaining 372 (78.5%) are equally well off.

The tailored policy carries potential benefits over simply offering commitment or mandating advance choice for all. The proportion of better off individuals exceeds those from offering commitment, and the relative proportion of winners to losers is greater regardless of the preference measure. Relative to mandated advance choice for all, the tailored mandate dramatically reduces the proportion of individuals who are negatively affected by the policy, while maintaining a sizable proportion of beneficiaries.

Policymakers weigh more than just distributional welfare consequences of commitment policies. For example, they may wish to increase the number of fruits and vegetables consumed and reduce inconsistencies towards less healthful choices. In Table 5, we evaluate the predicted behavioral consequences of each policy presented in Figure 8. We examine how the foods that individuals end up with are influenced by the combination of policy and individual decisions.[29] We report estimated means for each

[29]Without a commitment policy, foods individuals end up with will be identical to those in immediate choice. Under offered commitment, foods are identical to those in advance choice for those who choose to commit and identical to immediate choice for those who don't. Under mandated advance choice, foods are identical to those in advance choice. Under the tailored mandate, foods are identical to advance choice for those individuals with $\frac{V_{U,i}(\mathbf{q}_{A,i})-V_{U,i}(\mathbf{q}_{I,i})}{|V_{U,i}(\mathbf{q}_{A,i})|} > 0$ and identical to immediate choice

outcome and policy with standard errors clustered by individual using choices made when commitment is not available.

Table 5: Behavioral Evaluation

	(1) Fruits/Veg	(2) Sweets	(3) Salty Snacks	(4) Calories	(5) Fat (g)	(6) Carb (g)	(7) Protein (g)
No Intervention	5.857	2.648	1.487	3166.497	83.112	583.722	49.512
	(0.103)	(0.078)	(0.057)	(46.058)	(2.449)	(8.022)	(0.871)
Offer Commitment	5.922	2.601	1.468	3146.171	82.182	580.732	49.291
	(0.102)	(0.076)	(0.057)	(45.521)	(2.455)	(7.914)	(0.863)
Mandated Advance Choice	6.049	2.498	1.445	3107.023	79.486	577.741	48.767
	(0.099)	(0.073)	(0.055)	(43.650)	(2.338)	(7.761)	(0.801)
Tailored Mandate	5.973	2.568	1.451	3145.916	80.921	584.073	49.084
	(0.105)	(0.077)	(0.057)	(47.037)	(2.514)	(8.205)	(0.861)

Notes: The table reports predicted means using the relevant individual choices (advance or immediate) for each policy. Standard errors clustered at individual level reported in parentheses.

We predict that offering commitment would generate 0.065 more fruits and vegetables and about 20 fewer calories. Mandating advance choice leads to more perceptible differences, 0.2 more fruits and vegetables and 60 fewer calories.[30] The tailored mandate has intermediate behavioral effects with 0.11 more fruits and vegetables and 20 fewer calories. Recall that we restricted the choice set in our study to 20 foods and up to 4 exchanges – health effects may be larger in a setting with more foods with greater likelihood for temptation and exchange.

Our analysis demonstrates that people who subsequently demand commitment are less likely to exhibit dynamic inconsistencies, and have less inconsistency in their estimated preferences. Table 5 showed that the behavioral effects of a policy built on offering commitment are predicted to be limited. In Table 6, we examine whether our program of offering commitment alters consumption patterns. We augment the prior data of Table 2 with the final week(s) of decisions in which commitment was offered. At both study sites, offering commitment has virtually no effect on the extent of inconsistencies in terms of observable and nutritional characteristics. We find that subjects choose advance bundles containing fewer fruits and vegetables in the final weeks(s) of the study when commitment is offered (but before we tell subjects about the commitment offer). However, offering commitment has no effect on dynamic inconsistencies. This is despite the fact that 207 of 389 subjects (53.2%) ever demand commitment.

for those with $\frac{V_{U,i}(\mathbf{q}_{A,i})-V_{U,i}(\mathbf{q}_{I,i})}{|V_{U,i}(\mathbf{q}_{A,i})|} \leq 0$.

[30]These differences are identical to those estimated in Table 2 Panel C.

Table 6: Behavior and Offered Commitment

	(1) Fruits/Veg	(2) Sweets	(3) Salty Snacks	(4) Calories	(5) Fat (g)	(6) Carb (g)	(7) Protein (g)
Panel A: Chicago Study							
Immediate Choice	-0.220***	0.161***	0.060**	61.573***	4.051***	5.661***	0.338**
	(0.034)	(0.029)	(0.024)	(12.436)	(0.717)	(1.857)	(0.148)
Commitment Offered	-0.450***	0.390***	0.050	112.917***	7.335***	9.931**	0.653
	(0.117)	(0.098)	(0.077)	(35.987)	(2.461)	(4.883)	(0.406)
Immediate X	0.028	-0.009	-0.018	-22.982	-0.720	-4.204*	0.133
Commitment Offered	(0.045)	(0.043)	(0.038)	(16.520)	(0.947)	(2.530)	(0.189)
Constant	5.390***	2.628***	1.968***	2723.890***	89.658***	462.236***	39.414***
	(0.140)	(0.103)	(0.078)	(40.256)	(2.785)	(5.132)	(0.444)
# Observations	872	872	872	872	872	872	872
# Subjects	218	218	218	218	218	218	218
Panel B: Los Angeles Study							
Immediate Choice	-0.168***	0.141***	0.027	57.686**	3.263**	6.254	1.092**
	(0.042)	(0.039)	(0.031)	(25.608)	(1.359)	(3.826)	(0.474)
Commitment Offered	-0.394***	0.427***	-0.027	125.160*	1.094	27.289**	1.268
	(0.117)	(0.107)	(0.076)	(72.670)	(3.600)	(11.416)	(1.219)
Immediate X	0.074	-0.100**	0.025	13.410	-0.163	3.541	-0.082
Commitment Offered	(0.053)	(0.046)	(0.038)	(34.015)	(1.827)	(5.107)	(0.595)
Constant	6.745***	2.263***	0.986***	3354.537***	67.616***	665.328***	55.596***
	(0.116)	(0.099)	(0.060)	(60.223)	(3.156)	(8.925)	(1.071)
# Observations	854	854	854	854	854	854	854
# Subjects	171	171	171	171	171	171	171
Week Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Panel C: Pooled Data							
Immediate Choice	-0.192***	0.150***	0.042**	59.474***	3.626***	5.981***	0.745***
	(0.028)	(0.025)	(0.020)	(14.935)	(0.803)	(2.232)	(0.265)
Commitment Offered	-0.378***	0.401***	-0.017	135.162*	1.211	29.489***	1.246
	(0.115)	(0.107)	(0.072)	(70.961)	(3.467)	(11.145)	(1.195)
Immediate X	0.043	-0.047	0.004	-6.594	-0.396	-0.859	-0.037
Commitment Offered	(0.035)	(0.032)	(0.028)	(18.834)	(1.026)	(2.839)	(0.312)
Constant	6.757***	2.258***	0.979***	3353.643***	67.435***	665.464***	55.769***
	(0.116)	(0.098)	(0.060)	(59.517)	(3.119)	(8.804)	(1.064)
# Observations	1726	1726	1726	1726	1726	1726	1726
# Subjects	389	389	389	389	389	389	389
Location X Week Control	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Ordinary least squares regression. Standard errors clustered on individual level in parentheses. Levels of significance: * 0.10, ** 0.05, *** 0.01.

3.3 Robustness Tests and Additional Exercises

Our exercise interprets dynamically inconsistent behavior as evidence of dynamically inconsistent preferences. Though our structural exercise examines the possibility that inconsistent behavior exists with consistent preferences, this is done through the lens of the model. In this sub-section, we provide additional evidence that dynamic inconsis-

tency is a product of preferences rather than an alternate force such as the resolution of uncertainty, changing environmental factors or noise. We also evaluate the stability of inconsistency and commitment.

3.3.1 'Want' versus 'Should' Foods

Models of dynamically inconsistent preferences are often organized around a narrative of temptation. There are foods decision-makers should be consuming and those that they want to consume. In our Los Angeles study site, we provided subjects with two forms of food rating data. In the first, subjects were asked how much they liked eating the food, including aspects such as how the food tastes.[31] We term this the 'want' ranking. In the second, subjects were asked how often they felt they should eat each food.[32] We term this the 'should' ranking.

Table 7 follows the structural exercise from actual food choices to contrast the preferences implied by the 'want' and 'should' rankings. Column (1) shows differences between 'want' and 'should' preferences in line with choices. Fruits and vegetables are valued according to 'should' preferences, but receive lower weight in 'want' preferences. In column (2), we restrict attention to the 125 Los Angeles subjects who provided both 'want' and 'should' rankings for all foods and find similar results. In columns (3)-(6), we examine differences in 'want' and 'should' preferences by commitment choice and dynamic inconsistency. Interestingly, individuals who are inconsistent and individuals who do not commit have smaller percentage differences in their 'want' and 'should' preferences for fruits and vegetables than those who are consistent and those who demand commitment. These data are in line with the interpretation that those with larger self-control problems are less aware thereof and hence are less likely to commit.

3.3.2 Stability of Inconsistency and Commitment

Our data demonstrate evidence of dynamic inconsistency when comparing advance and immediate decisions. Though the data patterns are indicative of a change in preference rather than shocks, specific forms of resolution of uncertainty may lead to apparent time inconsistencies. For example, perishable foods such as fruits and vegetables may appear less attractive than packaged foods such as sweets and salty snacks on the day of delivery. Our Los Angeles study foods were chosen with this critique in mind. The similarity in results between Chicago and Los Angeles helps alleviate this concern.

[31] This rating was provided on a 1-7 scale from 'Dislike Very Much' to 'Like Very Much.'
[32] This rating was provided on a 1-5 scale from 'Never' to 'Every Day.'

Table 7: 'Want' Versus 'Should' Utility Estimates in Los Angeles Study

	(1)	(2)	(3)	(4)	(5)	(6)
	All Subjects		*Complete Rankings*			
			Inconsistent = 0	Inconsistent = 1	Commit = 0	Commit = 1
Fruit/Vegetable	0.391***	0.423***	0.452***	0.400***	0.522**	0.398***
	(0.071)	(0.087)	(0.127)	(0.119)	(0.204)	(0.097)
Perishable	1.578***	1.532***	1.508***	1.551***	1.223***	1.623***
	(0.104)	(0.116)	(0.169)	(0.161)	(0.249)	(0.131)
Fat	-0.013***	-0.014***	-0.013***	-0.015***	-0.011***	-0.015***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)
Carbohydrates	0.002***	0.003***	0.003***	0.002***	0.001	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Protein	0.010**	0.012**	0.009	0.015*	0.022	0.009*
	(0.005)	(0.005)	(0.006)	(0.008)	(0.014)	(0.006)
Want Ranking						
X Fruit/Vegetable	-0.356***	-0.404***	-0.510***	-0.309**	-0.238	-0.437***
	(0.091)	(0.109)	(0.160)	(0.151)	(0.244)	(0.123)
X Perishable	-0.553***	-0.505***	-0.262	-0.704***	-0.680***	-0.468***
	(0.107)	(0.122)	(0.175)	(0.165)	(0.243)	(0.141)
X Fat	-0.004**	-0.005**	-0.003	-0.007**	-0.014**	-0.003
	(0.002)	(0.002)	(0.003)	(0.003)	(0.006)	(0.003)
X Carbohydrates	0.000	0.000	0.000	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
X Protein	0.008	0.009	0.003	0.014	0.032*	0.003
	(0.006)	(0.008)	(0.013)	(0.010)	(0.016)	(0.009)
# Observations	6550	5000	2280	2720	1000	4000
# Rankings	331	250	114	136	50	200
# Clusters	171	125	57	68	25	100
Log-Likelihood	-12473.98	-9548.37	-4334.81	-5208.43	-1949.62	-7584.94
H_0: Want = Should	$\chi^2(5) = 78.40$ ($p < 0.01$)	$\chi^2(5) = 62.05$ ($p < 0.01$)	$\chi^2(5) = 24.41$ ($p < 0.01$)	$\chi^2(5) = 43.46$ ($p < 0.01$)	$\chi^2(5) = 21.69$ ($p < 0.01$)	$\chi^2(5) = 47.82$ ($p < 0.01$)

Notes: Rank Order Logit regression results. Standard errors clustered on individual level in parentheses. Levels of significance: * 0.10, ** 0.05, *** 0.01. The 'want' rating was provided on a 1-7 scale from 'Dislike Very Much' to 'Like Very Much.' The 'should' rating was provided on a 1-5 scale from 'Never' to 'Every Day.' Null hypothesis tests equality of 'want' and 'should' preferences from interacted rank order logit regression of choices on nutritional characteristics with different coefficients for 'want' rankings. Test corresponds to all interaction terms being equal to zero.

Additional exercises can be taken to ensure that observed dynamic inconsistencies are not simply driven by changes to environmental factors. First, we can examine whether individuals who are inconsistent at one delivery remain so at future deliveries. Of our 389 subjects, 182 never chose commitment. For these subjects, the correlation between inconsistency before and after commitment is offered is $\rho = 0.33,\ (p < 0.01)$. This positive association through time suggests some stability at the individual level. Additionally, 85 subjects in Los Angeles made two allocation decisions prior to being offered commitment. For these subjects the correlation in dynamic inconsistency over the two weeks is $\rho = 0.20,\ (p = 0.07)$. This lower correlation is driven by a growing tendency of inconsistency over the two weeks: 28 of 34 individuals were inconsistent in the first week were again inconsistent, but 29 of 51 individuals who were not inconsistent in the first week became inconsistent.

Second, we can examine whether changes to the decision environment relate to

observed inconsistencies. For example, for people with children, decisions may be made with or without children present. For 343 of our 389 subjects, we have a survey response to their total number of children. Ninety of 343 (26%) report having no children. The correlation between having no children and dynamic inconsistency prior to commitment is $\rho = -0.05,\ (p = 0.38)$, indicating that those less likely to experience the environmental change of having children present in the household are no more or less likely to exhibit inconsistencies. Further, in Los Angeles, our study staff recorded the number of children present at registration at first delivery for all 171 subjects. The correlation between having more kids present at delivery than registration and dynamic inconsistency is $\rho = 0.03,\ (p = 0.67)$.

Another possible source of environmental change is the decision-maker's current level of hunger. In our Los Angeles study, 170 of 171 subjects rated their current hunger level on a 4-point scale from 'Very Hungry' to 'Not At All Hungry' both at registration and delivery. The correlation between changing one's report to 'Very Hungry' from a lower hunger level and dynamic inconsistency is $\rho = 0.07,\ (p = 0.37)$. Additionally, in our Los Angeles study, we used a series of questions to measure food security – i.e., levels of access to food due to lack of resources – at registration and delivery (Blumberg, Bialostosky, Hamilton and Briefel, 1999). The correlation between growing more food insecure from registration to delivery and dynamic inconsistency is $\rho = 0.003,\ (p = 0.97)$.

A final potential change to the decision environment is the resources available to the decision-maker. In our Los Angeles study, subjects were asked about their remaining Supplemental Nutritional Assistance Plan (SNAP) dollar balance at both registration and delivery. Fifty-seven of 171 Los Angeles subjects provided these reports, and having less available balance at delivery than registration is actually negatively correlated with dynamic inconsistency, though not significantly so, $\rho = -0.19,\ (p = 0.17)$. Taken together, these findings indicate that observable changes in decision environment are unlikely to drive our observed inconsistencies.

Our Los Angeles data also allow us to examine the stability of commitment demand. Eighty-six of our 171 Los Angeles participants were asked if they desired commitment for both their second and third delivery. The correlation between demanding commitment across these two deliveries is $\rho = 0.46,\ (p < 0.01)$. Of the 69 subjects who demanded commitment for their second delivery, 61 subsequently demanded commitment for their third delivery. This gives further indication of commitment as a deliberate choice taken by a set of subjects who have relatively small self-control problems.

4 Discussion and Conclusion

In two field experiments, we provide evidence on dynamic inconsistency and commitment demand in food choice. We show that dynamic inconsistencies are prevalent, with over 40% of subjects exhibiting inconsistency in choice. The direction of inconsistency is systematically towards less healthy foods: compared to advance choice, immediate choice decreases the amount of fruits and vegetables selected and increases calories and fat content. Using structural estimation, we find welfare effects of dynamic inconsistency on the order of around 5% of total utility, with the size of the effect depending on the welfare criterion used.

We also find substantial demand for commitment, with over half of subjects voluntarily restricting themselves to their advance choice. Importantly, we document a negative correlation between dynamic inconsistency and subsequent commitment demand. This suggests that those with the largest self-control problems may lack sufficient awareness to demand commitment.

Our results contrast with prior studies which find a weak positive correlation between commitment demand and present bias. Since our negative correlation is observed in both of our experiments, we believe it is unlikely to be due to chance alone. Instead, it is possible that the different results between our work and prior work are due to the context (we study food choice, other studies focus on other environments), or due to the fact that our study is conducted in a more natural environment, wherein subjects were not told that they were under observation. Existing puzzles related to commitment demand in field settings may benefit from a deeper understanding of this correlation, with our findings providing one key observation.

Interestingly, at both study sites, subjects who demand commitment also make more healthy advance decisions even when commitment is not available. This result resonates with one recent finding on commitment demand in gym attendance by Royer, Stehr and Sydnor (2015), who find greater commitment demand among subjects who are already exercising regularly. These findings suggest that those whose behavior (and welfare) would be most affected by commitment may be the least likely to take it up. More research is needed in field environments to understand the nuanced relationship between preferences, dynamic inconsistency and awareness.

Our research is critical for understanding the behavioral impacts and welfare consequences of commitment policies. In our studies, we use individuals' advance choices, immediate choices and unambiguous choices to evaluate the behavioral and welfare consequences of various policies. An important application is comparing a policy that offers commitment to a policy that mandates advance choice for a subset of individuals with

unambiguous costs to inconsistency. A common concern with mandated advance choice is that while it may have large effects on behavior, it may reduce welfare compared to offering commitment. Our welfare analysis in this context is perhaps surprising. We find that offering commitment does little to change behavior or improve welfare, with those who benefit from the program roughly equalling those who lose depending on the welfare measure. However, a tailored policy of mandated advance choice would increase healthy choices while maintaining a distribution of welfare consequences tilted towards those who benefit from the program under all welfare measures.

Finally, our results give insights to innovations in food policy. For example, our results add to our understanding of the impact of a policy change now under consideration at the USDA that would allow pre-ordering under SNAP. Our study provides an understanding of how this policy change would affect the food choice and welfare of consumers.

References

Allcott, Hunt and Dmitry Taubinsky, "Evaluating behaviorally motivated policy: experimental evidence from the lightbulb market," *The American Economic Review*, 2015, *105* (8), 2501–2538.

_ , Sendhil Mullainathan, and Dmitry Taubinsky, "Energy policy with externalities and internalities," *Journal of Public Economics*, 2014, *112*, 72–88.

Andreoni, James and Charles Sprenger, "Estimating time preferences with convex budgets," *American Economic Review*, 2012, *102* (7), 3333–3356.

Ashraf, Nava, Dean Karlan, and Wesley Yin, "Tying Odysseus to the Mast: Evidence from a Commitment Savings Product in the Philippines," *Quarterly Journal of Economics*, 2006, *121* (1), 635–672.

Augenblick, Ned and Matthew Rabin, "An experiment on time preference and misprediction in unpleasant tasks," *Review of Economic Studies*, forthcoming.

_ , Muriel Niederle, and Charles Sprenger, "Working over time: Dynamic inconsistency in real effort tasks," *The Quarterly Journal of Economics*, 2015, *130* (3), 1067–1115.

Beggs, S., S. Cardell, and J. Hausman, "Assessing the potential demand for electric cars," *Journal of Econometrics*, 1981, *16*, 1–19.

Bernheim, B. Douglas and Antonio Rangel, "Toward choice-theoretic foundations for behavioral welfare economics," *American Economic Review, Papers and Proceedings*, 2007, *97* (2), 464–470.

_ and _ , "Beyond Revealed Preference: Choice-Theoretic Foundations for Behavioral Welfare Economics," *The Quarterly Journal of Economics*, 2009, *124* (1), 51–104.

Bernheim, B Douglas, Andrey Fradkin, and Igor Popov, "The welfare economics of default options in 401 (k) plans," *American Economic Review*, 2015, *105* (9), 2798–2837.

Blumberg, Stephen J, Karil Bialostosky, William L Hamilton, and Ronette R Briefel, "The effectiveness of a short form of the Household Food Security Scale.," *American Journal of Public Health*, 1999, *89* (8), 1231–1234.

Chabris, Christopher F., David Laibson, and Jonathon P. Schuldt, "Intertemporal choice," in Steven N. Durlauf and Larry Blume, eds., *The New Palgrave Dictionary of Economics*, London: Palgrave Macmillan, 2008.

Chetty, Raj, Adam Looney, and Kory Kroft, "Salience and taxation: Theory and evidence," *American Economic Review*, 2009, *99* (4), 1145–1177.

Cubitt, Robin P. and Daniel Read, "Can intertemporal choice experiments elicit preferences for consumption," *Experimental Economics*, 2007, *10* (4), 369–389.

de Oliveira, Angela and Sarah Jacobson, "(Im)patience by proxy: Making intertemporal decisions for others," Technical Report, Department of Economics, Williams College 2017.

Duflo, Esther, Michael Kremer, and Jonathan Robinson, "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya," *American Economic Review*, 2011, *101* (6), 2350–2390.

Efron, Bradley, "The efficiency of Cox's Likelihood Function for censored data," *Journal of the American Statistical Association*, 1977, *72* (359), 557–565.

Fudenberg, Drew and David K Levine, "A dual-self model of impulse control," *American economic review*, 2006, *96* (5), 1449–1476.

Geiselman, Paula J, Ashley M Anderson, Marcia L Dowdy, David B West, Stephen M Redmann, and Steven R Smith, "Reliability and validity of a macronutrient self-selection paradigm and a food preference questionnaire," *Physiology & behavior*, 1998, *63* (5), 919–928.

Gruber, Jonathan and Botond Kőszegi, "Is addiction "rational"? Theory and evidence," *The Quarterly Journal of Economics*, 2001, *116* (4), 1261–1303.

Gul, Faruk and Wolfgang Pesendorfer, "Temptation and self-control," *Econometrica*, 2001, *69* (6), 1403–1435.

Halevy, Yoram, "Time consistency: Stationarity and time invariance," *Econometrica*, 2015, *83* (1), 335–352.

Herrnstein, Richard J, George F Loewenstein, Drazen Prelec, and William Vaughan, "Utility maximization and melioration: Internalities in individual choice," *Journal of behavioral decision making*, 1993, *6* (3), 149–185.

Kaur, Supreet, Michael Kremer, and Sendhil Mullainathan, "Self-control at work," *Journal of Political Economy*, 2015, *123* (6), 1227–1277.

Laibson, David, "Golden eggs and hyperbolic discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–477.

_ , "Why don't present-biased agents make commitments?," *American Economic Review, Papers and Proceedings*, 2015, *105* (5), 267–272.

O'Donoghue, Ted and Matthew Rabin, "Doing it Now or Later," *American Economic Review*, mar 1999, *89* (1), 103–124.

_ **and** _ , "Optimal sin taxes," *Journal of Public Economics*, 2006, *90* (10), 1825–1849.

Read, Daniel and Barbara Van Leeuwen, "Predicting hunger: The effects of appetite and delay on choice," *Organizational behavior and human decision processes*, 1998, *76* (2), 189–205.

Rees-Jones, Alex and Dmitry Taubinsky, "Heuristic perceptions of the income tax: Evidence and implications for debiasing," Technical Report, National Bureau of Economic Research 2016.

Royer, Heather, Mark Stehr, and Justin Sydnor, "Incentives, commitments, and habit formation in exercise: evidence from a field experiment with workers at a Fortune-500 company," *American Economic Journal: Applied Economics*, 2015, *7* (3), 51–84.

Sayman, Serdar and Ayse Onculer, "An Investigation of Time Inconsistency," *Management Science*, 2009, *55* (3), 470–482.

Sprenger, Charles, "Judging Experimental Evidence on Dynamic Inconsistency," *American Economic Review, Papers and Proceedings*, 2015, *105* (5), 280–285.

Strotz, Robert Henry, "Myopia and inconsistency in dynamic utility maximization," *The Review of Economic Studies*, 1955, *23* (3), 165–180.

Taubinsky, Dmitry and Alex Rees-Jones, "Attention variation and welfare: theory and evidence from a tax salience experiment," *Review of Economic Studies*, forthcoming.

Thaler, Richard H and Hersh M Shefrin, "An economic theory of self-control," *Journal of Political Economy*, 1981, *89* (2), 392–406.

Toussaert, Séverine, "Eliciting temptation and self-control through menu choices: a lab experiment," Technical Report, Working Paper 2015.

Appendix: Not For Publication

A Additional Figures and Tables

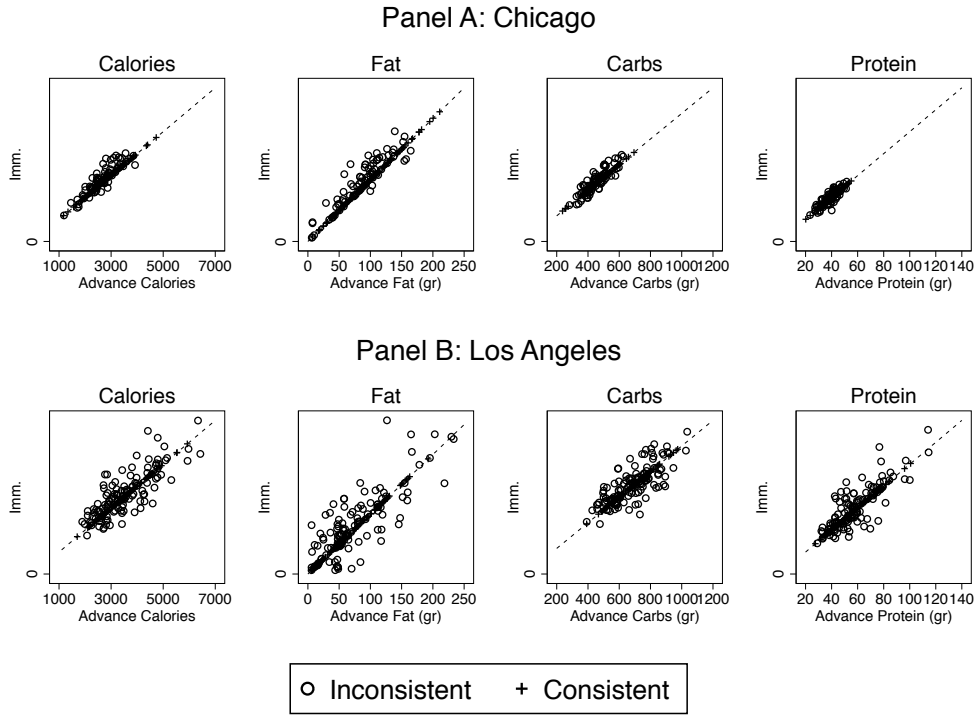


Figure A1: Frequency of Calories, Fat, Carbohydrates and Protein in Advance and Immediate Bundles

Notes: Each participant is represented by a point. Figures include calories, fat grams (1 fat gram = 9 calories), carbohydrate grams (1 carbohydrate gram = 4 calories) and protein grams (1 protein gram = 4 calories). Subjects who choose more of each nutrient in advance versus immediate choice lie below the 45° line. This graph includes a 5% jigger.

Table A1: Nutritional Information

Food	Calories	Fat (g)	Saturated Fat (g)	Carbohydrates (g)	Fiber (g)	Natural Sugar (g)	Added Sugar (g)	Protein (g)
Panel A: Chicago								
3 bananas	405	1.50	0.51	104.16	12.00	55.77	0	4.98
1 cucumber	68	0.50	0.17	16.39	2.30	7.54	0	2.93
2 Granny Smith apples	238	0.78	0.00	56.08	11.60	39.52	0	1.82
2 green peppers	131	1.12	0.38	30.44	11.20	15.74	0	5.64
2 oranges	216	0.55	0.07	54.05	11.00	43.01	0	4.32
2 Bosc pears	294	0.40	0.00	70.52	13.60	44.80	0	1.58
2 plums	120	0.72	0.02	30.16	3.60	26.20	0	1.84
2 Red Delicious apples	250	0.84	0.00	59.62	9.80	44.44	0	1.14
1 red pepper	74	0.72	0.06	14.36	5.00	10.00	0	2.36
1 tomato	33	0.36	0.05	7.08	2.20	4.79	0	1.60
2 bags Cheetos	360	24.75	3.38	29.25	2.25	0.00	0	2.25
1 bag Cheez-Its	210	11.00	2.50	24.00	1.00	0.00	0	5.00
2 bags Doritos	315	18.00	2.25	36.00	2.25	0.00	0	4.50
2 fudge brownies	780	34.00	10.00	112.00	2.00	0.00	62	6.00
2 Honey Buns	680	30.00	16.00	90.00	2.00	0.00	50	10.00
2 bags potato chips	360	22.50	3.38	33.75	2.25	2.25	0	4.50
4 Nutter Butter cookies	250	10.00	2.50	37.00	2.00	0.00	15	4.00
6 Oreo cookies	270	11.00	3.50	41.00	2.00	0.00	23	2.00
1 PayDay bar	240	13.00	2.50	27.00	2.00	0.00	21	7.00
1 Snickers bar	250	12.00	4.50	33.00	1.00	0.00	27	4.00
Panel B: Los Angeles								
16 oz bag baby carrots	159	0.59	0.10	37.38	13.15	21.59	0	2.90
4 bananas	484	1.79	0.61	124.32	14.15	66.56	0	5.93
14.5 oz can tomatoes	86	0	0	13.60	3.29	10.19	0	0.85
2 cucumbers	193	1.41	0.48	46.65	6.42	21.46	0	8.35
4 oz cup diced peaches	81	0	0	19.06	1.02	18.06	0	0.99
2 Gala apples	194	0.41	46.54	7.82	35.28	0	0.85	
12 oz bag salad	82	0	0	16.02	4.08	7.99	0	4.01
2 green apples	140	0.46	0	32.94	6.78	23.21	0	1.06
2 oranges	213	0.54	0.07	53.30	10.89	42.41	0	4.26
16 oz bag red grapes	313	0.73	0.24	82.10	4.08	70.21	0	3.27
3 chocolate chip cookies	163	7.23	3.57	24.06	0.66	0	13.58	1.33
4 oz bag Doritos	577	32.98	6.18	70.10	4.12	0	8.25	
8 oz gelatin cup	865	0	0	205.39	0	0	205.39	10.81
1 palmier	110	6.00	3.99	11.99	0	0	11.99	1.00
1 raspberry roll	157	4.84	1.73	26.89	0.42	0	15.67	1.46
8 oz rice pudding cup	245	4.88	2.74	41.71	0.68	0	26.28	7.32
1 Salvadorian bread	1,496	68.48	18.45	189.92	2.80	0	99.60	28.48
2 sweet buns	462	14.59	2.75	71.04	2.89	0	15.75	11.87
4 oz bag Takis chips	567	30.24	9.45	64.26	7.59	0	0	7.56
6 oz bag tortilla chips	845	37.79	4.72	114.55	7.99	0	2.06	11.25

Notes: Calculations of nutrition based on $1 quantities in study. Natural and added sugar calculated separately for healthy and unhealthy food items.

Table A2: Unambiguous Utility Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	All Subjects			*Inconsistent Subjects*		
	Chicago	Los Angeles	Pooled	Chicago	Los Angeles	Pooled
Fruit/Vegetable	0.004	0.053*	0.211***	-0.049	0.046	0.172***
	(0.049)	(0.030)	(0.028)	(0.090)	(0.049)	(0.044)
Perishable		0.512***			0.436***	
		(0.039)			(0.057)	
Fat	-0.007***	-0.003**	-0.004***	-0.009**	-0.003**	-0.005***
	(0.002)	(0.001)	(0.001)	(0.004)	(0.002)	(0.001)
Carbohydrates	0.001***	0.003***	0.002***	0.001	0.003***	0.002***
	(0.001)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
Protein	0.030***	-0.007*	-0.001	0.034***	-0.008	-0.000
	(0.006)	(0.004)	(0.003)	(0.011)	(0.006)	(0.005)
# Observations	4211	4851	9062	1491	2151	3642
# Rankings	218	256	474	82	121	203
# Clusters	218	171	389	82	95	177
Log-Likelihood	-8777.216	-9850.884	-18737.92	-3026.495	-4259.793	-7325.432

Notes: Rank Order Logit regression results from unambiguous ordering r_U, ignoring all foods ever exchanged. Standard errors clustered on individual level in parentheses. Levels of significance: * 0.10, ** 0.05, *** 0.01.

Table A3: Utility Estimates and Commitment

	(1)	(2)	(3)	(4)	(5)	(6)
	Chicago		*Los Angeles*		*Pooled*	
	Commit = 0	Commit =1	Commit = 0	Commit =1	Commit = 0	Commit =1
Fruit/Vegetable	0.062	0.009	0.052	0.078**	0.073*	0.368***
	(0.061)	(0.079)	(0.071)	(0.034)	(0.042)	(0.035)
Perishable			0.345***	0.543***		
			(0.074)	(0.043)		
Fat	-0.004*	-0.012***	-0.003	-0.002*	-0.006***	-0.003**
	(0.002)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)
Carbohydrates	0.000	0.003***	0.003***	0.003***	0.001***	0.003***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.000)
Protein	0.034***	0.025**	-0.007	-0.008*	0.010**	-0.009**
	(0.008)	(0.011)	(0.008)	(0.004)	(0.005)	(0.004)
Immediate Choice						
X Fruit/Vegetable	-0.080***	-0.061***	-0.116***	-0.031***	-0.065***	-0.035***
	(0.019)	(0.022)	(0.038)	(0.012)	(0.016)	(0.008)
X Perishable			-0.001	-0.011		
			(0.028)	(0.013)		
X Fat	0.000	-0.002*	-0.002	0.000	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)
X Carbohydrates	0.001**	0.000*	-0.001**	0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
X Protein	-0.007*	0.001	0.011**	-0.001	0.000	-0.001
	(0.004)	(0.003)	(0.005)	(0.002)	(0.003)	(0.001)
# Observations	5800	2920	2560	7680	8360	10600
# Rankings	290	146	128	384	418	530
# Clusters	145	73	37	134	182	207
Log-Likelihood	-12267.53	-6164.32	-5370.32	-15921.93	-17674.27	-22219.70
$H_0 : \beta_A(Commit = 0) = \beta_A(Commit = 1)$	$\chi^2(4) = 8.01$ ($p = 0.09$)		$\chi^2(5) = 7.13$ ($p = 0.21$)		$\chi^2(4) = 46.15$ ($p < 0.01$)	
$H_0 : \beta_I(Commit = 0) = \beta_I(Commit = 1)$	$\chi^2(4) = 10.53$ ($p < 0.05$)		$\chi^2(5) = 12.35$ ($p < 0.05$)		$\chi^2(4) = 56.21$ ($p < 0.01$)	
Fraction Inconsistent	0.441	0.247	0.813	0.359	0.555	0.328
ρ(Commit, Inconsistent)	-0.190 ($p < 0.01$)		-0.393 ($p < 0.01$)		-0.228 ($p < 0.01$)	
$V_A(\mathbf{q_A})$	1.398	1.343	3.352	5.111	0.960	3.590
$V_A(\mathbf{q_I})$	1.371	1.347	3.261	5.053	0.924	3.555
$\frac{V_A(\mathbf{q_A}) - V_A(\mathbf{q_I})}{V_A(\mathbf{q_A})}$	0.019	-0.003	0.027	0.011	0.037	0.010
$V_I(\mathbf{q_A})$	0.951	1.088	2.521	4.786	0.646	3.331
$V_I(\mathbf{q_I})$	0.947	1.097	2.476	4.734	0.631	3.299
$\frac{V_I(\mathbf{q_A}) - V_I(\mathbf{q_I})}{V_I(\mathbf{q_A})}$	0.004	-0.008	0.018	0.011	0.023	0.010
$V_U(\mathbf{q_A})$	1.200	1.260	3.295	5.113	0.827	3.555
$V_U(\mathbf{q_I})$	1.184	1.267	3.217	5.054	0.801	3.521
$\frac{V_U(\mathbf{q_A}) - V_U(\mathbf{q_I})}{V_U(\mathbf{q_A})}$	0.013	-0.006	0.024	0.012	0.031	0.010

Notes: Rank Order Logit regression results. Standard errors clustered on individual level in parentheses. Levels of significance: * 0.10, ** 0.05, *** 0.01. Null hypothesis test stationarity of preferences from interacted rank order logit regression of choices on nutritional characteristics with different coefficients for immediate choice. Test corresponds to all interaction terms being equal to zero.

B Experiment Script

B.1 Chicago

Recruitment and Item Selection (Week 1)

Thank you for participating in the store promotion. Only certain items are eligible. To see which items are eligible you should look at this promotion sheet (see Figure 1). Each box is worth $1. Pick 10 items for a basket worth $10/

Delivery Dates

1. Your basket will be delivered in ONE WEEK. Please specify on the back side which dates and times you will be available to receive it. You MUST be at home to get your basket: we cannot leave the basket for you.

2. At the end of the day, we will call you to confirm a delivery date and time

Special Promotion

1. Your $10 basket is FREE OF CHARGE

2. In addition, you will get $20 just for participating in our store promotion and completing the questionnaires. But you MUST BE HOME both times for the basket delivery.

Delivery Confirmation Call (Week 1)

Hi, this is [NAME] from Louis' Groceries. I'm calling for [NAME]. (Or, Is this [NAME]?). You have signed up for the FREE food basket delivery program. I'm just confirming that we have you scheduled to receive the basket of items that you picked out in store on [DATE].

1. Remember, you MUST be home to receive your basket, we are not able to leave it at your door. Does this still work for you? [If not, try to reschedule them within 2 days]

2. Great, we will see you next week on [DATE] between [TIME START] and [TIME END].

Delivery Reminder Call (Week 2)

Hi, this is [NAME] from Louis' Groceries. I'm calling for [NAME]. (Or, Is this [NAME]?). I am calling to remind you that your FREE food basket delivery is scheduled on [DATE] between [TIME START] and [TIME END]. You MUST be home to receive your basket and participate in the promotion that earns you $20 after 2 weeks.

First Delivery and Item Selection for Second Week (Week 2)

Hello, I am here with your basket. Please take a look [Bring open basket, allow person to look through]. We also have some extra items available. If you like, you can exchange

any one item in your basket for one of these items [show extra items on tray]. I brought four additional items, so you can make up to 4 exchanges. Do you want to make any exchange? [Great thanks, let me note that on your order sheet.]

Remember next week you will also get a basket. Here is a Week 2 basket order sheet and the promotion items [hand to person.] Will you please go ahead and fill this out? I will wait in the car and prepare the next round of deliveries while you do that. When you are done, just come outside and we will get your order sheet from you.

Remember:

1. Your delivery will be next [DATE] between [TIME START] and [TIME END].
2. You MUST be home to receive your basket next week
3. Your $10 basket is FREE OF CHARGE
4. In addition you will get $20 just for participating in our store promotion and completing the questionnaires. Next week when I come back and after you complete the questionnaire I will give you a voucher to pick up $20 in store.

Reminder Call and Commitment Elicitation (Week 3)

Hi, this is [NAME] from Louis' Groceries. [I'm calling for [NAME]. Or, Is this [NAME]?]. I am calling to remind you that your FREE food basket delivery is scheduled on [DATE] between [TIME START] and [TIME END]. You MUST be home to receive your basket and participate in the promotion that earns you $20 after 2 weeks.

Last time, we brought some extra items for you so you could exchange if you changed your mind from your previous choices. This time, we can also bring extra items, but I wanted to check if you'd like that or not. It is up to you: would you like me to bring extra items this time, or not?

Second Delivery (Week 3)

Hello, I am here with your basket. Please take a look [Bring open basket, allow person to look through].

[If they wanted an exchange] We also had some extra items from the deliveries, If you like, you can exchange any one item in your basket for one of these items [show extra items on tray]. I brought four additional items, so you can make up to 4 exchanges. Do you want to make any exchange? [Great thanks, let me note that on your order sheet.]

Here is a questionnaire we hope you can fill out about the promotion. After you are done, please bring this questionnaire back to the store to receive your $20 IN CASH just for your participation in the promotion.

B.2 Los Angeles

Recruitment and Item Selection (Week 1)

Do you want to participate in our promotion?

- You will get 3 weeks of free food deliveries valued at $10 each.
- You will get to pick the foods you want from a list.
- You need to be at home to receive your deliveries.
- You will answer a few questions when you sign up and each time you get a delivery.
- You will get a $25 Northgate gift card for completing all the steps.

[Interviewer next records participant name, address, phone number and scheduled date and time of delivery for about one week in the future.]

Thank you for signing up for the promotion! You'll first make food selections for your delivery. Then I'll also ask you a few questions. [Interviewer will place visual aids in front of respondent so they can point to answer or answer verbally. Later in survey, the interviewer will have the survey in front of them, and face the respondent.]

You will now select foods for the FREE basket that will be delivered to your house next week. You are deciding just for next week. Next week, you will decide for the following week and you will do the same the week after for a total of 3 weeks of deliveries. These are the foods that are available. [show food sheet MENU]. The foods will come from Northgate market when possible. Here are the foods. Each item on this list is worth $1 and you can select 10 - for a FREE basket worth $10. You can choose each item more than once. This also tells you how MANY of each item you will get with each $1 order. You can say or point to the items and I will write down what you selected.

Delivery Reminder Call (Week 2, 3 and 4)

Hi, this is [NAME] from the Northgate Delivery Promotion. I'm calling for [NAME]. (Or, Is this [NAME]?). I am calling to remind you that your FREE food basket delivery is scheduled on [DATE] between [TIME START] and [TIME END]. You MUST be home to receive your basket and participate in the promotion that earns you $20 after 2 weeks.

First/Second/Third Delivery and Item Selection for Second/Third Delivery (Week 2 and 3)

Hi, I am from the Northgate Delivery Promotion and am here with a food delivery. Are you [NAME]?

- If Yes, 'Great, may I see an ID just so I can verify that?'
- If No, 'Is [NAME] home? I can only leave the delivery with [NAME].' Arrange to come back at a time when [NAME] is available (either then or by phone later).

Today I will give you your food delivery, you will decide on foods for next week, and then I'll also ask you a few questions. Here is your food delivery [show box]. Please take a look [bring open basket, allow person to look through].

The below is only asked for everyone on their first delivery (Week 2) or anyone who did not commit for the subsequent deliveries (Weeks 3-4).

We also have some extra items available. If you like, you can exchange any one item in your basket for one of these items [show extra items in tray]. I brought all the menu items, and you can make up to 4 exchanges. Do you want to make any exchange? [Great thanks, let me note that on your order sheet/ BE SURE TO RECORD WHAT WAS SWITCHED ON ORDER SHEET AND TAKE BACK ORDER SHEET].

Is [NEXT WEEK SAME DATE] and [NEXT WEEK SAME TIME] still good for you?

- If Yes, 'Great'
- If No, 'I can reschedule for [AROUND SAME TIME +/- 2 days]'

You will now select foods for the FREE basket that will be delivered to your house next week. You are deciding just for next week. Next week, you will decide for the following week for a total of 3 weeks of deliveries. These are the foods that are available. [show food sheet MENU]. The foods will come from Northgate market when possible.

Each item on this list is worth $1 and you can select 10 - for a FREE basket worth $10. You can choose each item more than once. This also tells you how MANY of each item you will get with each $1 order.

You can say or point to the items and I will write down what you selected. Please go ahead and start. [Record below]

Commitment Question (Week 2 for half of participants, Week 3 for all participants)

For this week's delivery, you had the option to change your mind by exchanging items in your basket. This time, you can choose whether you want the option to make exchanges, or whether you want to stick to your pre-ordered choices. It is no trouble for us either way, it is entirely up to you. Do you want to have the option to make exchanges, or do you want to stick to your pre-ordered choices?